SCIENTIFIC AMERICAN

HANG-GLIDER AERODYNAMICS

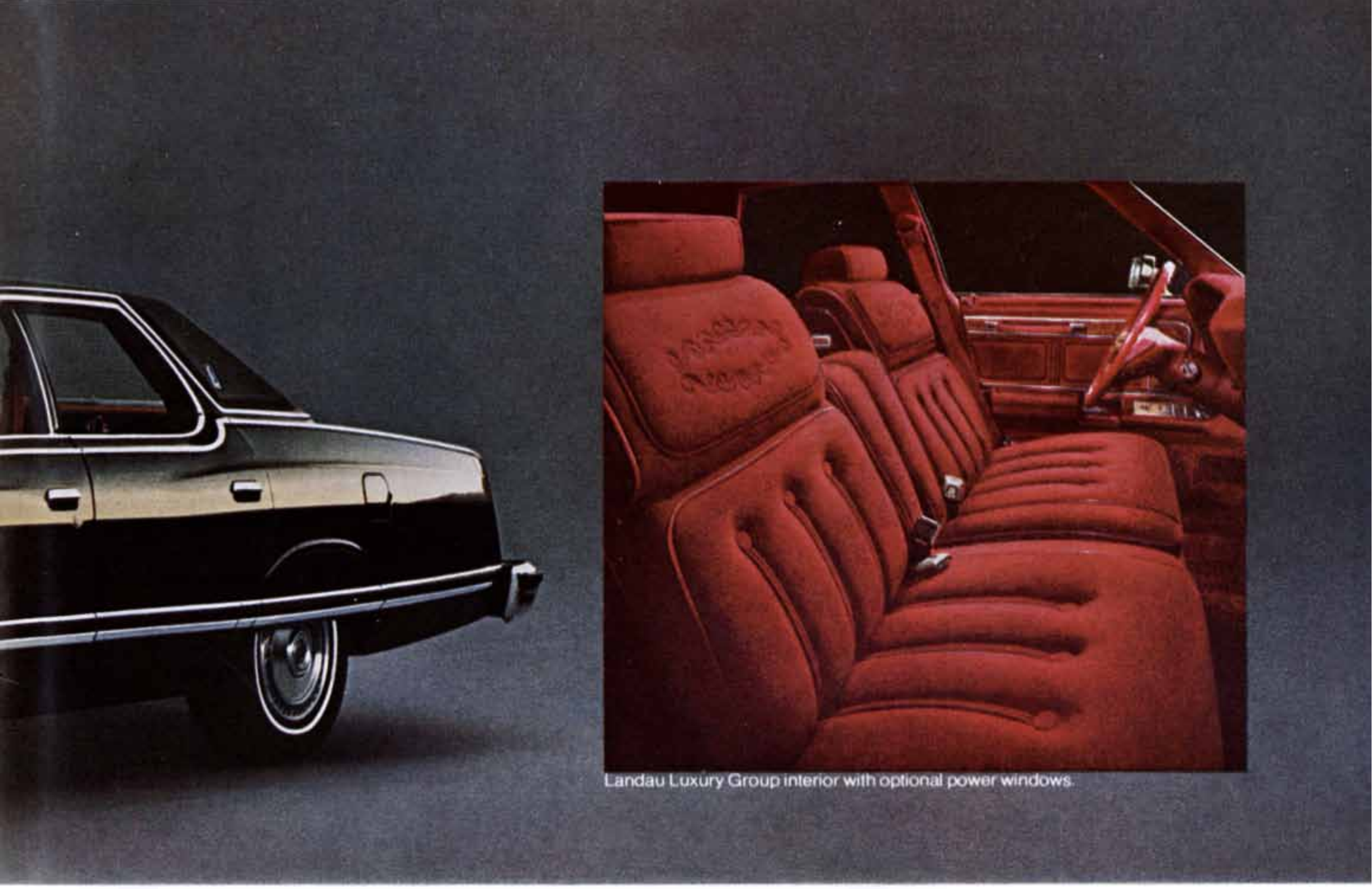
$1.00

December 1974

New Ford LTD Landau 4-Door Pillared Hardtop shown with optional Landau Luxury Group, fender skirts and electric rear window defroster.

Ford introduces two new luxurious standard-size cars. So well-made, Ford challenges you to look close and compare.

1975 Ford LTD with optional vinyl roof, deluxe bumper group and wheel covers and convenience group.



Landau Luxury Group interior with optional power windows.

1975 Ford LTD Landau.

A logical alternative to longer, heavier, more expensive cars like Buick Electra Limited and Oldsmobile 98 Regency.

We think that if you'll not only compare the workmanship in this new automobile but also compare it for size, comfort and luxury, you'll find that you can get features you want from more expensive luxury cars at a sticker price that's hundreds of dollars less.

Of course, the main thing you don't get is hundreds of pounds of extra weight and almost a foot of extra length.

In addition to LTD's standard equipment, here are the features that make LTD Landau so special. Compare them with more expensive cars.

Exterior features

- Hidden headlamps for distinctive styling
- Wide color-keyed vinyl protective side molding
- Front cornering lamps for improved side illumination.

Interior comfort and convenience

- Flight bench seat (with front center armrest)
- Lush knit cloth and vinyl trim
- Padded door panels
- Color-keyed cut-pile carpeting
- Right and left hand remote control mirrors for better visibility
- Automatic parking brake release
- Electric trunk lid release.

Landau Luxury Group option

For those who want the little bit extra in comfort and convenience:

- Split bench seat with passenger recliner
- Upholstery in new 100% nylon luxury cloth that feels and looks like cashmere.
- 22 oz. shag carpeting, stylish and soil resistant
- Front and rear door armrests
- Front and rear door pull straps, front seat-back assist straps for easy access and exit
- Quartz crystal digital clock for precise accuracy
- And more.

1975 Ford LTD.

Traditional LTD quality, value priced.

"LTD Quality" is synonymous with a well-made, comfortable, full-size car. For 1975, the newly styled Ford LTD comes to you as the lowest-priced full-size Ford while retaining a high level of LTD quality.

Consider LTD's reputation for quiet ride. Consider its resale value. Consider its steel-belted radials and solid-state ignition. Consider the peace of mind you get from its solid construction.

Designed and sticker-priced as it is, we think, it offers a great opportunity to buy a lot of automobile at a value price.

What you get for your money.

An impressive list of standard equipment and features.

For basic value

- Select Shift Cruise-O-Matic transmission
- 351 CID V-8 engine
- Power steering
- Power front disc brakes
- Power ventilation

For economy of operation

- Solid-state ignition
- Steel-belted radials

For luxury details that make an LTD an LTD

- Brocade cloth upholstery trim
- Color-keyed cut-pile carpeting
- And more.

Everyone says compare–we tell you how.

The Ford "Closer You Look Book" (available at your Ford Dealer) tells you how and what to look for in a well-made, durable car. And whether you examine the standard LTD, the LTD Brougham or the luxurious LTD Landau, we think you'll find a car that measures up in workmanship and detail.

FORD LTD

FORD DIVISION

Freedom for the Press.

The Harris 2500 Newspaper System and Phototypesetters

American newspapers have long been caught between 19th century production methods and 20th century costs. Now, Harris electronic video typewriters, editing terminals, and high-speed phototypesetters are helping to set them free.

Write: Harris Corporation,
55 Public Square, Cleveland, Ohio 44113

HARRIS
COMMUNICATIONS AND INFORMATION HANDLING

Established 1845 **SCIENTIFIC AMERICAN** December 1974 Volume 231 Number 6

ARTICLES

DEPARTMENTS

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $12 PER YEAR, U.S., POSSESSIONS AND CANADA; $15 PER YEAR, ALL OTHER COUNTRIES. EUROPEAN SUBSCRIPTION COPIES MAILED C/O PUBLICATIONS DISTRIBUTION SERVICE, ST. KATELIJNEVEST 14, B2000 ANTWERP, BELGIË.

THE COVER

The painting on the cover portrays a "hang" glider, which enables its pilot to fly with artificial wings, as Icarus did, but to do so with considerably more satisfactory results. (Icarus flew too close to the sun, so that the wax in the wings of his craft melted and he plunged to his death in the sea.) The craft on the cover was designed and built by Michael A. Markowski, an aerospace engineer of Marlboro, Mass. (see "The Amateur Scientist," page 138). Its principal features are a specially designed wing made of Dacron, a tubular aluminum frame, a control bar in the shape of an inverted Y and the harness from which the pilot hangs. On a typical flight the pilot takes off by running down a slope into the wind. With suitable wind conditions and sufficient flying skill he can stay aloft for a considerable period of time, making turns of as much as 180 degrees. The craft is named EAGLE-III.

THE ILLUSTRATIONS

Cover painting by Ted Lodigensky

Page	Source
23	*The New York Times* (*left and right*), New York Police Department (*center*)
24	Robert Buckhout (*top*), Graphic Presentation Services, Inc. (*bottom*)
25–27	Graphic Presentation Services, Inc.
28–29	Mark Greenwald
30–31	Graphic Presentation Services, Inc.
33–34	George V. Kelvin
35	Hale Observatories
36–43	George V. Kelvin
45	Alfred Gierer, Max Planck Institute for Virus Research, Tübingen
46–51	Tom Prentiss
52	Graphic Presentation Services, Inc.
53	Tom Prentiss
54	Graphic Presentation Services, Inc.
69	Fritz Goro
70	D. J. Paolillo, University of Illinois (*top*); R. Chollet and D. J. Paolillo, University of Illinois (*bottom left*); C. J. Arntzen, University of Illinois (*bottom right*)
71–80	Allen Beechel
89	Merton C. Flemings
90–94	Dan Todd
95	Ben Rose
96–97	Ralph Morse
98–101	Bunji Tagawa
102	Ralph Morse (*top*), Bunji Tagawa (*bottom*)
103–106	Bunji Tagawa
109–114	Gabor Kiss
115	European Organization for Nuclear Research (*top*), Gabor Kiss (*bottom*)
116–119	Gabor Kiss
120	Chase Manhattan Bank Numismatic Collection and American Numismatic Society
122	Ilil Arbel
123–127	George C. Boon, National Museum of Wales
128	George C. Boon (*top*), Ilil Arbel (*bottom*)
129	Ilil Arbel
130	George C. Boon
133	New York Public Library
134–135	Edward Bell
136	Graphic Presentation Services, Inc.
138–139	Michael A. Markowski
140–143	Jerome Kuhl

What would you do with a Statistics and Number-Crunching omputer that starts at $7,400,* has 16K Hardwired Basic Language and 28 Major Peripherals?

Plenty!

The new Wang System 2200 is a System. It gives you the raw power and the peripherals you must have for a wide range of problem solving. For under $7,500 you get a CPU with 16K bytes of BASIC language instructions hardwired into the electronics . . . plus a 4K operating memory. You also get a big 16 lines (of 64 characters each) CRT display, a console mag tape drive and your choice of either alpha or BASIC Keyword keyboards.

Some Words About Language: The hardwired MOS ROM language in your System 2200 finally ends your dollar tradeoffs . . . economy systems that are costly to program or very expensive systems that are relatively easy to program. Many, if not most, of your people already know BASIC. They'll be solving problems the day your system is delivered (and, we can deliver in about two weeks). Most of your budget will go into problem solving; not system support.

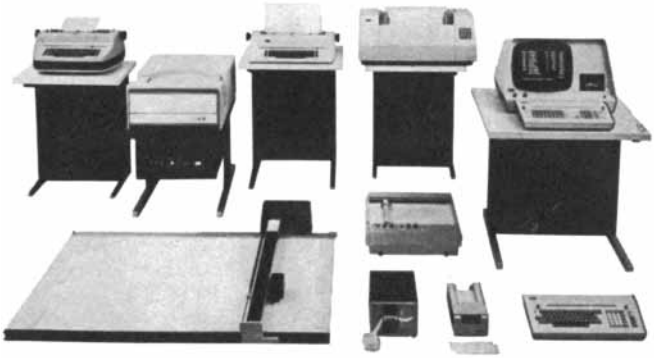
Try To Out-Grow It: Main memory is field expandable in 4K increments (at $1,600 per 4K). Up to 32K. You can choose from three kinds (and 7 price ranges) of printers . . . one even has a stepping motor for very precise 4-quadrant incremental plotting. Speaking of plots, we have a new, very large flatbed (31" × 48") for only $8,000 or a smaller one if you plot small. Both print alphanumerics and plot under full program control. Been appalled lately by disk prices? Starting at just $4,500, we offer you our new "floppy" disk in single, double and triple disk configurations (.25, .50 and .75 MB's). For big disk power, you can have 1, 2 or 5 megabyte fixed/removable disk systems. All peripherals, including punched or mark sense hopper card readers, paper tape readers and on- line BCD or ASCII controllers are easily added-on in the field so your System 2200 will grow with your needs.

The Wise Terminal: If you are now or may soon be getting into terminals, we have several new products that will instantly upgrade your System 2200 for telecommunications with any other System 2200 or a mainframe computer. And, you still have a powerful stand-alone system. Another approach, of course, is to justify it as a powerful terminal and get a "free" stand-alone computer. Wise?

We Do A Lot For You: System 2200 is backed by over 250 factory-trained Wang Service Technicians in 105 U.S. cities. Naturally, we guarantee or warranty everything you buy from us. If you want, there are free programming/operating schools here in Tewksbury, Massachusetts, almost every week. We have a growing program library on a wide range of statistics and math/science applications. Our user group (with the unlikely name of "SWAP") could help you cut programming costs even further. We do a lot for you.

* All prices U.S. List. If you're the entrepreneur type, we've just announced a new 7-module Basic Accounting System software package for the business end of your business like payroll, invoicing, inventory, receivables and some really fancy management reports.

WANG



Even if you call the Wang System 2200 a small system . . . you have to admit it's a big idea.

☐ Please send literature

☐ Please have your representative phone me:________

Name________________ Title________

Company________________

Address________________

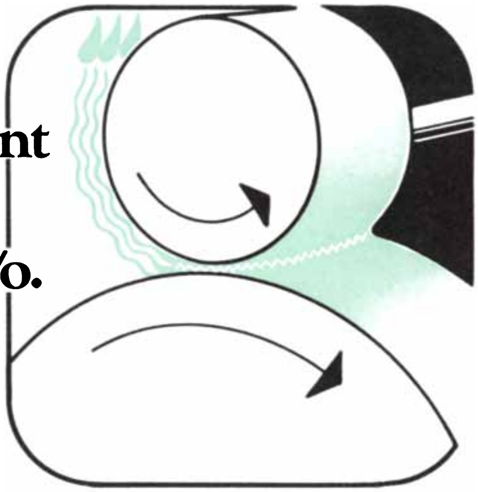
City________________

State________________ Zip________

Wang Laboratories, Inc., SA-12

836 North St., Tewksbury, Ma. 01876, Tel. (617) 851-4111

SANTOTRAC: the new synthetic lubricant that increases traction coefficients 35%-50%.

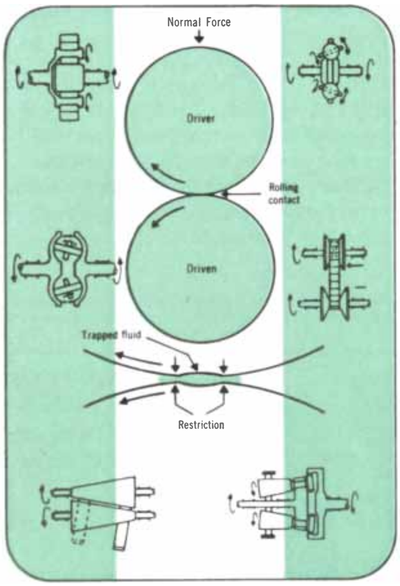


Astute mechanical engineers are warming up their pocket calculators and ticking off computer runs to take a fresh look at power transmission and control with traction drives. Something new has entered the picture: an odd-ball cycloaliphatic molecule. It is an excellent lubricant – with grip. Patented by Monsanto and trademarked Santotrac,® it gives lubrication a new dimension. It's a unique engineering tool that is putting a new gleam in the eye of design engineers. And forming storm clouds on the horizons of variable speed electric motors, complex gear boxes and conventional hydraulic drives. Because it's a lubricant with grip it makes traction drives work better, longer, and handle high loads never before possible.

Traction Drives — Virtues & Problems

Early in any first course on power transmission devices, a budding machinery engineer is introduced to traction drives. Basically, they transmit power between two (or more) – driving and driven – smooth rolling surfaces. These rolling elements are held in contact by a force exerted normal to the relatively miniscule area of contact. The greater the traction at that point of rolling contact, the higher the power transfer efficiency. Obviously, metal-to-metal contact is verboten. Lubrication is essential. So the power transmitted from the driver roller to the driven boils down to: how much force can be applied at the point of contact and how does the lubricant there handle its job. It must protect the metal, fight slip, and not come apart under the heat and pressure.

A baker's dozen of ingenuous designs have been invented to step-up, step-down transmitted power and speed, but the basic elements of a traction drive are shown in the illustration.



The basic elements of a traction drive. The principle of the driving force transmitted through smooth contact with driven elements prevail in all geometries for fixed or variable speeds.

The virtues of traction drives are many: stepless, continuous speed changes made at full load: precise speed settings that change negligibly with load variations and take little power to change, plus rapid speed shifts limited only by external inertial loads. Add to these, noiseless operation even at rolling speeds over 500 ft. per sec., and lower cost with components cheaper than complex gears. The uninitiate's first idea pops: ideal for automobiles! But the mechanical engineer sees much farther: textile machinery drives, variable load compressors, printing presses, shoe machinery. . . In fact, traction drives are contenders for any power system that requires varying levels of speed or torque power.

But there's an IF. The if, up to now, has been the lubricant at the driver-driven contact point.

'Santotrac' Works Here

Glance at the arrow tagged Normal Force in the schematic illustration, the pressure that holds the driver against the drivee. The greater this force, the higher the efficiency in passing on the power. Obviously, to forestall wear and metal fatigue, the contact point cannot be metal-to-metal, nor pressured high enough to cause plastic flow. This contact point must be lubricated, and there's the rub – or lack of it. The slip and wear at this contact point is the stumbling block. Up to now, it hobbled widespread use of traction drives. What's needed at that tiny, hot, high pressure, enormous shear-stress point is a lubricant that doesn't slip! – a protective, high-strength film with drag. To up the power transmission efficiency, the contact point must have a film that tremendous pressure (Hertz Stress) cannot rupture, high shear can't degrade, intense heat won't decompose.

Enter Santotrac.

Up to this point, the best traction drive lubricants were naphthenic oils. The traction coefficient was 5.5 to 6.5%. Santotrac weighed in offering a traction coefficient of 9.6% or higher! Right off, this meant a power transmission boost for any existing traction drive on the order of 40 to 80%. Moreover, the pure and stable synthetic hydrocarbon shrugs off heat, works at 300°F, and takes temperatures that start carbonization in most petroleum stocks.

'Santotrac' Becomes A Transitory Solid

To mechanical engineers, the physical phenomenon of the Santotrac lubricant's action at a rolling point of contact, figuratively speaking, is nothing less than mouth-watering. The higher the pressure, the more viscous the fluid film becomes. At pressures in the range of 200 to 400,000 psi – the film virtually turns into a glassy solid! But only for the fleeting millionth-of-a-second under pressure. Then it instantly reverts to its normal viscosity at that particular temperature. There appears to be no practical limit to the contact pressure it withstands or resistance to degradation from shear stress. Tests show the film still separates metal-to-metal contact and maintains its integrity at 500,000 psi. And piling-it-on performance-wise, despite giving excellent lubrication the traction coefficient of Santotrac drops negligibly at rolling speeds between 1 to 10,000 ft./min. The relative slip with Santotrac compared to mineral oils is in the range 50 to 75% less at a given torque transmission. In a ball and disk traction drive with two rolling surfaces operating at an output torque of 10 ft./lbs. – the total slip through the transmission using a standard naphthenic oil is 4.0%; with Santotrac 50, slip drops to 1.0%. To quantify this improved margin – when slip between individual rolling contact surfaces exceeds about 2%, the driven member grinds to a halt.

The horizon this performance opens before machinery designers is a glowing blue indeed. Long-wearing traction drives can now be built to attain power transmission efficiencies in the range of 85 to 95%. Noiseless traction drives can offer higher shaft speeds that are immune to shock, with momentary overloads absorbed by slip. In short, the continuous power flow of the more efficient traction drive can provide the best match of torque/speed requirements for turbine and rotary engine power sources. It is this happy marriage that promises a gearless car-of-the-future.

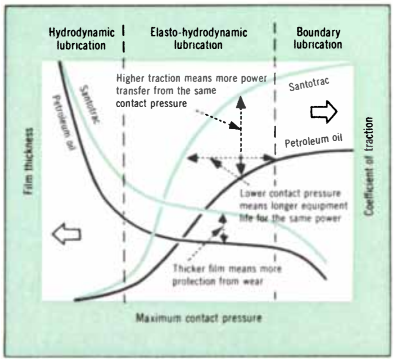
Smaller, Quieter, Cheaper

The size of a traction drive is inversely proportional to the traction coefficient. As traction coefficient goes up, size can come down. While a multitude of new designs will doubtless evolve to supplant variable-speed electric motors; gear-, chain-, and belt drives; and fluid power trains – the odd-ball molecule can add 50 to 100% more load capacity to traction drives currently operating. Plus greater power transmission efficiency and longer wear life. The basic compound is non-corrosive to metal, compatible with paints and seals. It has been formulated to three levels of liquid viscosity and to a high temperature and multi-purpose NLGI #2 grease (with a traction coefficient of 11.5%).

Excluding traction drives, the potential value in lubricating a variety of other equipment with Santotrac – from railroad car journal boxes to factory machinery running on ball- or roller bearings – is intriguing. All rolling contact components have a use-life limited by fatigue failure. Microscopic cracks are the precursor of break-down from cyclic stressing of the contact areas. In laboratory wear tests, Santotrac delivers a bearing wear-life approximately double that of other lubricants of like viscosity. Monsanto researchers hypothesize this as the reason: the chameleon lubricant must fill the high stress, hairline cracks, solidify under the high contact pressure and inhibit the crack propagation. That could explain why bearings last twice as long.

As with any multi-faceted technological development – commercial evaluation is a trackless forest. The fluid grip, the lubrication values, and bearing preservation of Santotrac lubricants are significant in thousands of applications, in scores of industries. They increase coefficient of friction in clutch packs and one-way roller clutches, raise the load capacity and extend wear life of linear ball bearings, improve performance of high-force positive roller feeds on machine tool drives. As a wire rope lubricant, they even reduce slip on elevator drives.

Santotrac is not cheap – but as a life-time single fill or the missing element that permits an engineering advance it is a bargain. A relatively high cost per pound lubricant is positively prudent when weighed economically against the alternatives.



The application range for the lubricant-with-grip is wide indeed. For more idea-generating information and specific data, request a copy of our new bulletin: *Santotrac Lubricants.*

To receive a copy of Santotrac Lubricants, address inquiries to:

F. H. Langenfeld
Monsanto Co.
800 N. Lindbergh Blvd.
St. Louis, Missouri 63166
Telephone: (314) 694-2523

Monsanto
the science company.

Problems can be opportunities in disguise.

The twin dilemmas of energy and the environment may, in the long run, provide man with his best chance to reassess some old assumptions and reorder his priorities for a safer, more pleasant life.

Even before these problems reached crisis proportions, Chiyoda—Japan's leading engineering firm—was developing technology that will help ensure both environmental integrity and efficient energy production.

For example, to minimize air pollution and alleviate the shortage of low-sulfur fuels, Chiyoda has built about half the fuel oil desulfurizers in Japan. We've developed an indoor fume controller to collect the most harmful fume, consisting of particles of less than 3.0 microns. And we've introduced the Chiyoda THOROUGHBRED series of pollution control systems: new processes for flue gas desulfurization, the simultaneous removal of sulfur dioxide and nitrogen oxides, and waste water treatment.

These days "technology assessment" has become a fashionable expression. In our fields, we've been practicing it for years.

RESPONSIBLE, GLOBAL ENGINEERING

CHIYODA CHEMICAL ENGINEERING & CONSTRUCTION CO.,LTD.

Yokohama, Japan. Telex: CHIYO J47726

Overseas Offices: U.S.A.: Chiyoda International Corporation 1300 Park Place Building, 1200 6th Ave., Seattle, Wash. 98101 **England: Chiyoda Representative Office** 16th Floor, St. Alphage House, Fore Street, London E.C. 2 **Singapore: Chiyoda Singapore (Pte.) Ltd.** 8-E, Chia Ping Road, Jurong Town, Singapore 22 **Iran: Chiyoda Iran Co. (Private)** No. 56 Takhte Jamshid Avenue, Teheran

LETTERS

Sirs:

Urie Bronfenbrenner, in his article "The Origins of Alienation" [Scientific American, August], states that today the number of children living in single-parent homes is almost twice as many as a decade ago, and that the number of working mothers has increased so much that a third of the mothers of preschool children are now employed outside the home. He appears to attribute the alienation of our young people to their lessened contact with their parents, or any other significant adult.

However, the evidence of this alienation which he cites in the article, namely the significant increase in the arrest rate amongst teenagers for drug abuse and for juvenile delinquency is based on statistics on the young people who were preschoolers a decade ago or more. Many of these young people were born during the postwar baby boom of the 1950's to two-parent families where the mother was a full-time homemaker.

Also the experience of the kibbutzim in Israel, where children are raised in community nursery schools that are more isolated from their parents than in this country indicates there is less rather than more juvenile delinquency. The problems with drugs that occur in Israel are usually with young people from other countries.

Betty Siegel

Detroit

Sirs:

Dr. Bronfenbrenner explores changes that have taken place in the American family, without recognizing the most disruptive force present in the home: television. He refers to it only in the context of physical environment.

The causes of juvenile delinquency remained fairly constant until this powerful new medium was introduced into American homes. Because of the emphasis on crime, television has become a prime cause of the enormous increase in juvenile delinquency and crime....

To "rebuild and revitalize the social context" requires that we stop permitting the constant image of violence that is before our children's eyes. Since it is unrealistic to expect parents, themselves unaware or absent, to turn off the sets, the broadcasters of crime for entertainment must be pressured to stop their misuse of the public airwaves. This will happen faster if people like Dr. Bronfenbrenner, who study the results and search for causes, will sit down for a few days and watch the images that are brainwashing the children....

Elizabeth Livingston

Fort Myers, Fla.

Sirs:

Urie Bronfenbrenner's article is an absurd example of historical ignorance. Anyone who has read Dickens—not to mention Mayhew or Victor Moberg—knows that the material, social and psychological condition of the poor has improved immensely in the last 100 years in America.

I do not wish to suggest that we live in Utopia and, in fact, strongly favor attempts to improve the quality of life from its present often unsatisfactory state. But please, let us not distort the case! This is not the Apocalypse.

Brucia Witthoft

Framingham State College
Framingham, Mass.

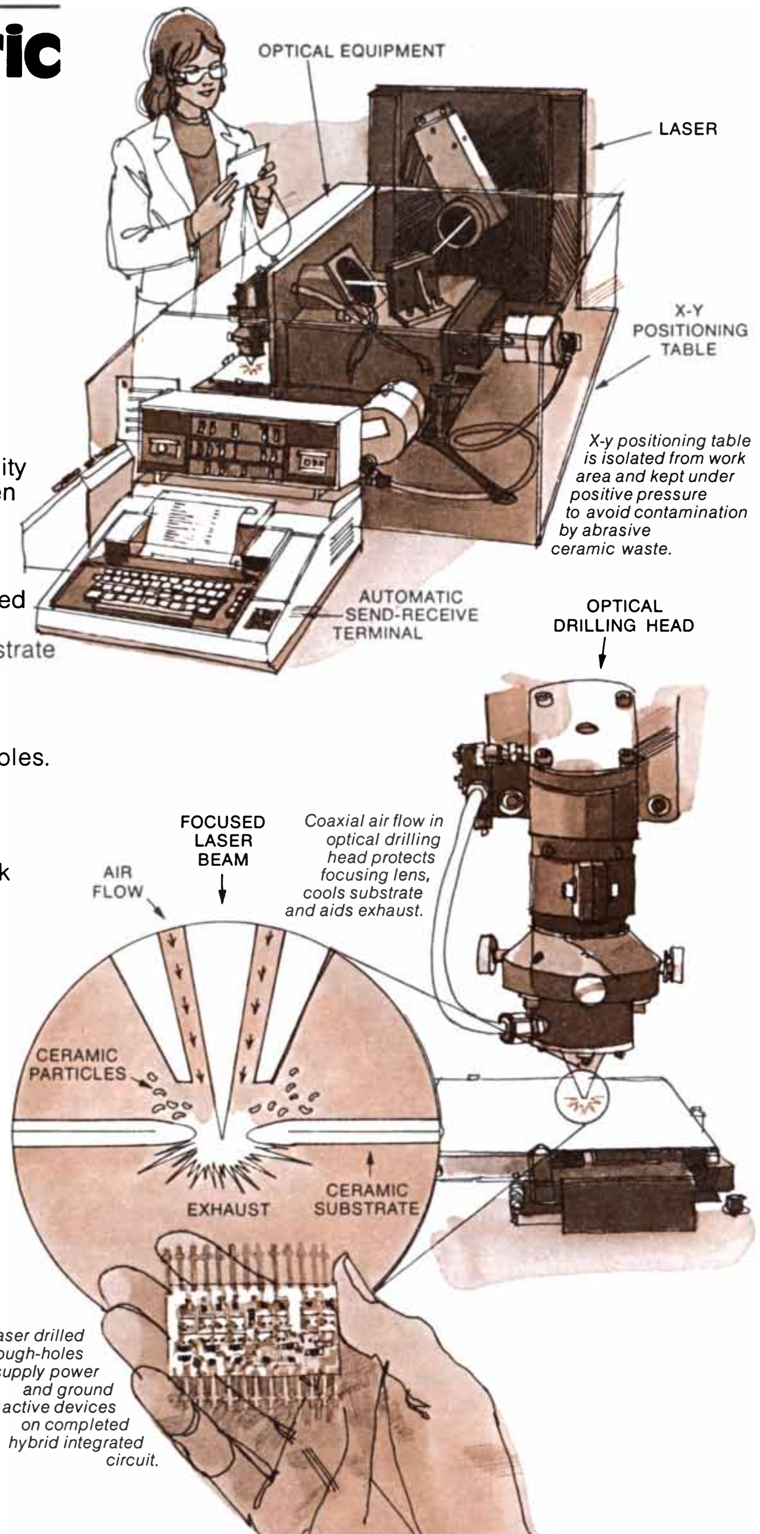
Sirs:

Ms. Siegel appears to assume that parental absence affects the child only during his early years. Studies of juvenile delinquency, drug abuse, school drop-outs and other forms of alienated behavior reveal, however, that a disproportionate number of teenagers engaging in such acts come from homes in which one or the other parent is missing or frequently absent *during the adolescent years.*

At such later ages parental absence is even more pronounced than in the preschool years. Thus in 1973, compared to 12 percent for youngsters under six, 15 percent of all family-reared children between six and 18 were living in single-parent families. According to figures released after my article went to press, by 1974 the percentage of all children under 18 raised by only one parent reached an all-time high of 16 percent (one out of every six children), and the gain over the past four years was twice that for the entire preceding decade (1960-1970). With respect to Ms. Siegel's second point, the popular impression that children in the kibbutz are not raised by their parents and have minimal contact with them is

Western Electric Reports:

Laser drilling. We do it with mirrors.



Thin-film circuit boards in high capacity telephone transmission systems often require hundreds of connections to power and ground sources.

Plated through-holes have proven an efficient way to make these connections. Coated with conducting material, they connect the circuitry carried on one side of a ceramic substrate with power and ground on the other side.

These holes could be punched in the ceramic before it is fired. But shrinkage during firing can move the positions of the holes.

And because of component density, the precise placement of each hole is critical. It can't be more than two mils off.

Engineers at Western Electric's Merrimack Valley Works in Massachusetts recently developed a high-speed method of drilling these holes *after firing* by using a conventional CO_2 laser.

A complex of mirrors on an x-y positioning table is shifted to play the laser beam across a stationary ceramic substrate in a predetermined pattern. The mirrors direct the beam from the laser head enclosure to the positioning table and manipulate it in the x-y axes. An optical drilling head coupled to the table focuses the beam onto the ceramic. The system is controlled by a mini computer coupled with an automatic send-receive terminal. Pattern storage on a cassette tape allows easy changeover and storage.

Benefit: Laser drilling of ceramic substrates after firing has greatly improved positioning accuracy of plated through-holes. And computer controlled laser drilling has doubled the production rate over conventional laser systems—up to five holes a second in closely spaced patterns.

Western Electric

We make things that bring people closer.

grossly in error. The importance accorded to the parental role is reflected in the fact that on the kibbutz all work stops in the late afternoon to enable parents to be at home in order to engage in activities with their children. Children are also frequently involved with their parents and other adults as they visit and take active part in the work of the kibbutz.

The effect of such living patterns is reflected in systematic research. For example, for the past four years, together with colleagues at Cornell University and the University of Tel Aviv, I have been conducting a research project comparing child-rearing among kibbutz and conventional families in Israel. In a large sample of 12-year-olds, kibbutz youngsters reported spending an average of two hours a day in some activity with both of their parents, a figure about equal to that for home-reared children in Israel and appreciably higher than the average for their American age-mates.

Ms. Siegel is correct in saying that there is less juvenile delinquency and drug use in Israel than in the U.S., and our results suggest that greater contact with parents and other adults in the former setting may be one of the factors contributing to this difference.

In reply to Ms. Livingston's letter, I would agree that the growth of television has contributed significantly to changes in American family life and has become a major "disruptive force present in the home," but for reasons that go beyond the impact of television violence. In an earlier version of my article I had included the following passage:

"And even when the parent is at home, a compelling force cuts off communication and response among the family members. Like the sorcerer of old, the television set casts its magic spell, freezing speech and action and turning the living into silent statues so long as the enchantment lasts. For example, one study reports that 78 percent of viewers indicated no conversation while the set was on (except briefly during commercials). The primary danger of the television screen lies not so much in the behavior it produces as the behavior it prevents—the talks, the games, the family festivities and arguments through which much of the child's learning takes place and his character is formed. Turning on the television set can turn off the process that transforms children into people."

In response to Ms. Witthoft's letter, as one who has read Dickens but not Victor Moberg or Mayhew (three volumes was too much), I call on Mr. Bumble as a witness. Oliver has just escaped the workhouse and its starvation diet to become an apprentice to Mr. Sowerberry, the coffinmaker. He now has something to eat, but, exploited and abused, he fights back in revolt.

" 'Oh, you know, Mr. Bumble, he must be mad,' said Mrs. Sowerberry.

" 'It's not Madness, Ma'm,' replied Mr. Bumble, after a few moments of deep meditation. 'It's Meat.'

" 'What?' exclaimed Mrs. Sowerberry.

" 'Meat, M'am, meat,' replied Bumble, with stern emphasis. 'You've overfed him, Ma'm. You've raised a artificial soul and spirit in him, Ma'm, unbecoming a person of his condition.... What have paupers to do with soul or spirit? It's quite enough that we let 'em have live bodies. If you had kept the boy on gruel, Ma'm, this would never have happened.'

" 'Dear, dear!' ejaculated Mrs. Sowerberry, piously raising her eyes to the kitchen ceiling: 'This comes of being liberal!' "

I plead guilty to the charge.

URIE BRONFENBRENNER

Cornell University
Ithaca, N.Y.

Scientific American, December, 1974; Vol. 231, No. 6. 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 Harry T. Morris, 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 30 cents.

Change of address (or other subscription correspondence) should be addressed to Circulation Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Please notify us four weeks in advance of change. If possible, kindly furnish an address imprint from a recent issue. Send both old and new addresses or complete and mail Post Office Form 3578.

☐ Please enter change of address.

☐ Please enter subscription.
U.S. and Canada:
☐ 1 year, $12 ☐ 2 years, $22 ☐ 3 years, $30
All other countries:
☐ 1 year, $15 ☐ 2 years, $27 ☐ 3 years, $37

NAME

NEW ADDRESS

OLD ADDRESS

To start with the obvious, consider the engine. It's a double overhead camshaft six, with a transistorized ignition system and a Solex compound carburetor. That engineering bulletin translates into the following:

• An overhead camshaft configuration is a sophisticated design which requires fewer moving parts and allows more precise valve operation than traditional engines. Result: Better durability and smoother performance.

• The transistorized ignition system helps reduce stress and wear on contact breaker points. Result: Better starting performance.

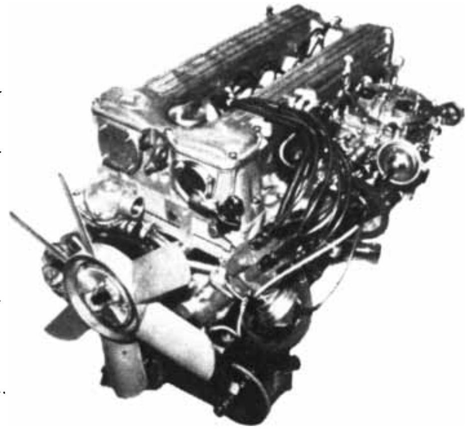
• The compound carburetor provides two-stage optimum fuel/air mixture at all speeds. Result: Good fuel mileage.

This efficient power plant is complemented by a 4-speed automatic transmission with torque converter—and by an aerodynamic silhouette and steel-belted radial tires.

Lessons from a classic

In every detail save power train, the new 280S is the direct descendant of the now classic 450SE Sedans.

The examples abound: Suspension is fully independent. Handling



The Mercedes-Benz double overhead camshaft six. Efficient power plant of the new 280S.

is simply superb, with extraordinary directional stability. You get quick response and surefooted road feel. You are in control—and you know that you are.

The power-assisted disc brakes on all four wheels sweep an area of 451 square inches. You stop again and again smoothly, securely and without disconcerting fade.

The all-welded, unitized body is almost rattle-proof. And enormously strong. The roof alone can stand up to five tons of pressure.

The complete sedan

The 280S is fully equipped. Quite literally, it's a *complete* automobile.

A full list of the features of the 280S is nothing short of an engineering tour de force. Key among these features: Self-clearing side windows and rear lights. Cruise control. Closed-circuit fuel system to avoid fuel evaporation. A central locking system that, at the turn of a key, locks all four doors...and the trunk...*and* the gas filler port. Anatomically designed seats. A multipurpose lever that lets you control the turn signals...and high and low beams...*and* the windshield wipers all without taking a hand off the wheel.

Safety first

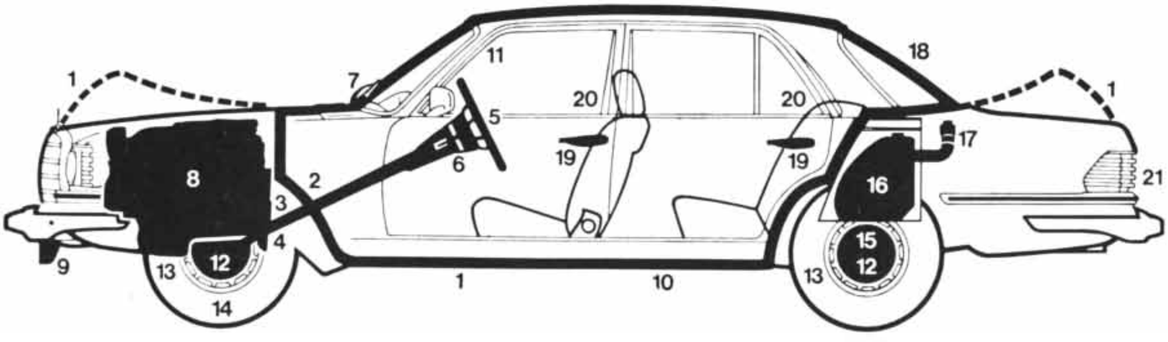
The total Mercedes-Benz commitment to safety is abundantly evident in the new 280S. The safety elements designed into the 280S represent the latest Mercedes-Benz technology. They derive from decades of meticulous research and development.

By any measurement—efficiency, safety, craftsmanship, 6-cylinder economy—the new 280S is a car well worth your close attention.

Contact your local Mercedes-Benz Dealer and schedule a test drive today.

Mercedes-Benz

Engineered like no other car in the world.



Significant safety developments in the 280S. All are standard equipment.

1. Crushable extremities/rigid passenger cell
2. Collapsible steering column
3. Variable ratio, servo-assisted steering
4. Steering box location
5. Safety steering wheel
6. Deformable steering hub
7. Articulated windshield wipers
8. Double overhead camshaft engine, 4-speed auto-transmission
9. Halogen fog lights
10. All-welded construction
11. Aerodynamic rain channels
12. 4-wheel disc brakes
13. Radial tires
14. Zero steering offset
15. Independent suspension
16. Steel-encased fuel tank
17. Side safety filler neck
18. Self-clearing rear glass
19. Safety cone door locks
20. Fully padded interior
21. Self-clearing rear lights

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

DECEMBER, 1924: "There is suspicion that our resources of coal and oil are being wasted, that their exhaustion is no more than around the corner of the next century, and that our civilization is threatened, in consequence, with an early and disastrous end. There is talk of possible power from the waves and the winds, and of the stores of power that some scientists believe may be obtainable from the atoms of matter. Before considering these alternatives, however, we must first take stock of what resources we have. That was done during the past summer at the World Power Conference in London. Hundreds of engineers and experts met to report on the power resources of their respective countries. The conference resulted, among other things, in the most complete and accurate survey of world power resources ever made. The total coal of the world, for example, was found to aggregate 7,397,553,000,000 metric tons. At the present rate of use this is enough for some 4,000 years. Of course, the rate of use is increasing."

"It is a common charge that American aviation, while successful in securing speed records and developing fighting qualities, has lagged seriously behind Europe in the perfection of commercial craft. The charge is not really justified. At a recent air meet in Dayton, Ohio, more than 200 airplanes flew in from other airports carrying visitors to the fair. There were only two or three forced landings. A Stout airplane flew from Detroit with a pilot, mechanic and six passengers. Furthermore, there is no doubt that speed records can be still further increased. At the beginning of the war the maximum speed of airplanes was about 100 miles per hour; last year at St. Louis, Lieutenant Al Williams flew at a speed of more than four miles a minute over a 120-mile course, and records of 300 miles per hour or even more seem quite within our reach."

"With the microscopic study of the cell and its constituent parts and the rediscovery of Mendel's law in 1900, two formerly independent lines of investigation have converged strikingly and have reinforced each other. We now know that living cells are not mere drops of jelly; each contains a complexity of parts, one of which is the easily dyed speck called the nucleus. The nucleus in turn contains a number of highly important microscopic constituents called chromosomes, of which each species of plant or animal possesses a characteristic number. Although the nature and behavior of these little chromosome particles are not fully understood, it is quite evident that they have very much if not everything to do with determining the great facts of heredity."

"During the past year there have been two vigorous attacks on the Einstein Theory of Relativity, one by Professor T. J. J. See, the other by Professor Charles L. Poor. Professor See's attack can be easily dismissed. He believes that Einstein's mathematical calculations are wrong; in fact it is See who is wrong. The argument of Professor Poor is a much stronger one. He disputes two of the famous proofs of the Einstein theory: the deviation of light rays that pass close to the sun and the explanation of the slight irregularity in the motion of the planet Mercury. Professor Poor points out that the deviation of light rays past the sun is a very difficult matter to measure with accuracy, and that the data obtained are susceptible to other interpretations. As to the irregularity of Mercury, he suggests that it is equally well explained by the assumption of some distributed matter in space, an assumption for which there is much other evidence. Neither professor considers, however, that evidence for the essential correctness of the theory lies not only in astronomy but also in atomic physics, where the successes of the theory show its greatest strength."

SCIENTIFIC AMERICAN

DECEMBER, 1874: "With regard to the constitution of the sun's central core, the opinion that now generally prevails, though not without some dissent, is that it is gaseous. According to the best determinations, the sun is about 320,000 times as heavy as the earth, and a million and a quarter times as large; it follows that its average density is less than one-fourth that of the earth. It is to be remembered also that since the force of gravity at the sun's surface is 28 times as great as on the earth, the effect of the weight of the strata near the surface in compressing and increasing the density of the central parts must be correspondingly more powerful. As things stand, then, there seems to be no possibility of admitting that the substances that compose the sun are mainly in the solid or liquid state, for in that case the mean density must almost necessarily far exceed that of the earth; yet the theory that they are in a gaseous state is not without difficulties. A few years ago it would have been urged with great plausibility that under such a pressure as must obtain at the center of the sun every gas would be liquefied. The recent researches of Andrews have shown, however, that a vapor or a gas, if above a certain critical temperature, refuses to be liquefied by any pressure whatever. Instead it grows denser and denser under the pressure, but still retains its gaseous characteristics."

"The manufacturers of firearms in this country are as busy as bees in clover time. Large orders from foreign governments are now being executed. Turkey is having 600,000 of the Peabody-Martini rifles made, Prussia lots of guns, Russia 100,000 of Smith and Wesson's pistols, while Spain calls for all that can be made of the Winchester and other breech-loaders."

"Professor Tyndall has dropped for the nonce the role of physicist for that of physician, and he deals in a recent publication with the subject of typhoid fever. He asserts positively that the weight of evidence is in favor of the view that the disease, like smallpox, arises wholly from contagion, and not spontaneously from fecal fermentation. He holds that the body is the seat of the development of the germ, and that the latter is not originated from noxious effluvia."

"The new and celebrated painting of the 'Roll Call' is now nightly exhibited in London to large audiences, by means of the oxyhydrogen light, or lime light, and all the colors of the picture are brought out with marvelous brilliancy, in fact with the same perfection as by daylight. The idea of illuminating art galleries in the evening by the lime light is an excellent one. The yellow color of the ordinary gas flame has the effect of revealing only a portion of the colors of the paintings. The reds and yellows are seen well enough, but the blues and greens, and their various tints, are sadly distorted, and the artistic effect lost. The use of the lime light or the electric light obviates such difficulties."

The Effects of Ice on Scotch

How fast a drink of Scotch whisky over rocks loses its flavor depends on the proof of the Scotch and the richness of its blend. These two factors are optimized for "on the rocks" Scotch drinkers in 90-Proof Famous Grouse, a venerable old brand from Scotland only recently introduced to America.

by Allen Mac Kenzie

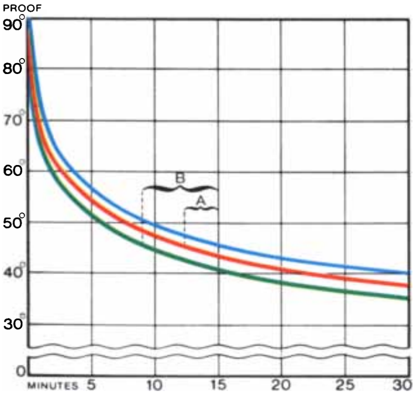
In countries where Scotch has been consumed for centuries, ice and whisky rarely mingle. But on this side of the Atlantic, the picture is quite different. While a small percentage of American Scotch drinkers take it neat, better than 35% drink it "on the rocks." The rest of us add varying amounts of water, club soda, etcetera. And ice. Always plenty of ice—the great American drink requisite.

It would seem then that the American Scotch devotee, particularly our on-the-rocks fancier, has a right to raise a serious question: *Is the Scotch I drink ideally suited to enjoyment over ice?*

Pursuing a Perfect Proof

Let's turn our attention first to the proof at which Scotch whisky is bottled. Consider the hypothesis that there is indeed a better proof for on-the-rocks Scotch drinking than that of the brand you currently favor.

Practically every Scotch sold in this country is bottled at 80, 86, or 86.8 Proof. So at the instant you pour Scotch over ice, it contains between 40% and 43.4% alcohol by volume. (Proof is double the percentage of alcohol.) The chilling effect of the ice is accompanied by dilution. And when your drink has been properly cooled—in 30 seconds to a minute—you achieve what one Scotch connoisseur refers to as



DILUTION BY ICE OF SCOTCH AT THREE DIFFERENT PROOFS (72°F). 80 Proof 86.8 Proof 90 Proof

"the ideal sip." From then on, the Scotch drinker's enjoyment typically runs downhill, as the drink loses its freshness.

While there is no way to preserve that fresh Scotch flavor indefinitely, we submit that you can sustain the freshness substantially longer with 90-Proof *Famous Grouse*. If you have never heard of this brand, we are not surprised. It is a well established name in Scotland, but only recently introduced to America. So far as we know, *Famous Grouse* is the only Scotch now available in this country at 90 Proof.

A Revealing Experiment

To demonstrate the merits of a slightly higher proof, we performed a simple experiment: 50 millilitres of Scotch (about 1.7 ounces) was chilled with 100 cc of ice. The ensuing dilutions at 80, 86.8 and 90 Proof are charted in the graph at left.

You'll notice that after 15 minutes on the rocks, the proof of *Famous Grouse* is diluted to a level which occurs after 12½ minutes when the Scotch is 86.8 Proof, and after 9 minutes when it is 80 Proof. In essence, the *Famous Grouse* brand has remained about 2½ minutes fresher than 86.8-Proof Scotch (Interval A on graph), 6 minutes fresher than 80-Proof Scotch (Interval B). If you "nurse" a drink beyond 15 minutes, the advantages of 90 Proof Scotch are even more pronounced.

Proof, of course, is not the only influence on the flavor of a blended Scotch. The proportion of malt to grain whiskies, origins of the malts, aging methods—these are also important factors determining the relative richness of Scotch flavor.

The makers of *Famous Grouse*—Matthew Gloag & Son of Perth, Scotland—have been producing Scotch in the same family for six generations. And they have performed their most noble feat in the rich blend they created for *Famous Grouse* Scotch. Its flavor—so remarkable at the outset—holds firmly to its character during prolonged contact with ice.

Knowledge of Scotch, however, cannot be indefinitely pursued in the abstract. Your learning process must ultimately include a leisurely sip of *Famous Grouse* on the rocks. For Scotch drinking is one of those pleasures enjoyed most, not in the pursuit, but in the conquest. Scotland's greatest bard, Robert Burns, said it best:

"Gie me a spark o' Nature's fire,
That's a' the learning I desire."

90 PROOF BLENDED SCOTCH WHISKY BOTTLED IN SCOTLAND
IMPORTED BY AUSTIN, NICHOLS & CO., LAWRENCEBURG, KY

Announcing a new generation of International Paper Supertrees. And not a moment too soon.

Do you have any idea how much paper and wood the average American will use by the year 2000?

Over a ton a year.

Now multiply that by 264.4 million Americans, and ask yourself where on earth all that paper and lumber will come from.

At International Paper, we're working hard to come up with some answers.

Like Supertrees.

The average Supertree grows faster, taller, straighter, and healthier than ordinary pine trees.

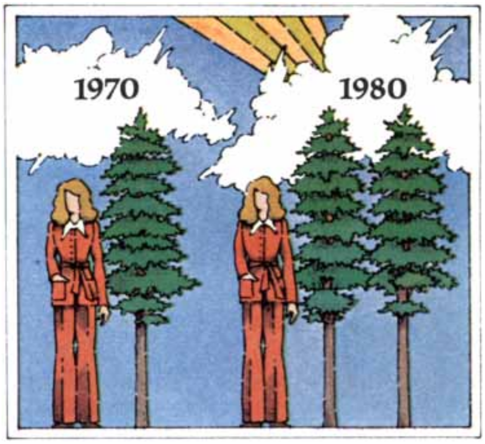
It produces more wood fiber for more paper. And like all trees, a Supertree is a renewable resource.

The best of the best.

How do you breed championship animals? Simply explained, you find two champions and mate them.

At International Paper, we've been doing the same thing with our Supertrees.

After years of testing, we've mated the best of our original Supertrees, and now we're beginning the second-generation selections. These new Supertrees will be even straighter. Even taller. Even healthier than their parents.



By 1980, International Paper will have planted two Supertrees for every person in America.

There will be more Supertrees, too. In fact, by 1980, International Paper will have planted two Supertrees for every man, woman and child in the country.

We've been working to build the world's fiber supply for decades. And not just with Supertrees.

Other things we're doing.

We've developed the Landowner Assistance Program, to help small landowners do a better job of managing their forests.

We're finding ways to get more wood fiber from our trees. For example, we're experimenting with a new machine that harvests tap roots that used to be left in the ground.

We're moving ahead on projects like cooperative nurseries. Tree farm programs. Forest research.

Will all this be enough to keep the world's fiber supply going strong?

It will help. But more must be done.

At International Paper, we believe forest products companies, private landowners and the Government should work together to find more solutions. To make better use of existing land. To abandon wasteful habits. And to develop better scientific management techniques, so all our forests can reach their full potential.

If you'd like more information about what has to be done to assure the world's fiber supply, just write Public Relations Department, Perspective, International Paper Company, 220 East 42nd Street, New York, N.Y. 10017.

INTERNATIONAL PAPER COMPANY

220 EAST 42ND STREET NEW YORK NEW YORK 10017

I am you.

My name is Masakazu Fukushima, and here I am trumping some neighbors at contract bridge. I play every chance I get. It beats watching television.

Not that I've got anything against watching television, mind you. Watching is fine. What really bothers me is the bulk of the sets themselves. You get a big-screen color TV in your living room and it's like you parked a small car in there. An ugly small car.

The fact is I work for a company that makes a lot of TVs... Hitachi. I do electronics research, so I thought I'd try and do something to make color TVs easier to live with.

All people really want from a color TV is a good picture. So we figured the set should be *shaped* like a picture. Flat. To be hung on the wall.

We figured it out, then went into the lab and put it together. We came up with what we wanted ... the world's first working prototype for a flat-profile color TV. It came out even better than we thought it would.

Not only is this new TV flat, the way it's designed is such that the screen can be just about however big you want it. There could be whole TV walls, say, in schools.

It opens up a lot of new, better possibilities for television, and that makes me happy. By doing myself a favor—by helping to invent something I'd enjoy—I've made something that can benefit an enormous number of people.

I am you.

HITACHI
Resourceful people

Hitachi America, Ltd.
Head Office: 437 Madison Avenue, New York, N.Y. 10022
Tel: (212) 758-5420
Branch Offices: San Francisco, Chicago, Atlanta.

● Power Generation & Transmission Equipment ● Industrial Machinery ● Transportation Equipment ● Communications Equipment ● Electronics & Home Appliances

THE AUTHORS

ROBERT BUCKHOUT ("Eyewitness Testimony") is associate professor of psychology and director of the Center for Responsive Psychology at Brooklyn College of the City University of New York. "I was born in Brooklyn," he writes, "and after leaving with a vow never to return, I am back–minutes away from where I was born. My career as a psychologist began as a major subject chosen after minoring in mistakes. After taking a master's degree at North Carolina State University I spent four years in the Air Force. I went to Ohio State University, obtaining my Ph.D. In 1963 I began teaching at Washington University and getting involved in the social problems of the time." Buckhout adds that he owes "more than I can express in words" to a group of undergraduates who helped him with the research on which his article is based. They include Andrea Alper, Susan Chern, Robinsue Frohboese, Richard Harwood, Daryl Figueroa, Ethan Hoff, Carolyn Hogan, Noreen Norton, Vincent Reilly, Betti Sachs, Glenn Silverberg, Miriam Slomovits and Lynne Williams. He writes that he is "also indebted to Leo Branton and Howard Moore, attorneys, who taught me to present psychology in English to a jury."

KIP S. THORNE ("The Search for Black Holes") is professor of theoretical physics at the California Institute of Technology and adjunct professor of physics at the University of Utah. He was graduated from Cal Tech in 1962 and received his master's and doctor's degrees from Princeton University in 1963 and 1965 respectively. "Eighteen months ago," he writes, "my brother and my wife built a small cabin on Palomar Mountain, a mile and a half from the nearest habitation. I have spent much of the past year there, amidst the solitude and beauty of the wilderness, reading, thinking, calculating and writing about relativistic stars, gravitational waves and Cygnus X-1. Once every few days, in the evening, I walked over to the Observatory to consult a few books and recharge my electronic calculator. It was a great year!"

ALFRED GIERER ("Hydra as a Model for the Development of Biological Form") is director of the microbiology division of the Max Planck Institute for Virus Research in Tübingen. He was graduated from the University of Göttingen in 1953 and obtained his Habilitation degree, which qualified him for a professorial appointment, at the University of Tübingen in 1958. From 1958 to 1964 he was successively lecturer in and professor of biophysics at Tübingen.

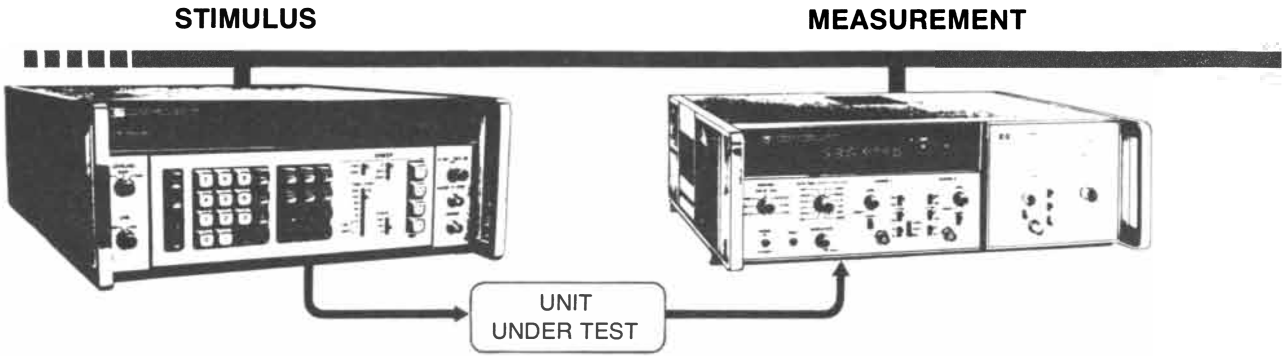
GOVINDJEE AND RAJNI GOVINDJEE ("The Primary Events in Photosynthesis") are at the University of Illinois at Urbana-Champaign; Govindjee is professor of botany and acting head of the department of botany and Rajni Govindjee is research associate in the department of physiology and biophysics. Govindjee, who uses only one name (it was his given name; his last name, Asthana, was dropped by his father as a protest against the caste system in India), obtained his bachelor's degree in 1952 and his master's degree in 1954 at the University of Allahabad. After teaching plant physiology there for two years he went to Illinois, where he obtained his Ph.D. (in biophysics) in 1960. Rajni Govindjee obtained her master's degree (in plant physiology) from Allahabad in 1955 and her Ph.D. (for work in photosynthesis) from Illinois in 1961.

MERTON C. FLEMINGS ("The Solidification of Castings") is Abex Professor of Metallurgy at the Massachusetts Institute of Technology. All his degrees are from M.I.T.: his bachelor's degree in 1951, his master's in 1952 and his doctorate in 1954. He worked for two years for the Abex Corporation and then returned to M.I.T., where he has been a member of the faculty since 1956. Flemings also works as a consultant to industry. "My interests in metallurgy," he writes, "lie almost entirely at the interface between theory and practice and particularly in the area of process innovation through the application of theory." He describes himself as "a dabbler" in other fields, including "educational innovation at the junior-college level, techniques of sculpture, tennis and climbing."

WILLIAM T. KEETON ("The Mystery of Pigeon Homing") is professor of biology at Cornell University, where he also serves as chairman of the section of neurobiology and behavior of the division of biological sciences. He was graduated from the University of Chicago in 1952 and obtained his master's degree four years later at the Virginia Polytechnic Institute. His Ph.D. (in entomology) is from Cornell, and for a year after receiving it in 1958 he taught entomology there. He became assistant professor of biology in 1959. Keeton writes that a neighbor introduced him to bird-watching when he was eight and that "bird-watching and pigeon-racing became boyhood hobbies."

DAVID B. CLINE, ALFRED K. MANN and CARLO RUBBIA ("Neutral Currents in Weak Interactions") are respectively professor of physics at the University of Wisconsin, professor of physics at the University of Pennsylvania and professor of physics at Harvard University. Cline did his undergraduate work at Kansas State University, obtaining his Ph.D. at Wisconsin in 1965. He is an experimental physicist studying elementary particles with bubble chambers and fast electronics. His main hobby is helping his wife in her profession of designing and manufacturing dresses. Mann's degrees are all from the University of Virginia; he received his Ph.D. there in 1947. He went to Pennsylvania in 1949 after two years of teaching at Columbia University. Rubbia was born in Italy and received his training in physics at the Normal School of Pisa and the University of Pisa. He taught at the university for a year before coming to the U.S. in 1959. He spent a year as a research fellow at Columbia, two years teaching at the University of Rome and eight years as a senior physicist at the European Organization for Nuclear Research (CERN) in Geneva before joining the Harvard faculty in 1970.

GEORGE C. BOON ("Counterfeiting in Roman Britain") is assistant keeper of archaeology at the National Museum of Wales. He writes that his interest in Romano-British studies began at the age of nine in Bristol, where he was born and grew up. His university education was in Latin. On obtaining his degree in 1950 he worked for six years as archaeological assistant at the Reading Museum. At the National Museum of Wales, which he joined in 1957, he has been concerned principally with the legionary fortress of Caerleon and other Roman sites in Wales and with numismatics, another interest that began in his boyhood. He writes that "technical aspects of Roman civilization–mining, pottery, glass, bronze work and the like –appeal to me." Boon is the author of four books; the most recent of them is *Silchester: The Roman Town of Calleva.* "I live," he notes, "in a crumbling Victorian house at Penarth, near Cardiff, where I reluctantly devote time off from work to gardening and decoration."



A standard communications link that facilitates conversation among instruments.

It wasn't long ago that all instruments were, in human terms, totally deaf and dumb. They could not hear instructions so you made them do their job by setting knobs and switches. And when the job was done, they could not tell you the results; the only way to find out was to read, and then analyze, their displays.

Many instruments have since learned to "talk." On command, they can output measurement results and transmit them remotely in code. More and more are being equipped to "listen": send them prearranged signals and they can program their own controls, remotely. Add a control function to such instruments—to tell them when to talk and when to listen—and they can communicate with each other automatically.

This sounds easy, but it hasn't been. Although the three basic elements for automatic instrumentation systems—talkers, listeners, and controllers—are readily available, one who sets out to design and assemble such a system quickly runs into severe frustrations. The different elements are rarely compatible; more often than not, they use different logic, speak a different language, and interconnect with different hardware.

Avoiding this electronic Tower of Babel is what the Hewlett-Packard Interface Bus (HP-IB) is all about. A standard interface system, the HP-IB forms a basic communications link that allows interconnected system components to communicate effectively, in an orderly and unambiguous manner. The interface system involves much more than the standardization of interconnecting cables; it also defines the interface logic capabilities within the system instruments, the scope of the data codes used on the interface, and the timing and control techniques for exchanging messages.

To talk or to listen: never a doubt.

In the HP Interface Bus, all system devices are exposed to all system communications. But a device can neither send nor receive a message unless told to do so by the system *controller*: at any given time, it can be either a *talker* or a *listener*, but not both. Listeners receive programming data from a controller or measurement data from talkers; talkers send measurement data to listeners. There can never be more than one active controller or one talker at the same time, but there can be as many as 14 concurrent listeners.

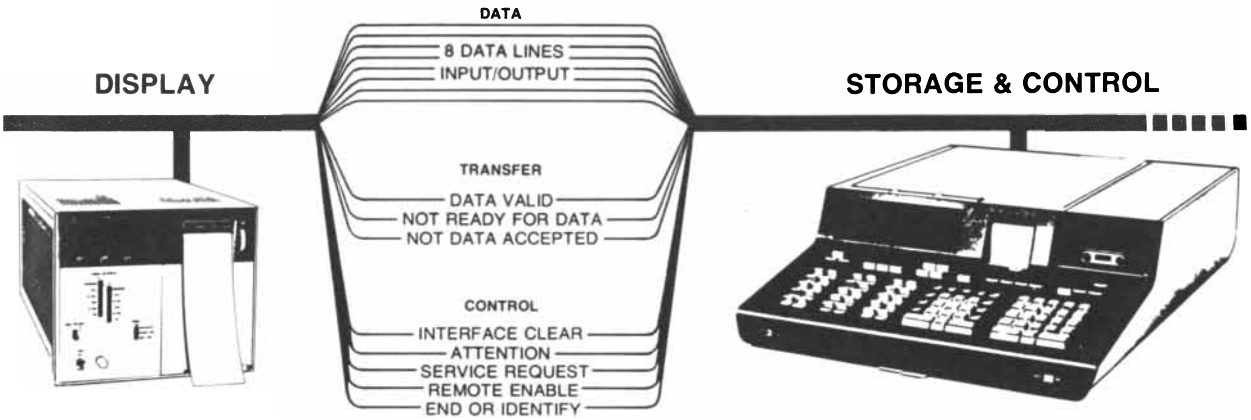
Depending on its capabilities, a device may play more than one role at different times. A calculator or computer, for example, can be talker, listener, or controller; a programmable digital voltmeter alternately talks when it outputs its measurement and listens when it's being programmed; a paper punch can only play the role of listener.

The bus: a common interconnection.

All system devices are interconnected on a common set of 16 signal lines. Eight of these lines form the *data bus* which carries all data messages bidirectionally between talkers and listeners, in bit-parallel byte-serial fashion. The *transfer bus* uses three lines to ensure that data is interchanged only from the intended talker to the designated listeners, through an interrogation and reply sequence. The remaining five signal lines constitute the *control bus*, by which the controller directs an orderly flow of information across the interface, sending commands to the devices and receiving service requests from them. Although system control is always delegated (never assumed), it may be shifted from one system device to another.

HP-IB simplifies systems, small or large.

An HP-IB system can consist of one talker, one listener, and *no* controller; for example, a counter and digital printer for semi-automatic data logging. At the other extreme, a completely auto-



matic system may include as many as 15 instruments possessing stimulus, measurement, display, storage, and control capabilities. Whether a calculator, computer, or the processor of a "smart" instrument, the controller operates the entire system through an interface connection (a single I/O card)—an obvious economy compared to non-bus systems that require one I/O card for each instrument.

System configuration: fundamental problems solved.

Although the HP-IB does not provide instant systems, it does solve the fundamental interface problems that have plagued instrumentation system designers and users until now. Designers no longer need to invent custom interfaces for each new product; users no longer need to familiarize themselves with an interface unique to each new product. Cable and connector problems are minimized by the use of a simple, passive cable interconnection system.

HP-IB protocol allows the designer to assign talk and listen addresses to each device to suit his purposes. Each address is set at the device to any desired value, through a switch on a rear panel, jumper wires on a PC board, or other convenient means.

The HP-IB imposes minimal functional restrictions on data transfer between a talker and a listener. For example, data bytes may consist of from one to eight bits. Once a device is addressed, data can be transferred using any coding and format convention appropriate to the application. The most commonly used codes are the printable characters of the ASCII code set, and the number representations are typically FORTRAN compatible.

Minimal timing restrictions are imposed on the data rates by the HP-IB. Data is transferred asynchronously at a rate that suits the devices involved; burst rates of 1 megabyte per second are possible over limited distances. Data may be transferred directly between devices, thus reducing message traffic on the bus.

More than a theory, HP-IB is a reality now.

Within Hewlett-Packard, the common interface concept has already been incorporated into a growing list of more than 20 instruments and accessory products as well as our programmable desk-top calculators. Outside the company, the HP-IB has served as a proposed interface standard now under active consideration by the IEC and the IEEE. Thus the possibility exists that this concept will become an international communication link applicable to instruments and peripherals without regard to manufacturer or nation of origin.

Obviously an idea whose time has come, the common interface is here now, still another aspect of the new measurement technology that is taking shape at Hewlett-Packard.

For more information, write to us. Hewlett-Packard, 1503 Page Mill Road, Palo Alto, California 94304.

00436

Mail to: Hewlett-Packard, 1503 Page Mill Road, Palo Alto, CA 94304

Please send me further information on the Hewlett-Packard Interface Bus.

Name ____________ Title ____________

Company ____________

Address ____________

City ____________ State ________ Zip ________

HEWLETT *hp* PACKARD

Sales and service from 172 offices in 65 countries.
Palo Alto, California 94304

Established 1845 SCIENTIFIC AMERICAN December 1974 Volume 231 Number 6

Eyewitness Testimony

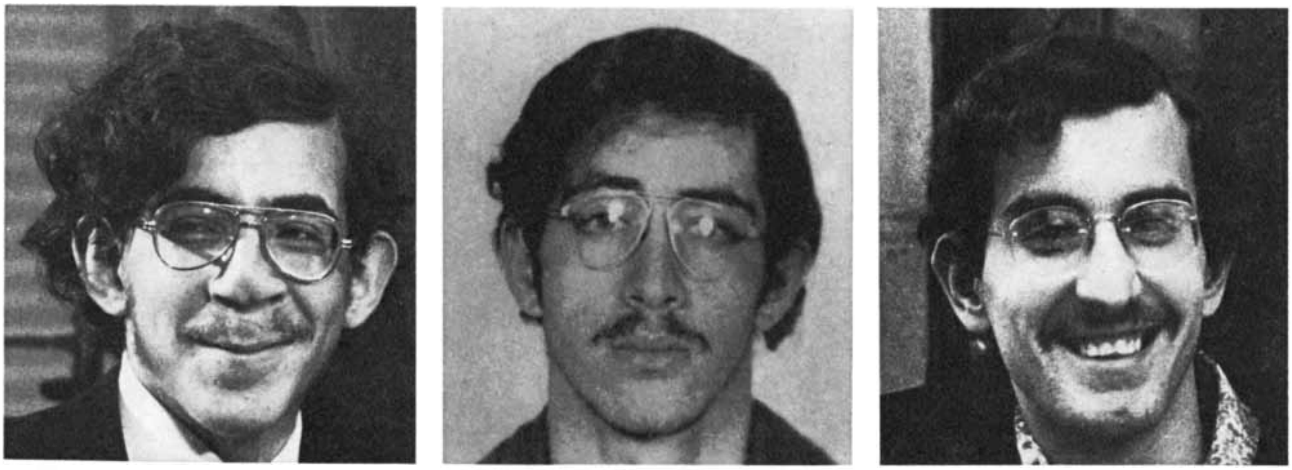
Although such testimony is frequently challenged, it is still widely assumed to be more reliable than other kinds of evidence. Numerous experiments show, however, that it is remarkably subject to error

by Robert Buckhout

The woman in the witness box stares at the defendant, points an accusing finger and says, loudly and firmly, "That's the man! That's him! I could never forget his face!" It is impressive testimony. The only eyewitness to a murder has identified the murderer. Or has she?

Perhaps she has, but she may be wrong. Eyewitness testimony is unreliable. Research and courtroom experience provide ample evidence that an eyewitness to a crime is being asked to be something and do something that a normal human being was not created to be or do. Human perception is sloppy and uneven, albeit remarkably effective in serving our need to create structure out of experience. In an investigation or in court, however, a witness is often asked to play the role of a kind of tape recorder on whose tape the events of the crime have left an impression. The prosecution probes for stored facts and scenes and tries to establish that the witness's recording equipment was and still is in perfect running order. The defense cross-examines the witness to show that there are defects in the recorder and gaps in the tape. Both sides, and usually the witness too, succumb to the fallacy that everything was recorded and can be played back later through questioning.

Those of us who have done research in eyewitness identification reject that fallacy. It reflects a 19th-century view of man as perceiver, which asserted a parallel between the mechanisms of the physical world and those of the brain. Human perception is a more complex information-processing mechanism. So is memory. The person who sees an accident or witnesses a crime and is then asked to describe what he saw cannot call up an "instant replay." He must depend on his memory, with all its limitations. The limitations may be unimportant in ordinary daily activities. If someone is a little



MISTAKEN IDENTIFICATIONS led to the arrests of two innocent men: Lawrence Berson (*left*) for several rapes and George Morales (*right*) for a robbery. Both men were picked out of police lineups by victims of the crimes. Berson was cleared when Richard Carbone (*center*) was arrested and implicated in the rapes. Carbone was convicted. Later he confessed to the robbery, clearing Morales.

REENACTMENT OF A MURDER was photographed at the same time of night as the murder and from the viewing position of an eyewitness who said he had been 120 feet away. The witness had identified a suspect charged with killing another man in darkened doorway.

MEASUREMENTS OF BRIGHTNESS (in millilamberts) at various points in the scene showed how little light was reflected from the doorway to the eyewitness. The light readings and the photograph (*top*) combined to cast doubt on the accuracy of the identification.

unreliable, if he trims the truth a bit in describing what he has seen, it ordinarily does not matter too much. When he is a witness, the inaccuracy escalates in importance.

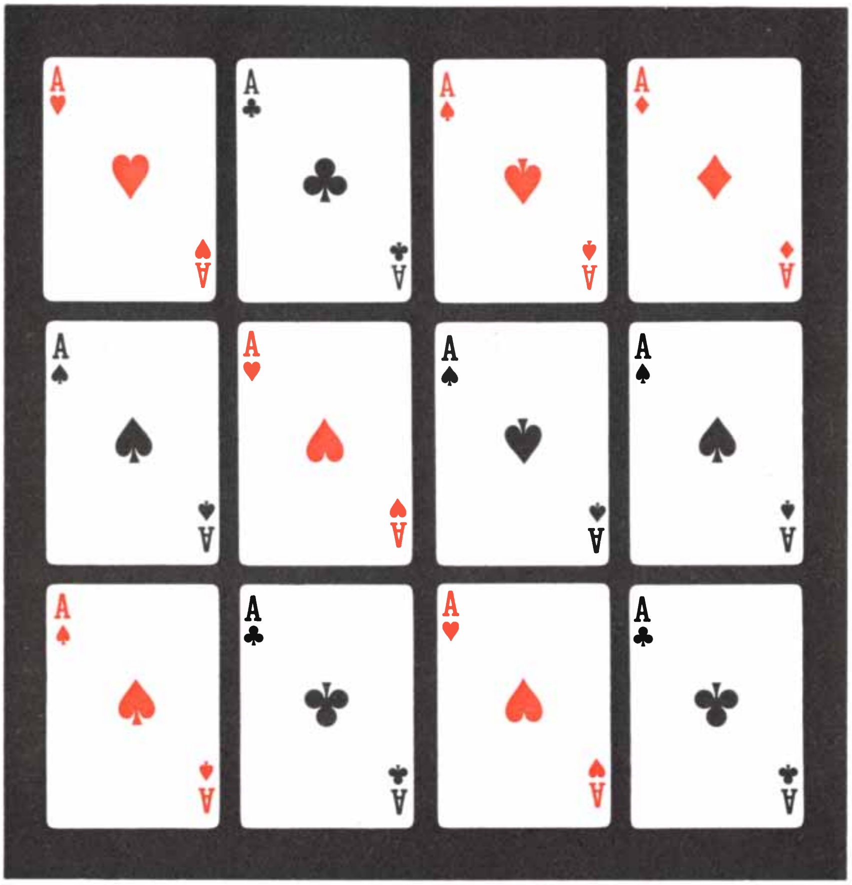
Human perception and memory function effectively by being selective and constructive. As Ulric Neisser of Cornell University has pointed out, "Neither perception nor memory is a copying process." Perception and memory are decision-making processes affected by the totality of a person's abilities, background, attitudes, motives and beliefs, by the environment and by the way his recollection is eventually tested. The observer is an active rather than a passive perceiver and recorder; he reaches conclusions on what he has seen by evaluating fragments of information and reconstructing them. He is motivated by a desire to be accurate as he imposes meaning on the overabundance of information that impinges on his senses, but also by a desire to live up to the expectations of other people and to stay in their good graces. The eye, the ear and other sense organs are therefore social organs as well as physical ones.

Psychologists studying the capabilities of the sense organs speak of an "ideal observer," one who would respond to lights or tones with unbiased eyes and ears, but we know that the ideal observer does not exist. We speak of an "ideal physical environment," free of distractions and distortions, but we know that such an environment can only be approached, and then only in the laboratory. My colleagues and I at the Brooklyn College of the City University of New York distinguish a number of factors that we believe inherently limit a person's ability to give a complete account of events he once saw or to identify with complete accuracy the people who were involved.

The first sources of unreliability are implicit in the original situation. One is the insignificance—at the time and to the witness—of the events that were observed. In placing someone at or near the scene of a crime, for example, witnesses are often being asked to recall seeing the accused at a time when they were not attaching importance to the event, which was observed in passing, as a part of the normal routine of an ordinary day. As long ago as 1895 J. McKeen Cattell wrote about an experiment in which he asked students to describe the people, places and events they had encountered walking to school over familiar paths. The reports were incomplete and unreliable; some individuals were very sure of details that had no basis in fact. Insignificant events do not motivate a person to bring fully into play the selective process of attention.

The length of the period of observation obviously limits the number of features a person can attend to. When the tachistoscope, a projector with a variable-speed shutter that controls the length of an image's appearance on a screen, is used in controlled research to test recall, the shorter times produce less reliable identification and recall. Yet fleeting glimpses are common in eyewitness accounts, particularly in fast-moving, threatening situations. In the Sacco-Vanzetti case in the 1920's a witness gave a detailed description of one defendant on the basis of a fraction-of-a-second glance. The description must have been a fabrication.

Less than ideal observation conditions usually apply; crimes seldom occur in a well-controlled laboratory. Often distance, poor lighting, fast movement or the presence of a crowd interferes with the efficient working of the attention process. Well-established thresholds for the eye and the other senses have been established by research, and as those limits are approached eyewitness accounts become quite unreliable. In one case in my experience a police officer testified that he saw the defendant, a black man, shoot a victim as both stood in a doorway 120 feet away. Checking for the defense, we found the scene so poorly lit that we could hardly see a person's silhouette, let alone a face; instrument measurements revealed that the light falling on the eye amounted to less than a fifth of the light from a candle. The defense presented photographs and light readings to demonstrate that a positive identification was not very probable. The members of the jury went to the scene of the crime, had the one black juror stand in the doorway, found they could not identify his features and acquitted the defendant.



HOW MANY ACES OF SPADES DID YOU SEE? After a brief glance at this display of playing cards most people report seeing three. Actually there are five. Because people expect aces of spades to be black, not red, they tend to see only the black ones and to miss the atypical red ones. Thus do prior conditioning and experience influence perception.

The witness himself is a major source of unreliability. To begin with, he may have been observing under stress. When a person's life or well-being is threatened, there is a response that includes an increased heart rate, breathing rate and blood pressure and a dramatic increase in the flow of adrenalin and of available energy, making the person capable of running fast, fighting, lifting enormous weight—taking the steps necessary to ensure his safety or survival. The point is, however, that a person under extreme stress is also a less than normally reliable witness. In experimental situations an observer is less capable of remembering details, less accurate in reading dials and less accurate in detecting signals when under stress; he is quite naturally paying more attention to his own well-being and safety than to nonessential elements in the environment. Research I have done with Air Force flight-crew members confirms that even highly trained people become poorer observers under stress. The actual threat that brought on the stress response, having been highly significant at the time, can be remembered; but memory for other details such as clothing and colors is not as clear; time estimates are particularly exaggerated.

The observer's physical condition is often a factor. A person may be too old or too sick or too tired to perceive clearly, or he may simply lack the necessary faculty. In one case I learned that a witness who had testified about shades of red had admitted to the grand jury that he was color-blind. I testified at the trial that he was apparently dichromatic, or red-green color-blind, and that his testimony was probably fabricated in the basis of information other than visual evidence. The prosecution brought on his ophthalmologist, presumably as a rebuttal witness, but the ophthalmologist testified that the witness was actually monochromatic, which meant he could perceive no colors at all. Clearly the witness was "filling in" his testimony. That, after all, is how color-blind people function in daily life, by making inferences about colors they cannot distinguish.

Psychologists have done extensive research on how "set," or expectancy, is used by the observer to make judgments more efficiently. In a classic experiment done in the 1930's by Jerome S. Bruner and Leo Postman at Harvard University observers were shown a display of playing cards for a few seconds and asked to report the number of aces of spades in the display [*see illustration above*]. After a brief glance most observers reported seeing three aces of spades. Actually there were five; two of them were colored red instead of the more familiar black. People are so familiar with black aces of spades that they do not waste time looking at the display carefully. The prior conditioning of the witness may cause him similarly to report facts or events that were not present but that he thinks should have been present.

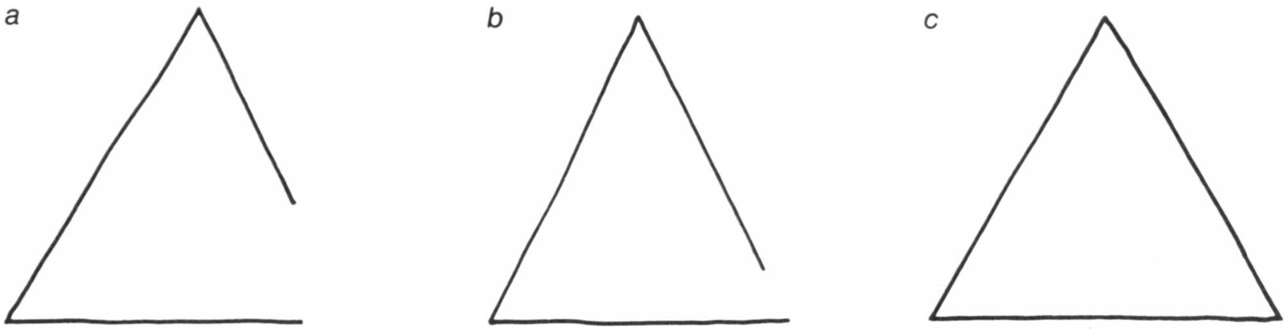
Expectancy is seen in its least attractive form in the case of biases or prejudices. A victim of a mugging may initially report being attacked by "niggers" and may, because of prejudice or limited experience (or both), be unable to tell one black man from another. ("They all look alike to me.") In a classic study of this phenomenon Gordon W. Allport of Harvard had his subjects take a brief look at a drawing of several people on a subway train, including a seated black man and a white man standing with a razor in his hand. Fifty percent of the observers later reported that the razor was in the hand of the black man. Most people file away some stereotypes on the basis of which they make perceptual judgments; such stereotypes not only lead to prejudice but are also tools for making decisions more efficiently. A witness to an automobile accident may report not what he saw but his ingrained stereotype about women drivers. Such short-cuts to thinking may be erroneously reported and expanded on by an eyewitness without his being aware that he is describing his stereotype rather than actual events. If the witness's biases are shared by the investigator taking a statement, the report may reflect their mutual biases rather than what was actually seen.

The tendency to see what we want or need to see has been demonstrated by numerous experiments in which people report seeing things that in fact are not present. R. Levine, Isador Chein and Gardner Murphy had volunteers go without food for 24 hours and report what they "saw" in a series of blurred slides presented on a screen. The longer they were deprived of food the more frequently they reported seeing "food" in the blurred pictures. An analysis of the motives of the eyewitness at the time of a crime can be very valuable in determining whether or not the witness is reporting what he wanted to see. In one study I conducted at Washington University a student dressed in a black bag that covered him completely visited a number of classes. Later the students in those classes were asked to describe the nature of the person in the bag. Most of their reports went far beyond the meager evidence: the bag-covered figure was said to be a black man, "a nut," a symbol of alienation and so on. Further tests showed that the descriptions were related to the needs and motives of the individual witness.

Journalists and psychologists have noted a tendency for people to maintain they were present when a significant historical event took place near where they live even though they were not there at all; such people want to sound interesting, to be a small part of history. A journalist once fabricated a charming human interest story about a naked woman stuck to a newly painted toilet seat in a small town and got it distributed by newspaper wire services. He visited the town and interviewed citizens who claimed to have witnessed and even to have played a part in the totally fictitious event. In criminal cases with publicity and a controversial defendant it is not uncommon for volunteer witnesses to come forward with spurious testimony.

Unreliability stemming from the original situation and from the observer's fallibility is redoubled by the circumstances attending the eventual attempt at information retrieval. First of all there is the obvious fact, supported by a considerable amount of research, that people forget verbal and pictorial information with the passage of time. They are simply too busy coping with daily life to keep paying attention to what they heard or saw; perfect recall of informa-

WHO HAD THE RAZOR? After a brief look at a drawing such as this one, half of the observers report having seen the razor, a stereotyped symbol of violence in blacks, in the black man's hand. Gordon W. Allport of Harvard University devised this experiment.



"FILLING IN" OF DETAILS was demonstrated by a simple drawing test. Observers were shown an incomplete but roughly triangular figure and immediately afterward were asked to draw what they had seen. The typical drawing was a good reproduction of the original (*a*). A month later observers asked to draw what they remembered produced more regular figures (*b*). Three months after the original viewing, again asked to draw what they remembered, they drew erroneously complete, symmetrical figures (*c*).

tion is basically unnecessary and is rarely if ever displayed. The testing of recognition in a police "lineup" or a set of identification photographs is consequently less reliable the longer the time from the event to the test. With time, for example, there is often a filling in of spurious details: an incomplete or fragmentary image is "cleaned up" by the observer when he is tested later. Allport used to have students draw a rough geometric shape right after such a shape was shown to them. Then they were tested on their ability to reproduce the drawing 30 days later and again three months later [*see illustration above*]. The observers tended first to make the figure more symmetrical than it really was and later to render it as a neat equilateral triangle. This finding was repeated with many objects, the tendency being for people to "improve" their recollection by making it seem more logical.

In analyses of eyewitness reports in criminal cases we have seen the reports get more accurate, more complete and less ambiguous as the witness moves from the initial police report through grand-jury questioning to testimony at the trial. The process of filling in is an efficient way to remember but it can lead to unreliable recognition testing: the witness may adjust his memory to fit the available suspects or pictures. The witness need not be lying; he may be unaware he is distorting or reconstructing his memory. In his very effort to be conscientious he may fabricate parts of his recall to make a chaotic memory seem more plausible to the people asking questions. The questions themselves may encourage such fabrication. Beth Loftus of the University of Washington has demonstrated how altering the semantic value of the words in questions about a filmed auto accident causes witnesses to distort their reports. When witnesses were asked a question using the word "smashed" as opposed to "bumped" they gave higher estimates of speed and were more likely to report having seen broken glass—although there was no broken glass.

Unfair test construction often encourages error. The lineup or the array of photographs for testing the eyewitness's ability to identify a suspect can be analyzed as fair or unfair on the basis of criteria most psychologists can agree on. A fair test is designed carefully so that all faces have an equal chance of being selected by someone who did not see the suspect; the faces are similar enough to one another and to the original description of the suspect to be confusing to a person who is merely guessing; the test is conducted without leading questions or suggestions. All too frequently lineups or photograph arrays are carelessly assembled or even rigged. If, for example, there are five pictures, the chance should be only one in five that any one picture will be chosen on the basis of guessing.

Frequently, however, one picture—the picture of the suspect—may stand out. In the case of the black activist Angela Davis one set of nine photographs used to check identification included three pictures of the defendant taken at an outdoor rally, two police "mug shots" of other women with their names displayed, a picture of a 55-year-old woman and so on. It was so easy for a witness to rule out five of the pictures as ridiculous choices that the test was reduced to four photographs, including three of Miss Davis. The probability was therefore 75 percent that a witness would pick out her picture whether he had seen her or not. Such a "test" is meaningless to a psychologist and is probably tainted as evidence in court.

Research on memory has also shown that if one item in the array of photographs is uniquely different—say in dress, race, height, sex or photographic quality—it is more likely to be picked out. Such an array is simply not confusing enough for it to be called a test. A teacher who makes up a multiple-choice test includes several answers that sound or look alike to make it difficult for a person who does not know the right answer to succeed. Police lineups and picture layouts are multiple-choice tests; if the rules for designing tests are ignored, the tests are unreliable.

No test, with photographs or a lineup, can be completely free of suggestion. When a witness is brought in by the police to attempt an identification, he can safely assume that there is some reason: that the authorities have a suspect in mind or even in custody. He is therefore under pressure to pick someone even if the officer showing the photographs is properly careful not to force the issue. The basic books on eyewitness identification all recommend that no suggestions, hints or pressure be transmitted to the witness, but my experience with criminal investigation reveals frequent abuse by zealous police officers. Such abuses include making remarks about which pictures to skip, saying, "Are you sure?" when the witness makes an error, giving hints, showing enthusiasm when the "right" picture is picked and so on. There is one version of the lineup in which five police officers in civilian clothes stand in the line, glancing obviously at the one real suspect. Suggestion can be subtler. In some experiments the test giver was merely instructed to smile and be very approving when a certain kind of photograph or statement was picked; such social approval led to an increase in the choosing of just those photographs even though there was no "correct" answer. A test that measures a need for social approval has shown that people who are high in that need (particularly those who enthusiastically volunteer information) are particularly strongly influenced by suggestion and approval coming from the test giver.

Conformity is another troublesome influence. One might expect that two eyewitnesses—or 10 or 100—who agree are better than one. Similarity of judgment is a two-edged sword, however: people can agree in error as easily as in truth. A large body of research results demonstrates that an observer can be persuaded to conform to the majority opinion even when the majority is completely wrong. In one celebrated experiment, first performed in the 1950's by Solomon E. Asch at Swarthmore College, seven observers are shown two lines and asked to say which is the shorter. Six of the people are in the pay of the experimenter; they all say that the objectively longer line is the shorter one. After hearing six people say this, the naïve subject is on the spot. Astonishingly the majority of the naïve subjects say that the long line is short—in the face of reality and in spite of the fact that alone they would have no trouble giving the correct answer [see "Opinions and Social Pressure," by Solomon E. Asch; SCIENTIFIC AMERICAN, November, 1955].

To test the effect of conformity a group of my students at Brooklyn College, led by Andrea Alper, staged a "crime" in a classroom, asked for individual descriptions and then put the witnesses into groups so as to produce composite descriptions of the suspect. The group descriptions were more complete than the individual reports but gave rise to significantly more errors of commission: an assortment of incorrect and stereotyped details. For example, the groups (but not the individuals) reported incorrectly that the suspect was wearing the standard student attire, blue jeans.

The effects of suggestion increase when figures in obvious authority do the testing. In laboratory research we find more suggestibility and changing of attitudes when the tester is older or of apparently higher status, better dressed or wearing a uniform or a white coat—or is a pretty woman. In court I have noticed that witnesses who work together under a supervisor are hard put to disagree with their boss in testifying or in picking a photograph. The process of filling in details can be exaggerated when the boss and his employee compare their information and the employee feels obligated to back up his boss to remain in his good graces. Legal history is not lacking in anecdotes about convict witnesses who were rewarded by the authorities for their cooperation in making an identification.

In criminal investigations, as in scientific investigations, a theory can be a powerful tool for clarifying confusion, but it can also lead to distortion and unreliability if people attempt, perhaps unconsciously, to make fact fit theory and close their minds to the real meanings of facts. The eyewitness who feels pressed to say something may shape his memory to fit a theory, particularly a highly publicized and seemingly reasonable one. Robert Rosenthal of Harvard studied this effect. He devised a test in


1


2


3


4


5


6

TWO LAYOUTS OF PHOTOGRAPHS like these were presented to witnesses to a staged assault. The actual attacker had been the young man labeled No. 5. In the unbiased spread (*left*) the portraits are aligned and show similar full-face views; in the biased spread (*right*) the culprit's head is tilted and he is grinning, and the portrait itself is placed at an angle. Witnesses were shown one

which people were supposed to pick out a "successful" face from a set of photographs. There was actually no correct answer, but the experimenter dropped hints to his assistants as to what he thought the results should be. When they subsequently administered the test the assistants unconsciously signaled the subjects as to which photograph to pick, thus producing results that supported their boss's theory. Any test is a social interaction as well as a test.

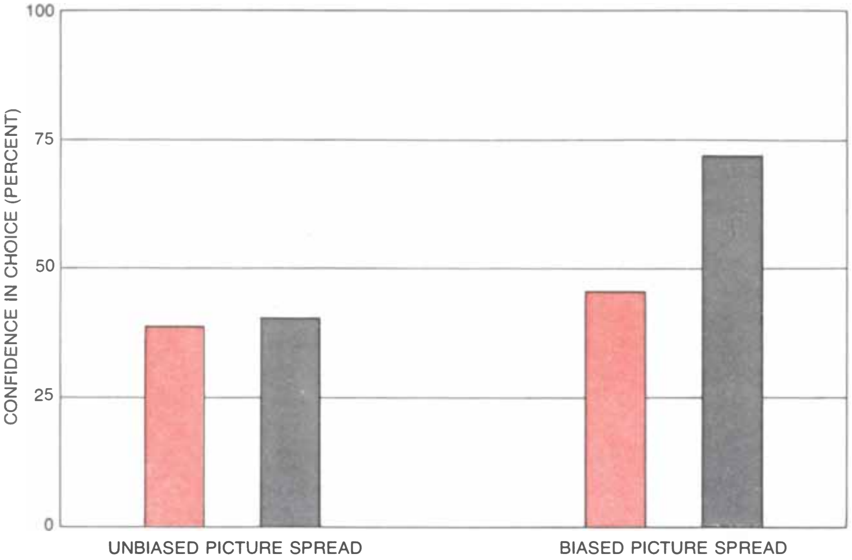
There is a nagging gap between data on basic perceptual processes in controlled research settings and important questions about perception in the less well-controlled real world. Inspired by the new approach to perception research exemplified in the work of Neisser and of Ralph Norman Haber of the University of Rochester, my colleagues and I have felt that this gap can only be bridged by conducting empirical research on eyewitness identification in a somewhat real world. In one such experiment we staged an assault on the campus of the California State University at Hayward: a student "attacked" a professor in front of 141 witnesses; another outsider of the same age was on the scene as a bystander. We recorded the entire incident on videotape so that we could compare the true event with the eyewitness reports. After the attack we took sworn statements from each witness, asking them to describe the suspect, his clothes and whatever they could remember about the incident. We also asked each witness to rate his own confidence in the accuracy of his description.

As we expected, the descriptions were quite inaccurate, as is usually the case in such situations. The passage of time was overestimated by a factor of almost two and a half to one. The average weight estimate for the attacker was 14 percent too high, and his age was underestimated by more than two years. The total accuracy score, with points given for those judgments and for others on appearance and dress, was only 25 percent of the maximum possible score. (Only the height estimate was close. This may be because the suspect was of average height; people often cite known facts about the "average" man when they are uncertain.)

We then waited seven weeks and presented a set of six photographs to each witness individually under four different experimental conditions. There were two kinds of instructions: low-bias, in which witnesses were asked only if they recognized anybody in the photographs and high-bias, in which witnesses were reminded of the attack incident, told that we had an idea who the suspect was and asked to find the attacker in one of two arrangements of photographs, all well-lit frontal views of young men including the attacker and the bystander. In the unbiased picture spread all six portraits were neatly set out with about the same expression on all the faces and with similar clothing. In the biased spread the attacker was shown with a distinctive expression and his portrait was positioned at an angle [*see illustration on these two pages*].

Only 40 percent of the witnesses iden-

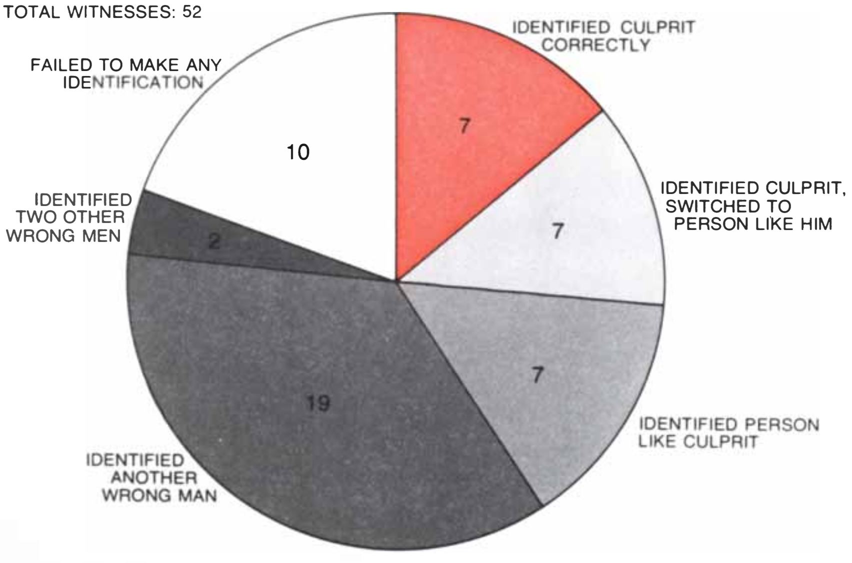
of the spreads after having been given one of two kinds of instructions: either low-bias (simply, "Do you recognize any of these men?") or high-bias (such as, "One of these men is a suspect in that assault we saw; it is important that you identify him for us"). Whereas 40 percent of all the witnesses picked No. 5, 61 percent of those who saw the biased spread and got biased instructions did so.



BIASED CONDITIONS gave observers more confidence in their ability to recognize faces in the picture layouts displayed on the two preceding pages. The bars show the degree of confidence expressed by those who picked from the unbiased spread (*left*) and the biased one (*right*) and after having been given low-bias (*color*) and high-bias (*gray*) instructions.

tified the suspect correctly; 25 percent of them identified the innocent bystander instead; even the professor who was attacked picked out the innocent man. The highest proportion of correct identifications, 61 percent, was achieved with a combination of a biased set of photographs and biased instructions. The degree of confidence in picking suspect No. 5, the attacker, was also significantly higher in that condition [*see illustration above*]. We have subsequently tested the same picture spreads with groups that never saw the original incident. We describe the assault and ask people to pick the most likely perpetrator. Under the biased conditions they too pick No. 5.

In another study undertaken at Brooklyn College a student team, led by Miriam Slomovits, staged a live purse-snatching incident in a classroom. We gave the witnesses the usual questionnaire and got the usual bad scores. This time, however, we were concerned with a specific dilemma: Why is recognition so much better than recall? In private



REAL CULPRIT in a staged purse-snatching incident was identified by only seven of 52 witnesses who viewed two videotaped lineups. Seven others picked the culprit first but then switched to a man who looked like him in the second lineup; seven picked only the man who looked like him. Most witnesses picked other people or were unable to choose.

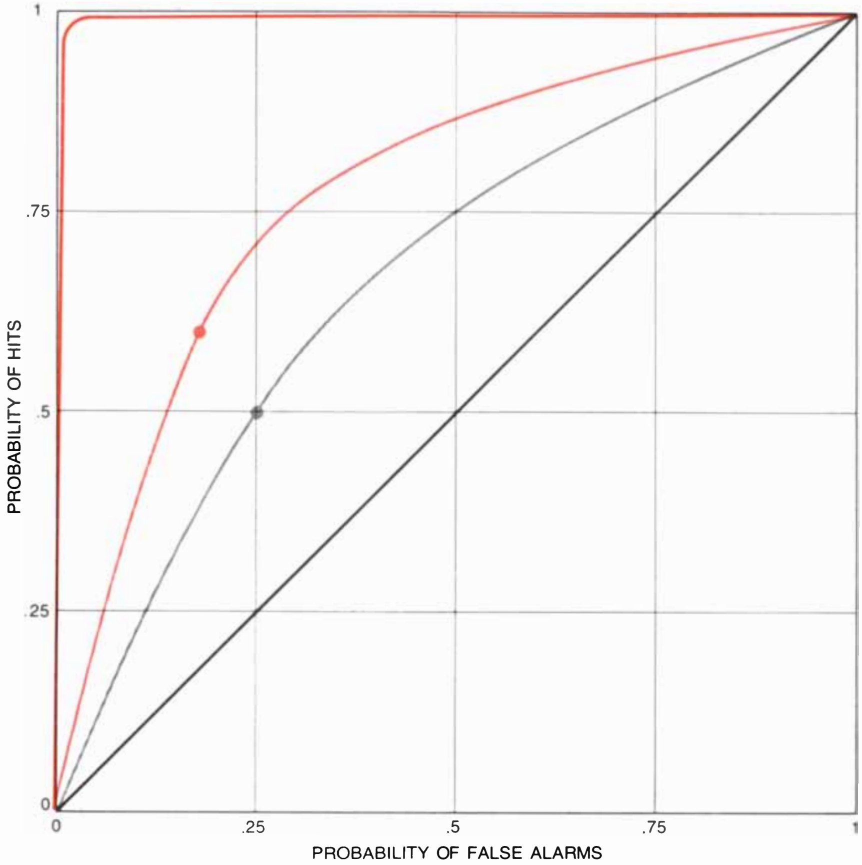
most lawyers and judges agree that the recall of a crime by a witness is very bad, but they still believe people can successfully identify a suspect. What we had to do was to break away from our demonstrations of how bad witnesses are at recalling details and search for what makes a witness good at recognizing a face. To do so we took the witnesses who had predictably given poor recall data and gave them a difficult recognition test. Our witnesses got not only a lineup with the actual purse-snatcher in the group but also a second lineup that included only a person who looked like the purse-snatcher. The question was: Would the witnesses pick only the real culprit and avoid making a mistaken identification of the person who looked like him?

We videotaped two lineups of five persons each and showed them in counterbalanced order to 52 witnesses of the purse-snatching. Very few witnesses were completely successful in making a positive identification without ambiguity. An equal number of witnesses impeached themselves by picking the man who resembled the culprit after having correctly picked the culprit. Most people simply made a mistaken identification [*see bottom illustration on this page*]. Our best witnesses had also been among the best performers in the recall test, that is, they had made significantly fewer errors of commission (adding incorrect details). They had not given particularly complete reports, but at least they had not filled in. The good witnesses also expressed less confidence than witnesses who impeached themselves. Finally, when we referred to the earlier written descriptions of the suspect we found our successful witnesses had given significantly higher, and hence more accurate, estimates of weight. People guessing someone's weight often invoke a mental chart of ideal weight for height and err substantially if the person is fat. Our purse-snatcher was unusually heavy, something the successful witnesses managed to observe in spite of his loose-fitting clothing. The others were guessing.

Once again we noted that witnesses tend not to say, "I don't know." Eighty percent of our witnesses tried to pick the suspect even though most of them were mistaken. The social influence of the lineup itself seems to encourage a "yes" response. This effect presented a disturbing problem that actually drove us back from these rather realistically enacted crimes to the more controlled, emotionally neutral environment of the laboratory. We hoped to design a test for eyewitnesses that could distinguish a good witness from a poor one, under circumstances in which we knew what the true facts were.

Pure measures of accuracy would not be adequate, since there are many different kinds of error, some of which come from the witness's desire to please the questioner with an abundance of details. Eventually we settled on adapting signal-detection theory, espoused by John A. Swets of the Massachusetts Institute of Technology, to the eyewitness situation. Signal-detection theory evolved in psychophysics as a means of coping with the fact that an observer's attitude "interferes" with his detection, processing and reporting of sensory stimuli. Limited to saying "yes" or "no" (I hear or see or smell it, or whatever), the observer applies criteria that vary with personality, experience, anticipated cost or reward, motivation to please the tester or to frustrate him and other factors. What the experimenter does, therefore, is usually to present noise about half of the time and signals plus noise about half of the time and to count correct "yes" answers (hits) and incorrect "yes" answers (false alarms), combining the scores statistically into a single measure of observer sensitivity. This quantifies an estimate of the observer's criteria for judging his immediate experience. A very cautious person might have very few false alarms and a high proportion of hits, indicating that he says "yes" sparingly; a less than cautious person might say "yes" most of the time, scoring a large number of hits but only at the price of a large number of false alarms.

In our research at Brooklyn College Lynne Williams and I now show a film of a supposed crime and then present to the observers 20 true statements about the incident and the same number of false statements. The witness indicates "yes" or "no" as to the truth of each statement. We end up with a record of hits and false alarms which, after some complicated statistical processing, yields a curve called a receiver-operating-characteristic (ROC) curve [*see illustration on this page*]. A person whose hits and false alarms were equal, indicating that the answers had no relation to the true facts, would generate a straight diagonal ROC curve. A perfect witness would have all hits and no false alarms. Real people fall somewhere in between. We have found so far that witnesses with the better (which is to say higher) ROC curves go on to do better than other people at recognizing the suspect in a lineup. We are using the ROC function to test various hypotheses about how environmental conditions, stress, mental set, bias in interrogation, age, sex, and social, ethnic and economic group affect the accuracy and reliability of eyewitnesses.



PERFORMANCE AS A WITNESS is measured by plotting "hits" against "false alarms." Hits are "yes" answers regarding the truth of a true descriptive detail of a scene the witness has viewed; false alarms are "yes" answers regarding the truth of a false detail. "Witness sensitivity curves" are generated by the results for various witnesses or for answers given by the same witness with varying degrees of confidence. A perfect witness would be one who always scored hits and never scored false alarms (*solid color curve*); a "blind" witness would score as many hits as false alarms (*black*). In the author's experiment successful witnesses, who identified a suspect correctly in a lineup, had produced curves that were higher on the chart; they averaged 12 hits to 3.6 false alarms (*light color*). Unsuccessful witnesses had produced lower curves, averaging 10 hits to five false alarms (*gray*). Scores are plotted as fractions of the maximum possible scores: 20 hits or 20 false alarms.

Psychological research on human perception has advanced from the 19th-century recording-machine analogy to a more complex understanding of selective decision-making processes that are more human and hence more useful. My colleagues and I feel that psychologists can make a needed contribution to the judicial system by directing contemporary research methods to real-world problems and by speaking out in court (as George A. Miller of Rockefeller University puts it, by "giving psychology away").

It is discouraging to note that the essential findings on the unreliability of eyewitness testimony were made by Hugo Münsterberg nearly 80 years ago, and yet the practice of basing a case on eyewitness testimony and trying to persuade a jury that such testimony is superior to circumstantial evidence continues to this day. The fact is that both types of evidence involve areas of doubt. Circumstantial evidence is tied together with a theory, which is subject to questioning. Eyewitness testimony is also based on a theory, constructed by a human being (often with help from others), about what reality was like in the past; since that theory can be adjusted or changed in accordance with personality, with the situation or with social pressure, it is unwise to accept such testimony without question. It is up to a jury to determine if the doubts about an eyewitness's testimony are reasonable enough for the testimony to be rejected as untrue. Jurors should be reminded that there can be doubt about eyewitness testimony, just as there is about any other kind of evidence.

THE SEARCH FOR BLACK HOLES

Observations at the wavelengths of light, radio waves and X rays indicate that the X-ray source Cygnus X-1 is probably a black hole in orbit around a massive star

by Kip S. Thorne

Of all the conceptions of the human mind from unicorns to gargoyles to the hydrogen bomb perhaps the most fantastic is the black hole: a hole in space with a definite edge over which anything can fall and nothing can escape; a hole with a gravitational field so strong that even light is caught and held in its grip; a hole that curves space and warps time. Like the unicorn and the gargoyle, the black hole seems much more at home in science fiction or in ancient myth than in the real universe. Nevertheless, the laws of modern physics virtually demand that black holes exist. In our galaxy alone there may be millions of them.

The search for black holes has become a major astronomical enterprise over the past decade. It has yielded dozens of candidates scattered over the sky. At first the task of proving conclusively that any one of them is truly a black hole seemed virtually impossible. In the past two years, however, an impressive amount of circumstantial evidence has been accumulated on one of the candidates: a source of strong X-ray emission in the constellation Cygnus designated Cygnus X-1. The evidence makes me and most other astronomers who have studied it about 90 percent certain that in the center of Cygnus X-1 there is indeed a black hole.

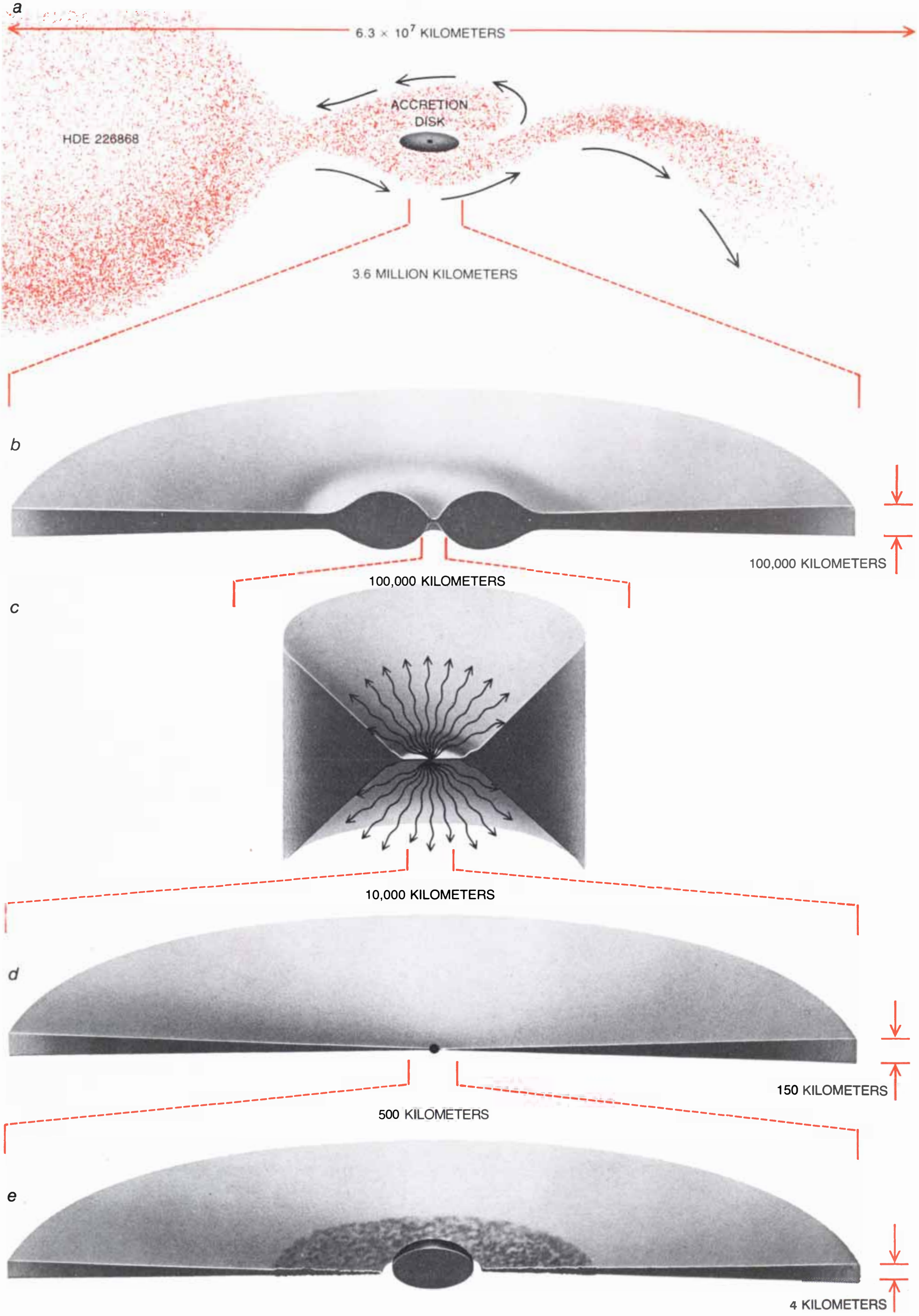
Before I describe the evidence that leads to this conclusion, let me lay some groundwork and indulge my theoretical proclivities by describing some of the predicted properties of black holes [see "Black Holes," by Roger Penrose; SCIENTIFIC AMERICAN, May, 1972]. Physicists educate themselves and their students by means of "thought experiments" whose results are predicted by theory. I shall resort to such an experiment to convey the basic reasoning that underlies the concept of the black hole.

Imagine that at some distant time in the future the human species has migrated throughout the galaxy and is inhabiting millions of planets. Having no further need for the earth, men choose to convert it into a monument: They will squeeze it until it becomes a black hole. To do the squeezing they build a set of giant vises, and to store the necessary energy they fabricate a giant battery. They then scoop out a chunk of the earth and convert its mass into pure energy, the amount of energy obtained being given by Einstein's equation $E = mc^2$, in which E is the energy, m is the mass and c is the speed of light. This energy is stored in the battery. The vises are arrayed around the earth on all sides and, powered by the battery, they squeeze the earth down to a quarter of its original size.

To check their progress the project engineers fabricate from a chunk of the earth a tight-fitting spherical jacket strong enough to hold the planet in its compressed state. They slip the jacket over the earth and open the vises. Then they measure the escape velocity of a rocket placed on the earth, that is, the velocity the rocket must attain in order to be able to coast out of the earth's gravitational field. Before the earth was compressed the escape velocity was the same as it is today: 11 kilometers per second. The compression of the earth, however, brings the earth's surface four times closer to its center, thereby quadrupling the kinetic energy that the rocket must have in order to escape. The escape energy is proportional to the square of the escape velocity; therefore the escape velocity after this first compression is doubled to 22 kilometers per second.

Satisfied that some progress has been made, the engineers repeat the process, compressing the earth still further until its original circumference of 40,000 kilometers is only 10 kilometers. I give the measurement of circumference in-

MODEL FOR A BLACK HOLE IN CYGNUS X-1 is a likely explanation for the observations made in the visible and X-ray regions of the spectrum. Gas is being pulled off the supergiant primary star HDE 226868 (*a*) by the gravitational attraction of the black hole. As the gas falls toward the black hole, the hole moves in its orbit out of the way, causing the gas to miss it. The gas nearest the black hole is whipped around it into a tight circular orbit, forming a thin accretion disk. The second illustration (*b*), at a scale of about 20 times smaller than the first, shows the expected shape of the accretion disk. The gravitational pull of the black hole compresses the disk, making it thin. At the same time thermal pressures in the gas react against the compression and try to thicken the disk. Only in the central bulge (*c*) are the pressures sufficient actually to thicken the disk. The large pressures in the bulge are caused by heat from X rays emitted near the black hole. In the core of the accretion disk (*d*) the thermal pressures are even higher than they are in the bulge; the gravity is so enormously strong, however, that it prevents the disk from thickening. The X rays observed from the earth are generated only in the innermost 200 kilometers of the core (*e*), which has the black hole itself at its center. In the innermost 50 or 100 kilometers the disk becomes translucent, violently turbulent and much hotter than it is elsewhere. The disk terminates near the black hole, where the gravitational field becomes so strong that gas can no longer move in an orbit but is sucked directly in. The termination point and the structure of the inner disk are sensitive to the black hole's speed of rotation (*see illustration on page 37*). Here it is assumed that the black hole is rotating very rapidly; if the rotation is slow, the X-ray-emitting region may be 400 kilometers in radius rather than 200.


a
6.3 × 10⁷ KILOMETERS
ACCRETION DISK
HDE 226868
3.6 MILLION KILOMETERS
b
100,000 KILOMETERS
100,000 KILOMETERS
c
10,000 KILOMETERS
d
150 KILOMETERS
500 KILOMETERS
e
4 KILOMETERS

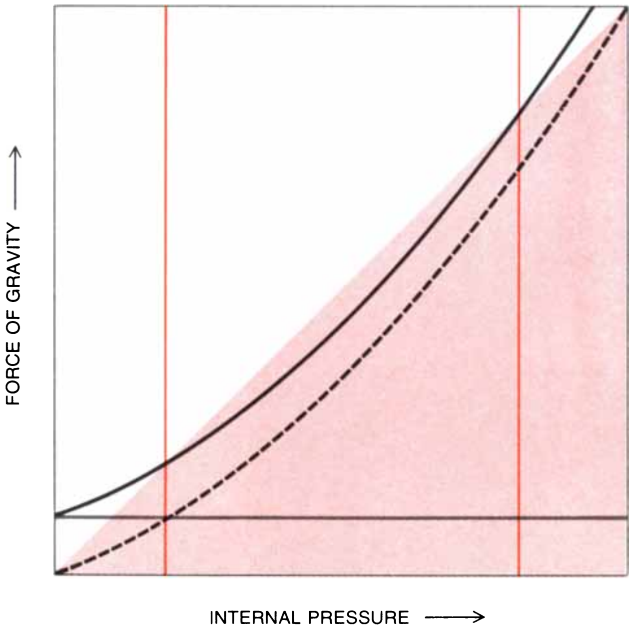
stead of diameter because in the presence of strong gravitational fields space is so highly curved that the object's diameter (d) is no longer related to its circumference (C) by the Euclidean formula $C = \pi d$; moreover, in the case of a black hole the diameter cannot be measured or calculated. This time the rocket needs an escape velocity of 708 kilometers per second in order to coast away from the earth.

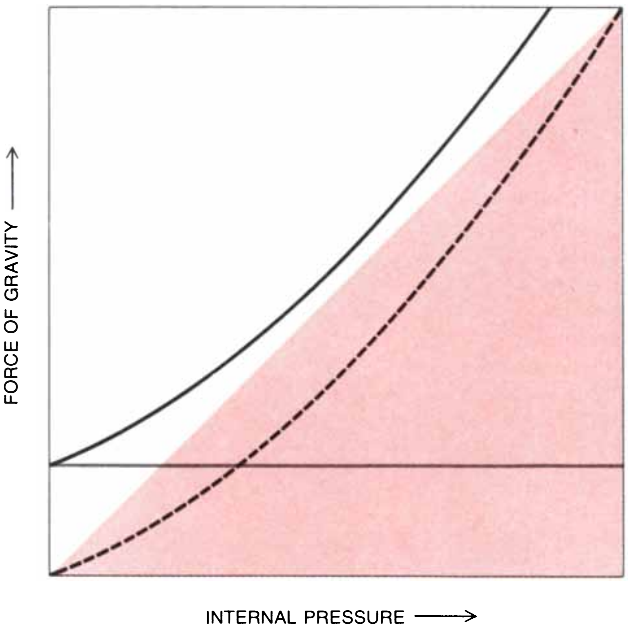
After several more compression stages the earth has been reduced to a circumference of 5.58 centimeters. The escape velocity is now 300,000 kilometers per second—the speed of light. One last little squeeze, and the escape velocity exceeds the speed of light. Now light itself cannot escape from the earth's surface, nor can anything else. Communication between the earth and the rest of the universe is permanently ruptured. In this sense the earth is no longer part of the universe. It is gone, leaving behind it a hole in space with a circumference of 5.58 centimeters. Outside the horizon, or edge, of the hole the escape velocity is less than the speed of light, and exceedingly powerful rockets can still get away. Inside the horizon the escape velocity exceeds the speed of light and nothing can escape. The interior of the hole, like the earth that gave rise to it, is cut off from the rest of the universe.

Let us return to the present and use the thought experiment to aid in understanding what happens in a star. There is a key difference between the earth and a massive star. For the earth to become a black hole external forces must be applied; for a star to become a black hole the necessary forces are provided by the star's own internal gravity. When a star of, say, 10 times the mass of the sun has consumed nuclear fuel through its internal thermonuclear reactions for a period longer than a few tens of millions of years, its fuel supply runs out. With its fires quenched the star can no longer exert the enormous thermal pressures that normally counterbalance the inward pull of its gravity. Gravity wins the tug-of-war, and the star collapses.

Unless the star sheds most of its mass during the collapse, gravity crushes it all the way down to a black hole. If, however, the star can eject enough material to reduce its mass to about twice the mass of the sun or less, then it is saved: nonthermal pressures, such as the electron pressures that make it difficult to compress rock, build up and halt the collapse. The star becomes either a white dwarf about the size of the earth or a neutron star with a circumference of some 60 kilometers. (A neutron star is a star where matter is so dense that its electrons have been squeezed onto its protons, converting them into neutrons.) In either case, as with the earth, to convert the object into a black hole one must apply external forces—forces that do not exist in nature [see "Gravitational Collapse," by Kip S. Thorne; Scientific American, November, 1967].

These predictions, which follow from the standard laws of physics, tell us that there is a critical mass for compact stars (stars with a circumference smaller than the earth's) of about two times the mass of the sun. Below the critical mass a compact star can be a white dwarf or a neutron star. Above the critical mass it can only be a black hole. The magnitude of the critical mass is a key link in the arguments that Cygnus X-1 is a black hole.

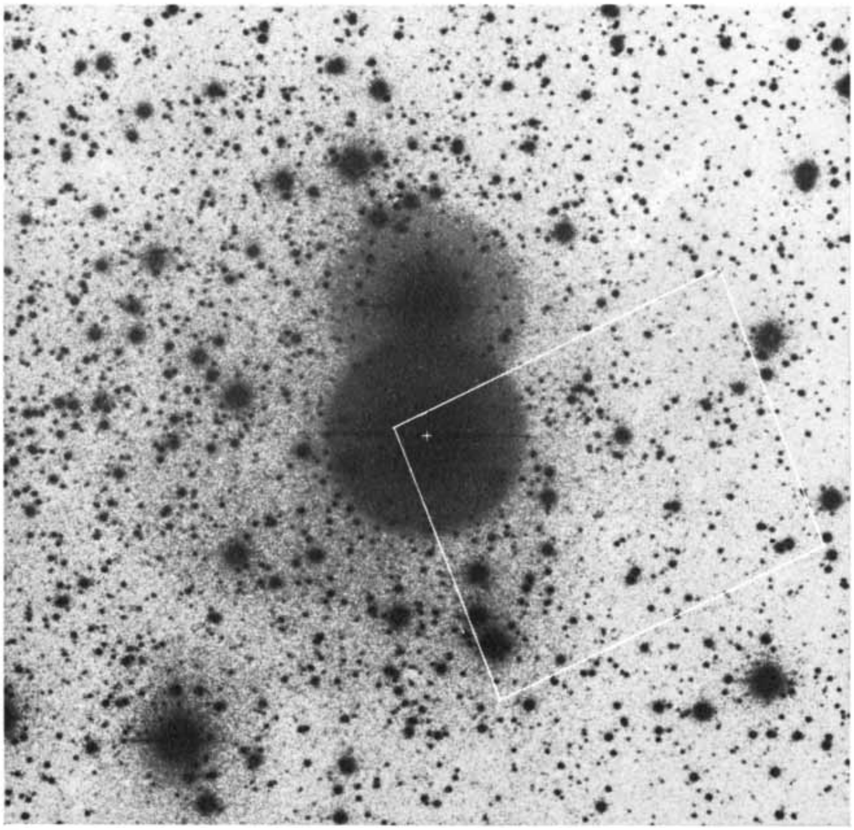




ONLY MASSIVE STARS BECOME BLACK HOLES, as is shown by these diagrams for a star with twice the mass of the sun (*left*) and a star with four times the mass of the sun (*right*). For a star to be stable the total force of gravity (*solid black line*) pulling a star's surface inward must be balanced by the star's internal pressures (*region in light color*) pushing the surface outward. If the force of gravity is stronger, the star collapses; if the internal pressure is stronger, the star explodes. For a compact star (a star smaller than the earth) the internal pressure is large enough to counterbalance gravity if the star's mass is less than three times the mass of the sun. The result is a stable white-dwarf star or a neutron star (a star in which matter is so dense that its electrons have been squeezed onto its protons, converting them into neutrons). If the star is more massive than three times the mass of the sun, however, and has a density greater than the density of the atomic nucleus, the internal pressure actually works against it. According to the general theory of relativity, gravity is produced not only by mass (*gray line*) but also by pressure (*broken black line*). At high pressures the force of gravity generated by the pressure is proportional to the square of the pressure. Thus if the star's internal pressure is very high, it gives rise to gravitational forces that overwhelm the internal pressure and the star collapses. For a compact star with a mass of less than three times the mass of the sun (*left*) there are intermediate pressures that can be counterbalanced by gravity (*vertical lines in color*). For stars more massive (*right*), however, force of gravity always wins and crushes star into a black hole in less than a second.

Therefore one would like to know the critical mass precisely. Precision is not possible, however, because we do not know enough about the properties of matter at the "supernuclear" densities of a white dwarf or a neutron star, that is, at densities above the density of the atomic nucleus: 2×10^{14} grams per cubic centimeter. Nevertheless, an upper limit on the critical mass is known: Remo Ruffini of Princeton University and others have shown that it cannot exceed three times the mass of the sun. In other words, no white dwarf or neutron star can have a mass greater than three times the mass of the sun.

From a physical and mathematical standpoint a black hole is a marvelously simple object, far simpler than the earth or a human being. When a physicist is analyzing a black hole, he need not face the complexities of matter, with its molecular, atomic and nuclear structure. The matter that collapsed in the making of the black hole has simply disappeared. It exerts no influence on the hole's surface or exterior. It makes no difference whether the collapsing matter was hydrogen, uranium or the antimatter equivalents of those elements. All the properties of the black hole are determined completely by Einstein's laws for the structure of empty space.



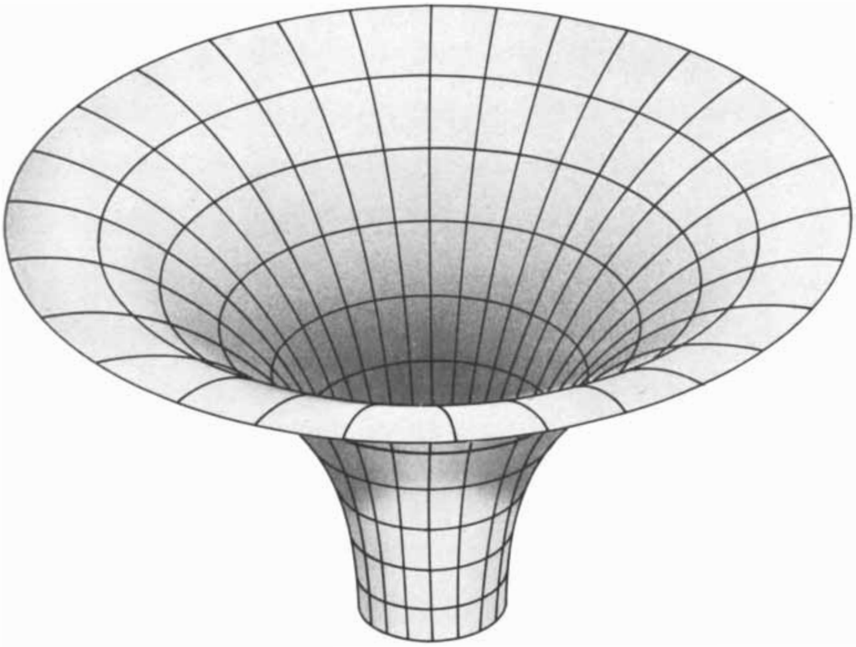
CYGNUS X-1 is believed to be associated with the star HDE 226868, the darkest and largest object in the center of this negative print of a small region of the sky. The photograph was made by Jerome Kristian with the 200-inch reflecting telescope on Palomar Mountain. The photographic plate was exposed so long that exceedingly faint stars (down to the 22nd magnitude) are visible. Superimposed on photograph are location of sources of radio emission (*small cross*) and of X-ray radiation (*white outline*). Position of Cygnus X-1 is not known very well from X-ray observations alone because X-ray telescopes have low resolution. During last week of March and first week of April in 1971, however, Cygnus X-1 underwent a cataclysmic and so far permanent change that caused it to begin emitting radio waves and to double the average energy of its X rays. Because the positions of radio sources can be measured accurately, the change in Cygnus X-1 assisted astronomers in identifying its location.

Exactly how simple black holes must be has been discovered by three physicists: Werner Israel of the University of Alberta and Brandon Carter and Stephen Hawking of the University of Cambridge. They have shown that when a black hole first forms, its horizon may have a grotesque shape and may be wildly vibrating. Within a fraction of a second, however, the horizon should settle down into a unique smooth shape. If the hole is not rotating, its shape will be absolutely spherical. Rotation, however, will flatten it at the poles just as rotation slightly flattens the earth. The amount of flattening and the precise shape of the flattened hole are determined completely by its mass and its angular momentum (speed of rotation). The mass and angular momentum not only determine the hole's shape; they also determine all the other properties of the hole. It is as though one could deduce every characteristic of a woman from her weight and hair color.

In calculations the angular momentum is replaced by a more convenient quantity: the rotation parameter. The rotation parameter (a) is equal to the speed of light (c) multiplied by the angular momentum (J), divided by the Newtonian gravitational constant (G) times the square of the hole's mass (m^2): ($a = cJ/Gm^2$). The rotation parameter always has a value between zero and one. For a rotation parameter of zero ($a = 0$) the hole is spherical and does not rotate; for a rotation parameter of 1 ($a = 1$) the hole is highly flattened and rotates extremely rapidly. There is no way to make a hole rotate any faster than $a = 1$; in fact, a hole with a rotation parameter very close to 1 should actually slow down until its rotation parameter is about .998 because of friction with the matter and radiation falling into it.

What are the most important properties that can be deduced from a hole's mass and rotation parameter? First, the gravitational field of the hole obeys the standard laws of Newton and Einstein: the hole's attraction for an object is proportional to its mass and inversely proportional to the square of the distance between it and the object if the distance is somewhat greater than the size of the hole. Second, a rotating hole creates a vortex in the empty space surrounding it, thereby swirling all particles or gas that approach it into whirlpool orbits. The greater the hole's rotation parameter, the stronger the vortex. Third, a black hole curves space and warps time in its vicinity. Fourth, a black hole has a clearly delineated horizon into which anything can fall but from which nothing can emerge. Fifth, the circumference of a black hole's equator is 19 kilometers multiplied by the mass of the hole and divided by the mass of the sun. Typical black holes should have masses between three and 50 suns, and circumferences between about 60 and 1,000 kilometers.

Such are the predicted properties of black holes from the viewpoint of the theorist. To the observational astronomer these properties present an exciting challenge: find a black hole and verify the predictions! Until the mid-1960's no one took the challenge seriously. Black



SPACE IS CURVED in the presence of a strong gravitational field. The curvature is depicted in what is known as an embedding diagram, in which three-dimensional space is represented by a flat plane that is warped by the presence of a star or any other kind of massive object. The amount of curvature is related to the strength of the object's gravitational field, and it affects the direction and travel time of rays of light and the measurement of distances.

holes were regarded as being strictly theoretical objects: objects that could be formed by the death of a star but probably never were formed. Even if they were, they could probably never be found observationally. Objects such as black holes and neutron stars were too bizarre to fit naturally into our tranquil universe. Somehow all massive stars would eject most of their mass before they died, thereby saving themselves the fate of becoming neutron stars or black holes. This climate of opinion was rarely verbalized explicitly, but it set the tenor of the times. I do not know of a single proposal to search for black holes before 1963.

In the 1960's, however, our view of the universe began to change radically. Exploding galaxies, rapidly varying radio galaxies, quasars, cosmic microwave radiation from the "big bang" explosion that formed the universe, flaring X-ray stars—all these and other observational discoveries taught us how violent and strange the universe can be. Gradually neutron stars and black holes began to seem more plausible. Then, in 1967, pulsars were discovered, and by late 1968 they were shown to be rotating neutron stars beaming radiation out into space. Since neutron stars really existed, then surely black holes must exist as well.

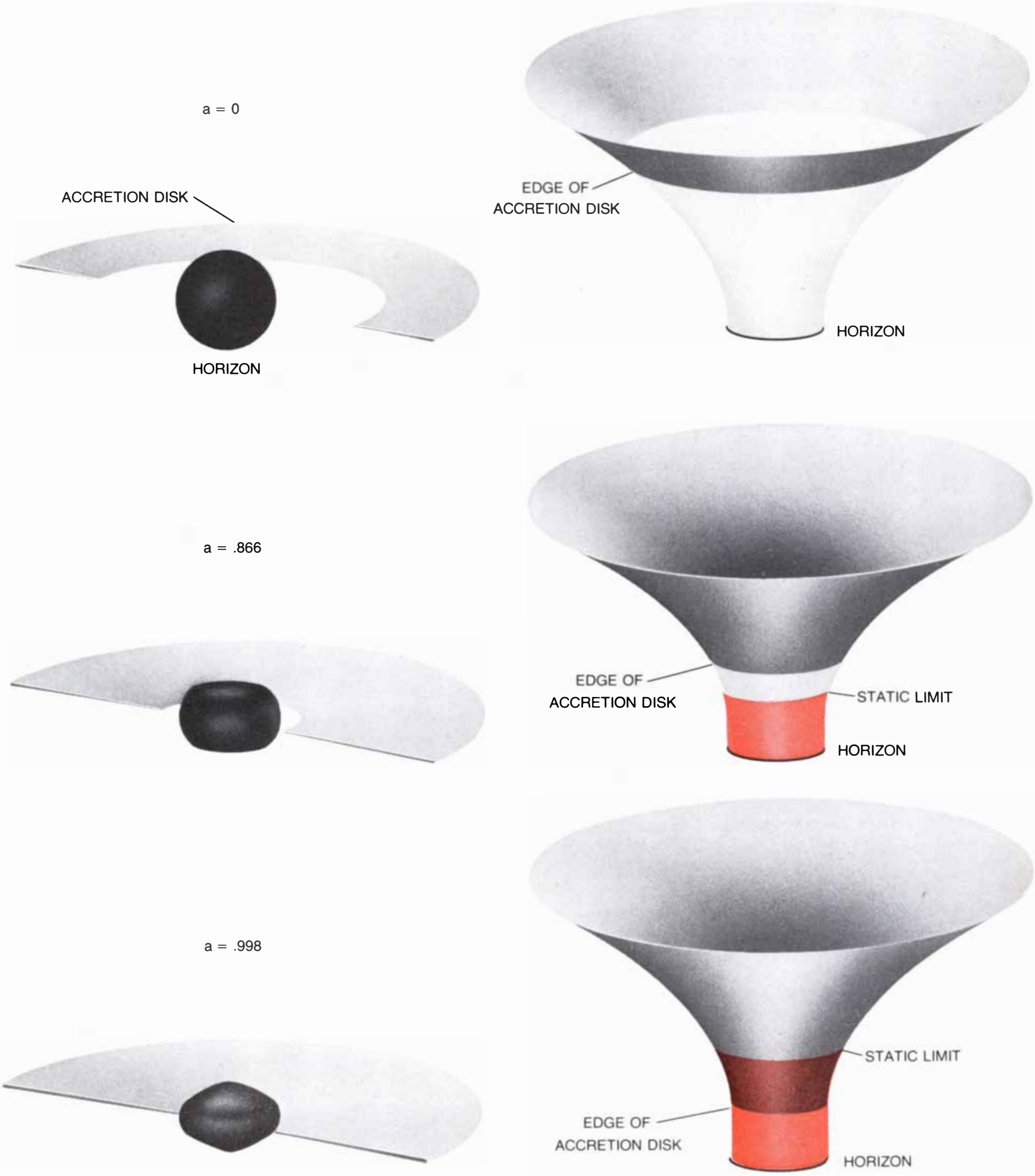
How could one go about searching for a black hole? If black holes are formed by dying massive stars, the nearest black hole should be no closer to the solar system than the nearest massive star: some 10 light-years away. Since most black holes would have a circumference of less than 1,000 kilometers, their resulting angular diameter in the sky would be a millionth of a second of arc. One could certainly not hope to find a black hole as a black spot in the sky.

Could one take advantage of the fact that a black hole's gravitational field can act as a lens and bend and focus light from a more distant star, thereby making the star look temporarily bigger and brighter? This is not a good way to search for black holes. If the hole were close to the star, the amount of focusing would be too small to be noticeable. If the hole and the star were widely separated, the amount of focusing would be large, but interstellar distances are so vast, that the necessary lining up of the earth, the hole and the star would be an exceedingly rare event—so rare that to search for one would be a waste of time. Moreover, even if such an event were observed, it would be impossible to tell whether the gravitational lens was a black hole or merely an ordinary but dim star.

Suppose the black hole had a companion star that could be seen and studied. Perhaps the presence of the hole could be deduced by its influence on the companion. With this idea in mind two Russian astrophysicists began the first search for black holes in 1964. Ya. B. Zel'dovich and O. Kh. Guseynov looked through catalogues of spectroscopic binary stars for systems that might be a black hole and a normal star revolving around each other. A spectroscopic binary system can look like a single star when it is viewed through even the most powerful telescope. The lines in its spectrum, however, shift periodically from the blue toward the red and back again as the observed star revolves around its darker companion. The periodic shift is produced by the Doppler effect: it is toward the blue as the star moves toward us in its orbit and toward the red as it moves away. The spectral lines of the companion may also be detected, shifting toward the red as the lines from the primary, or brighter, star are shifting toward the blue. In that case the companion is presumably not a black hole. If the companion star cannot be detected, however, it might be a black hole.

Star catalogues are full of binary systems for which the spectral lines of only one star are detected. Several hundred are known, and probably thousands more could be discovered if there were strong reasons to search for them. To shorten the list Zel'dovich and Guseynov investigated the mass of each dark companion. They could estimate the mass roughly from the amount the spectral lines were Doppler-shifted. The more massive the companion, the stronger its pull on the primary star and hence the greater the Doppler shift of the spectral lines. By requiring firm evidence from measurements of the Doppler shift that the mass of the dark companion is three times greater than the mass of the sun (and is therefore not a neutron star or a white dwarf), Zel'dovich and Guseynov brought their list down to a handful of spectroscopic binaries. In some of those systems the primary star was so bright that its dark secondary companion could very well be a normal star masked by the glare of the primary. After discarding those cases Zel'dovich and Guseynov were left with five good candidates for systems incorporating a black hole.

In 1968 Virginia Trimble, who was then working at the California Institute of Technology, and I revised and extended the Zel'dovich-Guseynov list. Unfortunately for us none of the eight good candidates on the new list we prepared presented a truly convincing case for a



THREE BLACK HOLES, their accretion disks and the curved space that surrounds them are compared to show how they are affected by their speed of rotation. A hole that is not rotating (*top*) has a horizon, or edge, that is spherical. Its rotation parameter *a* is zero, where the rotation parameter is defined as being equal to the speed of light times the black hole's angular momentum, divided by the Newtonian gravitational constant times the square of the black hole's mass. A hole rotating at a moderate speed (*middle*) with a rotation parameter of .866 will be perfectly flat at the poles but will still be rounded at the equator. Rotation does not affect the size of the equatorial circumference. For a black hole rotating at high speed (*bottom*) with a rotation parameter of .998, the horizon has a shape that cannot exist in the flat Euclidean space of everyday experience. In all three black holes the hot, gaseous accretion disk that is believed to surround the black hole in Cygnus X-1 is shown. Gas spirals inward through the disk toward the horizon, heating up by friction and emitting X rays along the way. At the inner edge of the disk the hot gas plunges into the hole. A whirlpool motion in space, created by the black hole's rotation, swirls the disk inward so that its inner edge is close to the horizon if the hole is rotating rapidly. The curved space around the hole and the relative positions of the hole and the accretion disk in the curved space are shown by means of an embedding diagram [*see illustration on the opposite page*]. The static limit is the point where the whirlpool motion of space becomes irresistible. For a rapidly rotating black hole the accretion disk extends far below the static limit. X rays emitted from such a disk may have specific characteristics that would reveal whether black-hole theory is correct.

black hole. In all eight cases Trimble was able to conjure up a semireasonable explanation for why the dark companion was invisible without resorting to the hypothesis that it was a black hole. For example, the dim star might itself be a multiple-star system and thus be less luminous than its mass would indicate. Alternatively the primary star might be more luminous than it appeared to be. Or complexities in the spectrum of the primary star might mask the spectral lines of the secondary star. At that point the search for black holes in binary systems seemed to be stymied.

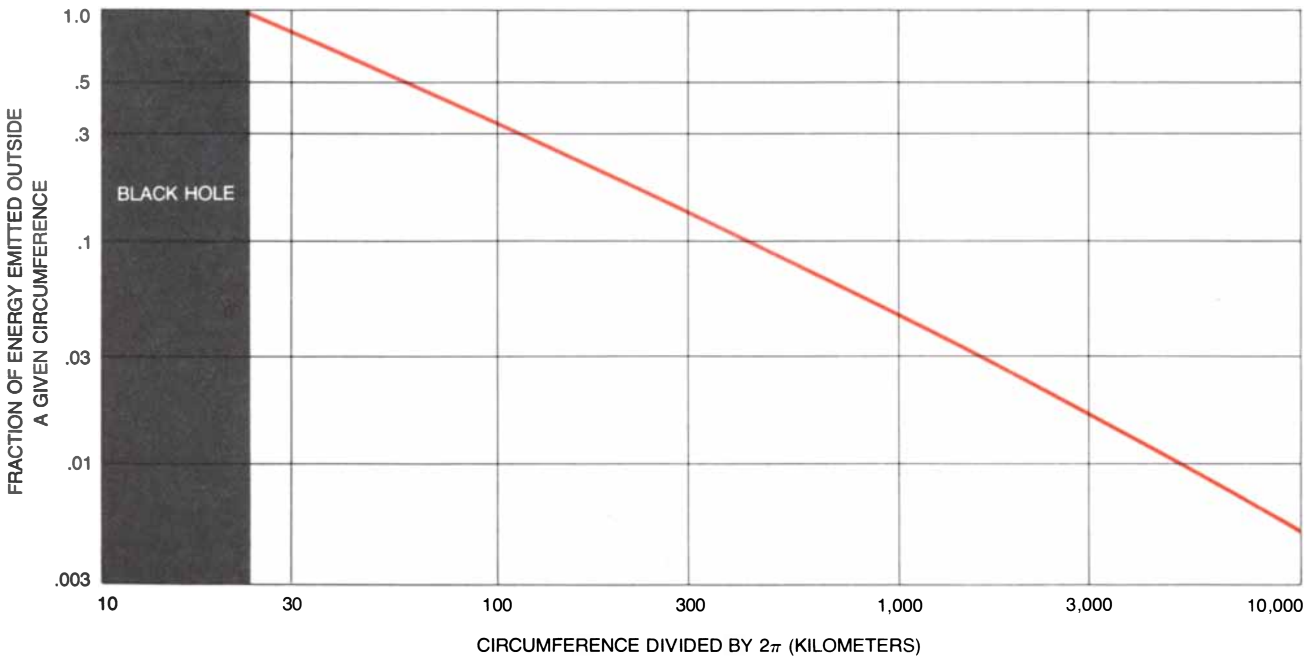
There was one major hope. As early as 1964 it had been realized that a black hole in a close binary system might pull gas off its companion star. As the liberated gas fell into the hole, it might heat up so much that it would emit X rays. Thus if any of the eight good candidates were found to emit X rays, the supposition that the dark companion was a black hole would become much more convincing.

A search for X rays emitted by binary systems could not be conducted with instruments on the ground because the X rays would be absorbed by the earth's atmosphere. One could use instruments aboard sounding rockets. Such a rocket, however, gets only a short peek at the sky before it falls back to the earth, so that its instruments can detect only the brightest of X-ray stars and can examine them only sketchily. Instruments carried aloft by balloons also get only short glimpses, and atmospheric absorption confines these views to the most penetrating of the X rays. For a definitive search an X-ray telescope aboard an artificial satellite would be needed.

The first such telescope was launched jointly by the U.S. and Italy aboard the *Uhuru* satellite on December 12, 1970. By the spring of 1972 *Uhuru* had gathered enough data to compile a detailed catalogue of 125 X-ray sources. To the astronomer searching for black holes the results from *Uhuru* were simultaneously disappointing and encouraging. They were disappointing because none of the X-ray sources coincided with any of the eight black-hole candidates. They were encouraging because at least six of the X-ray sources appeared to lie in other binary systems, typically systems that had not previously been recognized to be binary and had therefore been overlooked in the earlier searches.

Two of the six definite X-ray binary sources, Centaurus X-3 and Hercules X-1, clearly did not harbor a black hole. One could be certain of that because their X rays arrive in precisely timed periodic pulses: 4.84239 seconds between pulses for Centaurus X-3 and 1.23782 seconds for Hercules X-1. Nothing associated with a black hole can give rise to such regular behavior. Presumably each of these two binaries incorporates a rotating neutron star with its magnetic field inclined to its axis of rotation. Gas that is pulled off the companion star is funneled down the magnetic lines of force onto the magnetic poles of the neutron star, where heat from its impact generates a beam of X rays that sweeps across the sky as the star rotates. Each time the beam sweeps past the earth, the *Uhuru* satellite sees a burst of X rays. A black hole cannot produce such a beam because no off-axis structure such as a magnetic field can ever be anchored in a black hole. The hole would quickly destroy any such structure according to Einstein's laws of gravity.

The four remaining binary X-ray sources are designated 2U 1700-37, 2U 0900-40, SMC X-1 and Cygnus X-1; 2U refers to the second *Uhuru* catalogue and SMC stands for Small Magellanic Cloud, a companion galaxy of our own. Studies at visual wavelengths of each of these sources reveal a supergiant primary star with a telltale periodically varying Doppler shift. There is no sign of spectral lines from the secondary star. In all four cases, however, the visible spectrum shows lines emitted by gas flowing from the primary toward the unseen secondary. The X rays from the three systems 2U 1700-37, 2U 0900-40 and SMC X-1 are eclipsed each time the primary star passes between the earth and the unseen secondary. Therefore the secondary is almost certainly the source of the X rays. The X rays are most likely



ENERGY EMITTED by various regions of the core of the accretion disk around the black hole in Cygnus X-1 is expressed as a fraction of the total energy. Location on the disk is described in units of circumference divided by 2π, which is not equal to radius because space is highly curved in neighborhood of black hole. Half of energy is emitted from innermost 56 kilometers of disk, where temperature is higher than 30 million degrees Kelvin and energy of a typical X-ray photon exceeds 5,000 electron volts.

generated when the flowing gas is heated by falling into the secondary, just as they are when the gas falls on a neutron star; at least astronomers have not been able to invent any other quantitative explanation for the X rays.

To heat the falling gas to the temperatures necessary for the emission of X rays requires huge amounts of energy, energy that can be supplied only by a drop through a very strong gravitational field. The gravitational field surrounding a normal star or planet is not strong enough. Only three types of object have sufficiently strong fields: white dwarfs, neutron stars and black holes. And the only definitive way to distinguish among these three possibilities is to somehow measure the mass of the secondary star. If the mass exceeds 1.4 times the mass of the sun, the object cannot be a white dwarf. If the mass exceeds three times the mass of the sun, the object cannot be a neutron star. In the latter case it must be a black hole.

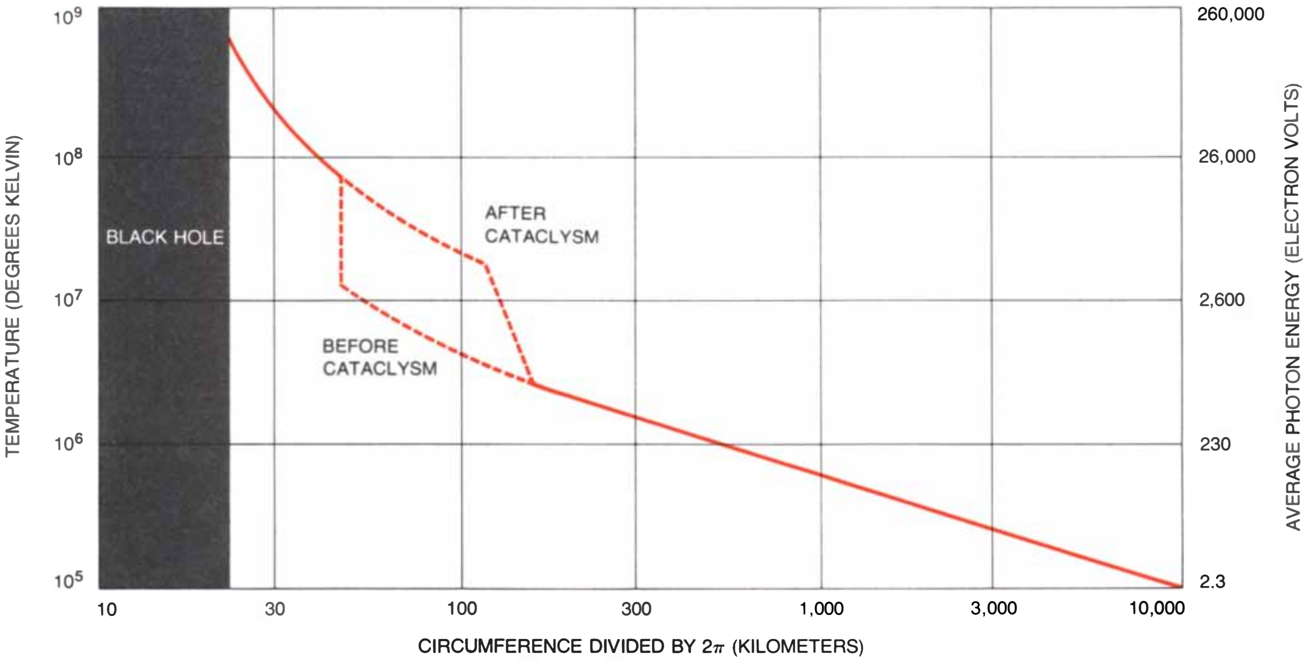
Obtaining a rough estimate of the mass of the secondary is not too difficult. One needs only a moderate amount of data from observations of the Doppler shift and fairly good information about the spectrum of the primary. On the basis of such data the mass of the unseen companion of 2U 1700-37 is 2.5 times the mass of the sun, the mass of the companion for 2U 0900-40 is three times the mass of the sun, the mass of the companion for SMC X-1 is twice the mass of the sun and the mass of the companion for Cygnus X-1 is about eight times the mass of the sun.

These figures suggest that at least one and perhaps all four of the X-ray binaries include a black hole. The estimates, however, are quite rough. Individual astronomers interpreting the data in various ways can differ by a factor of two or more in their estimates of masses. And any astronomer who wants to play the devil's advocate can reduce the estimates still further by introducing peculiar interpretations of the data.

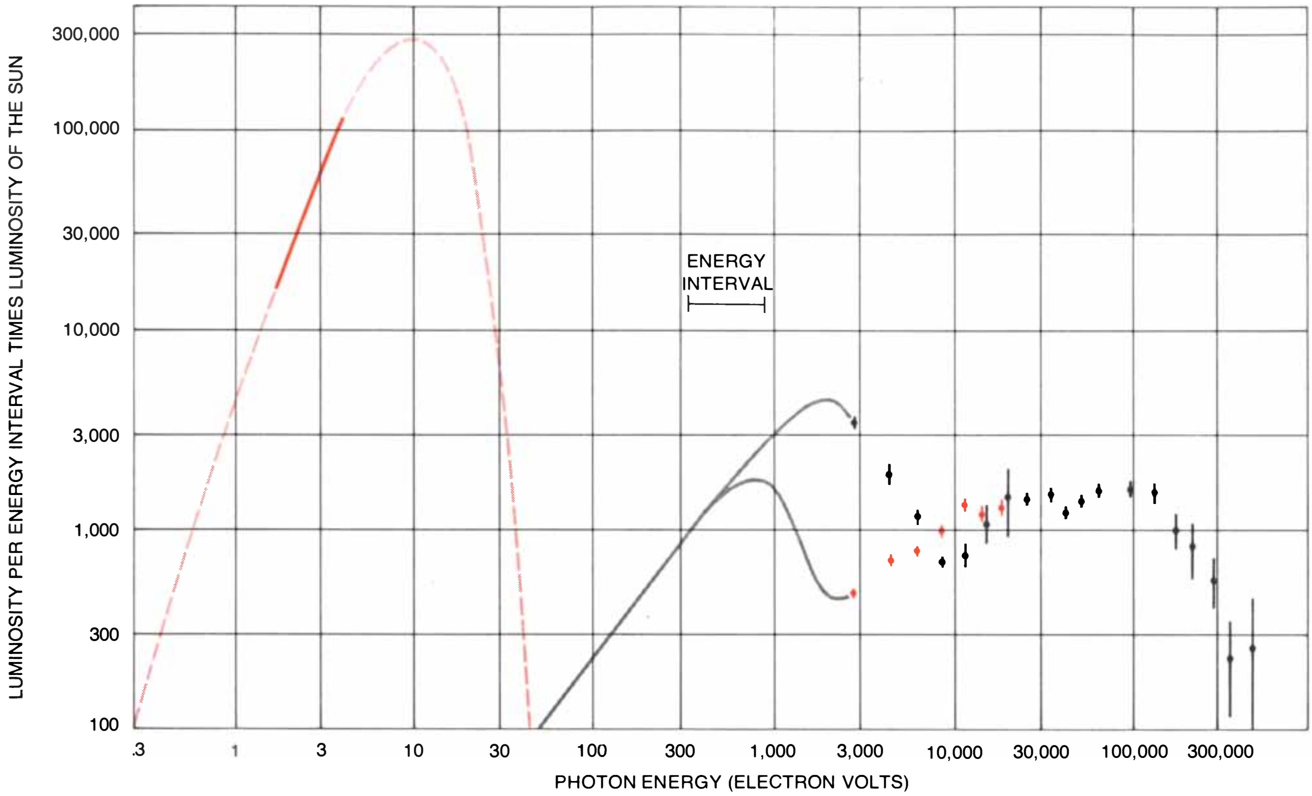
Only in the case of Cygnus X-1 do the devil's advocates face difficulties. Even a "worst-case" analysis of these data reveals that the unseen companion has a mass of no less than four times the mass of the sun. Therefore one can conclude that Cygnus X-1 does comprise a black hole. At least this conclusion is the most reasonable one.

Teams of devil's advocates led by John N. Bahcall of Princeton University and James Pringle of the University of Cambridge have invented viable, although less plausible, alternative models to explain the observations. The models assume that the massive secondary around which the bright primary travels is a normal but dim star. In one model the X rays come from a satellite neutron star in orbit around a massive normal secondary. In another model a neutron star emitting X rays circles in a wide orbit around an entire normal binary system. In a third model the X rays do not come from a neutron star or any other compact object at all. Instead between two normal stars stretch strong magnetic fields that are continually being twisted, tied into knots and broken as the two stars rotate. The knotting and breaking of the fields heats gas that is attached to them, and the hot gas emits the X rays. A fourth model assumes that the secondary emitting the X rays is not a black hole but is something even more exciting and bizarre: a massive "naked singularity" in the structure of space-time. Einstein's laws of gravity probably forbid the formation of naked singularities, but all attempts to prove that this is the case have failed.

We now face a situation that is common in astronomy. One model, that a black hole is in orbit around a normal star, can readily explain the observations: the light and the X rays from Cygnus X-1. The other leading models, proposing a secondary that is either a white dwarf or a neutron star, have been killed by the observations. Alternative contrived explanations, however, can still be made to fit the data. Such a situation leads astronomers into the final stage of the search for black holes: the attempt to accumulate more data of higher quality and of new types, data that (I hope!) will gradually kill the contrived models and clinch the case for a



TEMPERATURE OF GAS AND ENERGY OF PHOTONS emitted by core of the accretion disk is shown for various locations. Only X rays more energetic than 2,000 electron volts have been detected with X-ray telescopes available until recently; thus the curve for the less energetic photons is calculated from theory. Future X-ray telescopes will test theory by examining photons between 400 and 2,000 electron volts. Cataclysmic change in Cygnus X-1 may have been a change of state in the region between 50 and 100 kilometers.



SPECTRUM OF HDE 226868 AND CYGNUS X-1 is shown in terms of the absolute luminosity of the sun. The unit energy interval is the interval over which the energy increases by a factor of the base of natural logarithms, *e*, or about 2.718, for example at an energy between 1,000 electron volts and 2,718 electron volts. The curve in color shows the spectrum of HDE 226868; the solid portion of the curve is known from visual observations and the broken portion is extrapolated into the infrared and ultraviolet regions of the spectrum. The calculated emission from the core of the accretion disk of Cygnus X-1 is shown in gray. The points at energies above 2,700 electron volts are X-ray observations of Cygnus X-1 averaged over short fluctuations ranging from milliseconds to minutes; the vertical error bar on each point shows the statistical uncertainty due to the fact that the observations consisted of only a finite number of photons. Black dots are observations made before the cataclysmic change; colored dots show how the spectrum was altered.

black hole. In this final stage of the search it will be helpful to know what to look for. What kind of signatures should characterize the X rays and the light generated by a binary system whose secondary is a black hole accreting gas from a primary star?

Detailed theoretical studies of such binary systems were begun in 1971 before the observations from *Uhuru* had confirmed the fact that some binaries do emit X rays. The studies were initiated independently on the basis of Newtonian gravitation theory by Nikolai Shakura and Rashid Sunyaev of the Institute of Applied Mathematics in Moscow and by Pringle and Martin J. Rees at the University of Cambridge. Later analyses based on Einstein's general theory of relativity were undertaken by Donald Page and me at the California Institute of Technology, by Igor Novikov and Andrei Polnarev of the Institute of Applied Mathematics in Moscow and by Christopher Cunningham at the University of Washington. All of these studies reveal the same gross structure for the binary system and the flow of gas within it [*see illustration on page* 33].

The gravitational field of the black hole is continually pulling gas off of the supergiant star or out of its immediate vicinity and funneling it into orbits around the black hole. Centrifugal and gravitational forces flatten the orbiting gas into a thin disk around the black hole that is analogous to the rings around Saturn but is far larger. Once a gas filament is sucked into the disk, it would stay in orbit around the hole forever if there were no friction. Friction between adjacent gas filaments, however, forces the gas to spiral slowly inward. The inward velocity is much less than 1 percent of the orbital velocity. A few weeks or months after a filament enters the disk it has spiraled inward by several million kilometers and is approaching the disk's inner edge. There the black hole's gravity becomes irresistible. It sucks the gas filament away from the disk and, within a fraction of a second, through the black hole's horizon into the hole itself.

A by-product of the friction is heat. When a filament is first caught in the disk it might have a temperature of 25,000 degrees Kelvin, the same temperature as the surface of the supergiant primary. As the gas drifts inward through the disk friction heats it until in the last 100 kilometers of its spiraling descent it is hotter than 10 million degrees.

The hot gas radiates energy, about 80 percent of which is emitted from the inner 200 kilometers as X rays. Presumably these are the X rays detected by the telescope aboard *Uhuru*. The remaining 20 percent of the radiation, emitted from the comparatively cool outer parts of the disk, should be less energetic X rays that cannot be detected with the instruments currently available, together with ultraviolet radiation and light that would not be detectable against the glare of the supergiant star.

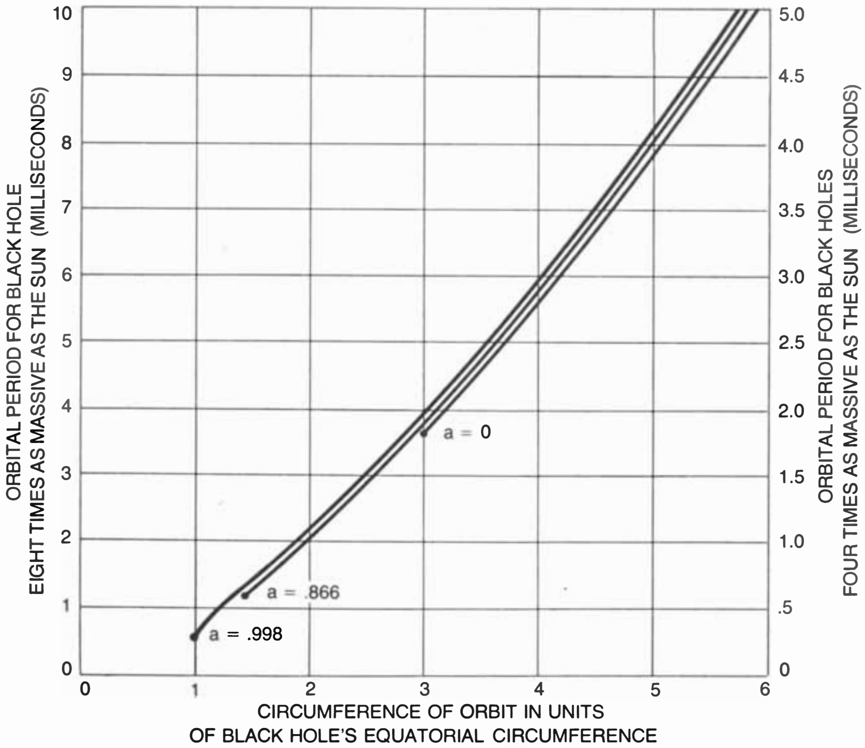
The case for a black hole in Cygnus X-1 would be much strengthened if theorists working with this model could calculate the properties of the X-ray emission in detail. The chief stumbling block at this point is friction in the disk. We do not know whether the friction is generated by turbulence in the spiraling gas, by magnetic fields embedded in the spiraling gas or by a combination of turbulence and magnetic fields. Even if we did know the source of the friction, we could not calculate its magnitude because we do not yet know enough about the general physical behavior of turbulent magnetized gases.

It is remarkable that in spite of our ignorance we are still able to calculate with confidence some of the important features of the disk. For example, from the laws of the conservation of energy and of angular momentum we can calculate how much energy each region of the disk radiates. We have concluded that most of the radiation must come from the hot inner 200 kilometers, no matter what the source or magnitude of the friction may be. We cannot, however, calculate the temperature in that inner region or the spectrum of the X rays it should emit. Instead we must discover what the spectrum is from observations and from it infer that the temperatures range between five million and 500 million degrees K. We go back to the models and see that such high temperatures are incompatible with a calm disk having little friction or turbulence but are quite reasonable if the inner region of the disk is violently turbulent and is optically thin, that is, translucent to radiation. We then return to the observations and note that the X rays do not arrive at the earth steadily. Their intensity fluctuates by a factor of two or three or even more over any length of time from milliseconds to days. Such fluctuations are also what one might expect from a turbulent disk.

By working back and forth between theory and observations in this way Richard Price of the University of Utah and I have built up a workable description of the structure of the inner disk of Cygnus X-1. Of course, success in building such a model is not much of a positive addition to the explanation that Cygnus X-1 includes a black hole. Too much was inferred from the observations and too little was calculated from the basic assumptions or the first principles of physics. On the other hand, things could be worse. The observations might have been incompatible with any type of model that assumed that the X rays were emitted by gas flowing into a black hole. In a sense, then, the black-hole model has survived a negative test. This type of test is the chief tool by which astrophysicists prove and disprove models. When a model has withstood a variety of negative tests that have destroyed all its competitors, astrophysicists begin to take the model very seriously.

Negative tests may not be the only way to prove or disprove the black-hole explanation of Cygnus X-1. Sunyaev has suggested one positive test and other theorists are searching for more. Sunyaev's test consists in looking for brief flares in the intensity of the X rays. Such X-ray flares would presumably be generated by temporary hot spots in the inner 100 kilometers of the disk [*see top illustration below*]. We have no reliable theory that accounts for the origin and destruction of such hot spots. Nevertheless it seems likely that once a hot spot is born it would live for more than one circuit in its orbit around the black hole. The X rays from the hot spot would be beamed in a direction that rotates as the hot spot circles the black hole. The beam-

RADIATION FROM A HOT SPOT on the accretion disk of a black hole would be beamed into a wide cone that would sweep out a circle in the sky above the disk each time the spot made an orbit around the hole. As the cone sweeps repetitively past the earth X-ray telescopes would observe pulses of X rays amidst a general increase in the total intensity of the X rays received. The result would be a flare of X rays that would have a pulsed substructure.



TIME INTERVAL BETWEEN PULSES emitted by a hot spot in orbit on the accretion disk around a black hole is uniquely and precisely determined by the circumference of the spot's orbit and by the mass and rotation parameter *a* of the hole. The more rapidly the hole rotates, the smaller the inner edge of its accretion disk is, and hence the shorter the minimum time would be between pulses in an X-ray flare. Observations of such pulses in X-ray flares can provide a way of measuring the rotation parameter of a black hole in Cygnus X-1.

ing might be caused in part by the process by which the radiation is emitted and in part by focusing of the radiation in the gravitational field of the black hole. Moreover, the Doppler shift would make the X rays more intense as the hot spot approached us in its orbit around the hole, and less intense as it receded. Hence the emission from the hot spot would not arrive steadily but should arrive in bursts, with the interval between bursts equal to the orbital period of the hot spot traveling around the black hole. Thus the model predicts that short X-ray flares are likely to show a substructure of pulses with an interval between pulses of a few milliseconds.

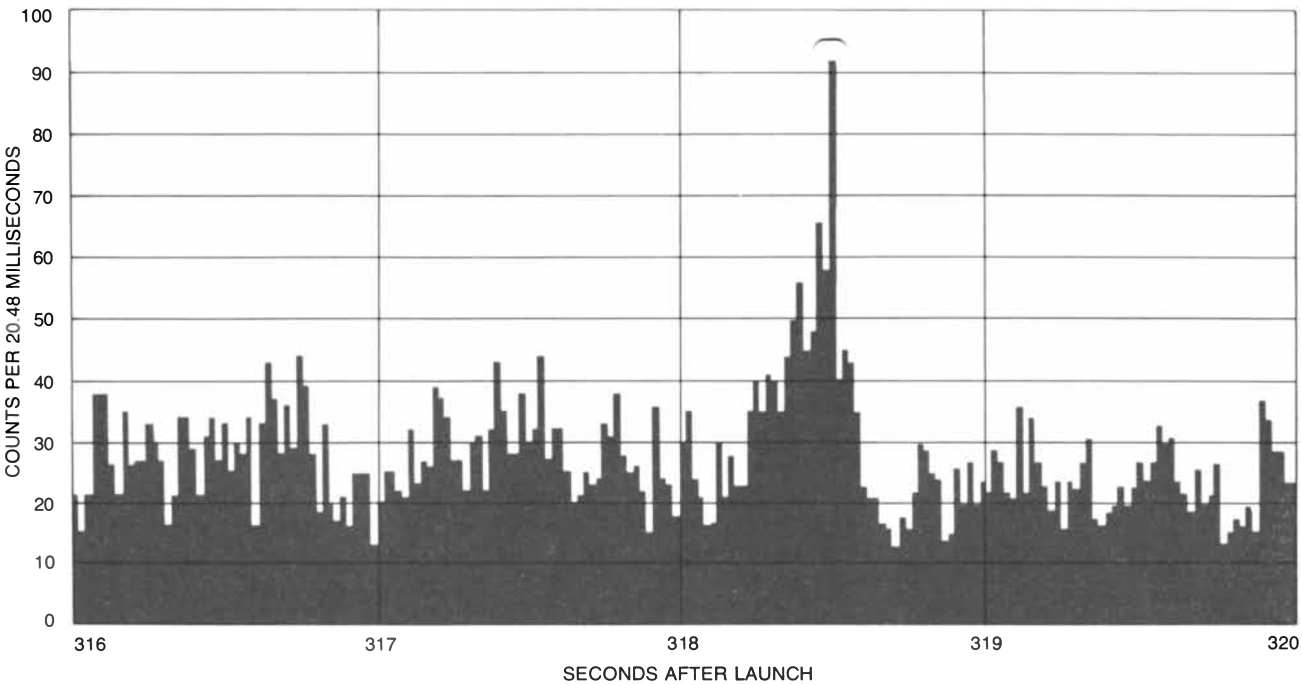
If such pulses were perceived and if the black hole's mass were known, then the interval between pulses would provide a way of computing the circumference of the hot spot's orbit around the hole. By observing many pulsed X-ray flares and determining the minimum interval between pulses we could learn what the circumference of the inner edge of the disk is, and from that we could infer the speed at which the hole is rotating. For a black hole with a mass eight times the mass of the sun the minimum interval between pulses must be between 3.6 milliseconds for a nonrotating hole (rotation parameter $a = 0$) and .6 millisecond for a rapidly rotating hole (rotation parameter $a = .998$).

Such a pulsed substructure is not, however, obligatory for X-ray flares. Flares without a pulsed substructure could originate in hot spots that are far enough from the black hole (perhaps more than 100 kilometers) so that they are not beamed by the hole's gravitational field and do not vary much because of the Doppler shift. They could also originate in spots that are very large and quickly get strung out into a doughnut around the disk by the orbital motion of the gas. Thus the observation of nonpulsed flares is also quite compatible with the black-hole explanation of Cygnus X-1. If pulsed X-ray flares can be detected, however, they should pulse only in the range predicted by Sunyaev's arguments.

The search for pulsed flares calls for an X-ray telescope that can make an observation in less than one millisecond and that has more than 10,000 square centimeters of area for collecting X-ray photons, so that many photons can be counted in each millisecond. Such a telescope will probably not be put into orbit around the earth until the National Aeronautics and Space Administration launches its first High Energy Astronomical Observatory satellite (HEAO-A), perhaps in 1977. Between now and then we must content ourselves with brief glimpses from the instruments on rockets and balloons.

The first such glimpse, in October, 1973, was promising. X-ray flares lasting about .1 second were observed with a telescope having a time resolution of .32 millisecond and the modest collecting area of 1,360 square centimeters. The largest of the flares does appear to have a pulsed substructure, but not enough photons were counted in the flare to be certain. So here we theorists sit, impatiently awaiting the next generation of X-ray telescopes, those of us in the "establishment" trying with great difficulty to build better black-hole models of Cygnus X-1 and those of us who are devil's advocates trying equally hard to build better non-black-hole models.

While some of us struggle with Cygnus X-1, others search elsewhere for black holes. The possibilities are plentiful, but none has yet yielded strong evidence for a hole. Several other spectroscopic binaries, including Epsilon Aurigae and Beta Lyrae, include a secondary star that has a mass of more than four times the mass of the sun and is surrounded by a huge opaque, or partially opaque, disk. As seen from the earth, the disk periodically blots out the light from the bright primary star. A. G. W. Cameron of the Harvard College Observatory and Edward Devinney of the University of South Florida have suggested that massive objects at the center of these disks might be black holes. Other astronomers find other explanations equally plausible.



X-RAY FLARE WAS OBSERVED IN CYGNUS X-1 by an X-ray telescope aboard a rocket built by the Goddard Space Flight Center and launched in October, 1973 (*left*). A closer analysis of the flare (*bracket*) itself (*right*) shows that it does appear to have a pulsed substructure (*color*) near the telescope's limit of sensitivity. (Apparent pulses in black are random fluctuations in the X-ray signal.)

The Russian astronomer V. F. Schwartzman is searching for black holes that have no binary companion. He calculates that the interstellar gas being sucked into such an isolated hole should emit light that flickers with a period of several milliseconds. Unfortunately for observers the light would not be very intense; it would have no more than 1 percent of the intensity of the sun's light if the sun were being observed at that distance. Since the nearest such black hole would be many light-years away, it would appear as a faint, rapidly flickering star. To detect the flicker and thereby make a strong case for the existence of an isolated black hole, Schwartzman needs sensitive electronics and a powerful telescope for observing at visible wavelengths. He is working at the Crimean Astrophysical Observatory, where a 240-inch telescope is nearing completion.

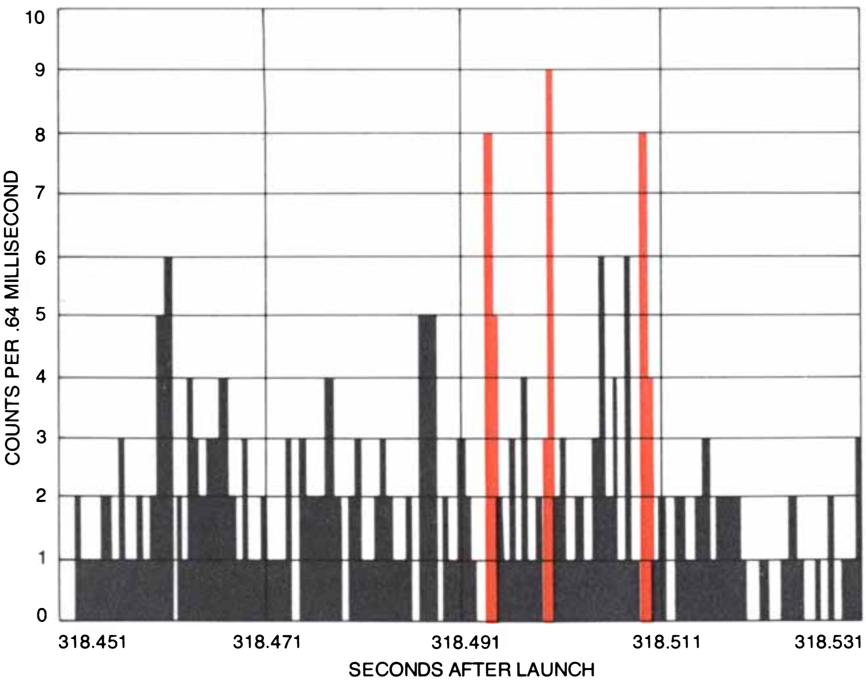
The collapse of the star that gives rise to a black hole should also generate a huge burst of gravitational waves. The present first-generation gravitational-wave antennas are only sensitive enough to detect such bursts from our own galaxy, where they would not be expected more than once every few years. Second-generation antennas now under construction at Stanford University, Louisiana State University and the University of Moscow might be able to detect such bursts from the cluster of 2,500 galaxies in the constellation Virgo. Even more sensitive third-generation antennas will surely be able to do so. Theorists expect that in the Virgo cluster there will be several black holes born every year. By detecting and analyzing the gravitational waves from such births one could not only verify the creation of a black hole but also study some intimate details of the newborn hole. It is a project a decade or so in the future.

Thus far I have described only normal black holes created by the collapse of normal stars ranging in mass from three to 60 times the mass of the sun. There are probably supermassive holes and possibly miniholes as well. Donald Lynden-Bell of the University of Cambridge has argued that the dense milieu of gas and stars that fuels grand-scale explosions in the nuclei of some galaxies must ultimately collapse to form a supermassive black hole. If this is true, a galaxy such as our own, which probably had explosions in its nucleus long ago, might possess a huge black hole in its nucleus today. That hole would be a "tomb" from a more violent past. Such a black hole in our own galaxy might be as massive as 100 million times the mass of the sun and have a circumference as large as two billion kilometers. The hole would suck gas from the surrounding galactic nucleus, perhaps forming a gigantic accretion disk analogous to the disks proposed for the spectroscopic-binary systems. Lynden-Bell and Rees calculate that such a disk would emit strong radio and infrared radiation but not X rays. The nucleus of our galaxy does give evidence of several bright infrared and radio "stars." Unfortunately for theorists an accretion disk around a supermassive black hole is not the only possible explanation of the observed objects, and so far no one has invented a definitive test for the hypothesis.

Miniholes far less massive than the sun cannot be created in the universe as it exists today. Nature simply does not supply the necessary compressional forces. The necessary forces were present, however, in the first few moments after the creation of the universe in the "big bang." If the big bang were sufficiently chaotic, then, according to calculations made by Hawking, it should have produced a great number of miniholes. Hawking has shown that miniholes behave quite differently from normal-sized holes. Any hole less massive than 10^{16} grams (the mass of a small iceberg) should gradually destroy itself by an emission of light and particles according to certain laws of quantum mechanics. Those laws, which are not important for the larger black holes, considerably modify the properties of the smaller holes. The result is that all the primordial black holes less massive than 10^{15} grams should be gone by now. Those with a mass between 10^{15} grams and 10^{16} grams are now dying. In its final death throes such a dying black hole would not be black at all. It would be a fireball powerful enough to supply all the energy needs of the earth for several decades yet small enough to fit inside the nucleus of an atom.

Hawking's results are less than a year old, and so their implications have not yet been explored in detail. They may motivate a flood of proposals for searching for miniholes. He and Page are exploring one possibility: that the bursts of cosmic gamma rays that have been detected by instruments on artificial satellites of the Vela series came from explosions of miniholes.

The present list of ways and places that black holes might be found is far from complete. With so many possibilities a theorist such as the author cannot help being excited—until he talks with his more down-to-earth experimenter friends. Then he realizes what a difficult job the search really is. We cannot expect quick results, but the future does not seem unpromising.



The time interval between pulses was no shorter than .005 second, corresponding to a circumference no smaller than four times the circumference of the black hole. That circumference is not small enough to determine the rotation parameter of the black hole.

Hydra as a Model for the Development of Biological Form

Cells isolated from this freshwater polyp can aggregate and form a complete new animal. Experiments with the system lend support to a physico-chemical scheme for the creation of biological pattern

by Alfred Gierer

The most interesting feature of a living organism, apart from its behavior, is its specific physical structure. Evolution has given rise to an untold variety of distinctive living forms, each capable of generating its own kind with high fidelity on reproduction. Some of the early biologists believed that the egg from which the organism develops must contain the plant or animal in some invisible miniature form. The actual explanation for the ability of like to produce like is a great deal more abstract. It is now known that the three-dimensional structure of every organism, in common with all other inherited properties, is encoded in the specific linear sequence of chemical subunits that constitute the genetic material DNA. In the past 20 years much has been learned about how the sequences of nucleotides in DNA we call genes are transcribed into equivalent sequences of "messenger" RNA and how these latter sequences are then translated into specific enzymes and other proteins. Biologists have also discovered mechanisms by which the transcription of genes can be turned on and off.

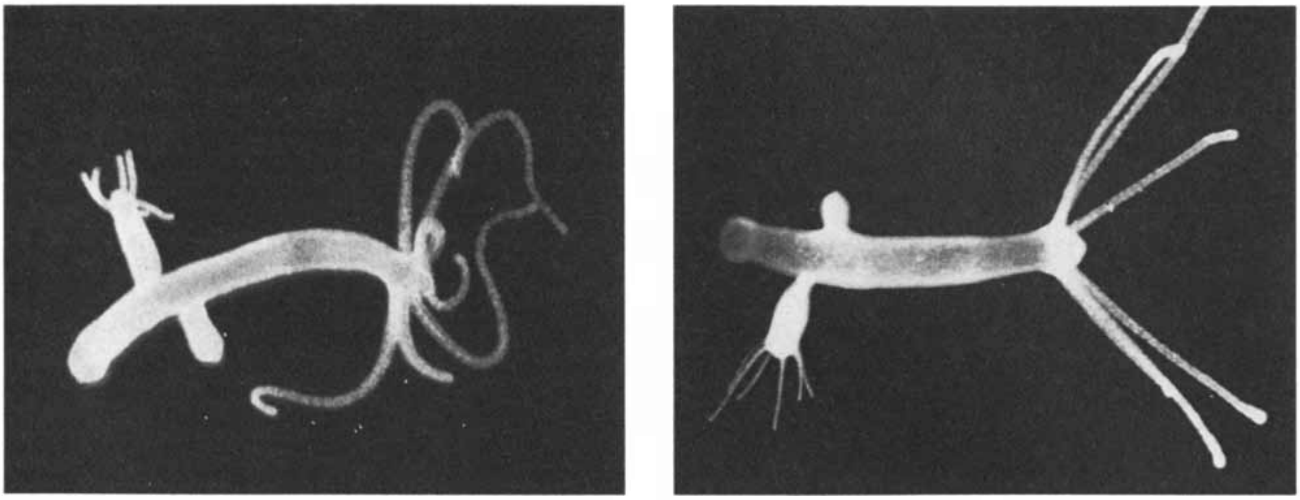
This understanding, however, has shed little light on morphogenesis: the origin and development of form. By what sequence of steps do the enzymes spun in individual cells modify the shape and function of the cells themselves and ultimately dictate the complex architecture in which thousands, millions and even billions of cells are marshaled? What are the biochemical and biophysical processes that generate macroscopic patterns and specify proportions? The task of the molecular biologist is to analyze the embryonic development of a plant or animal into combinations of elementary processes that can be explained by the physico-chemical properties of cells and molecules.

In the course of the development of a higher animal the original fertilized egg cell gives rise after a number of subdivisions to cells of all the various types needed to make muscle, gut, nerve, skin and so on. Although in general each cell is endowed with a complete copy of the genetic material, cells with distinctive properties are created by the turning on or off of particular sets of genes. Cell types differ not only in shape and size but also in composition and internal structure. Cell surfaces incorporate substances that cause cells to adhere or interact in other ways. Substances liberated by certain cells may cause other cells to move within the tissue. Often a tissue that is at first nearly homogeneous develops a complex pattern in which different regions consist of different cell types. The tissue may fold inward or outward in a precise manner to build a structure on an initially unfolded surface. These and other types of processes, ultimately directed and regulated by the genes, are the elementary mechanisms of pattern formation that one hopes to explain by molecular and cellular interactions.

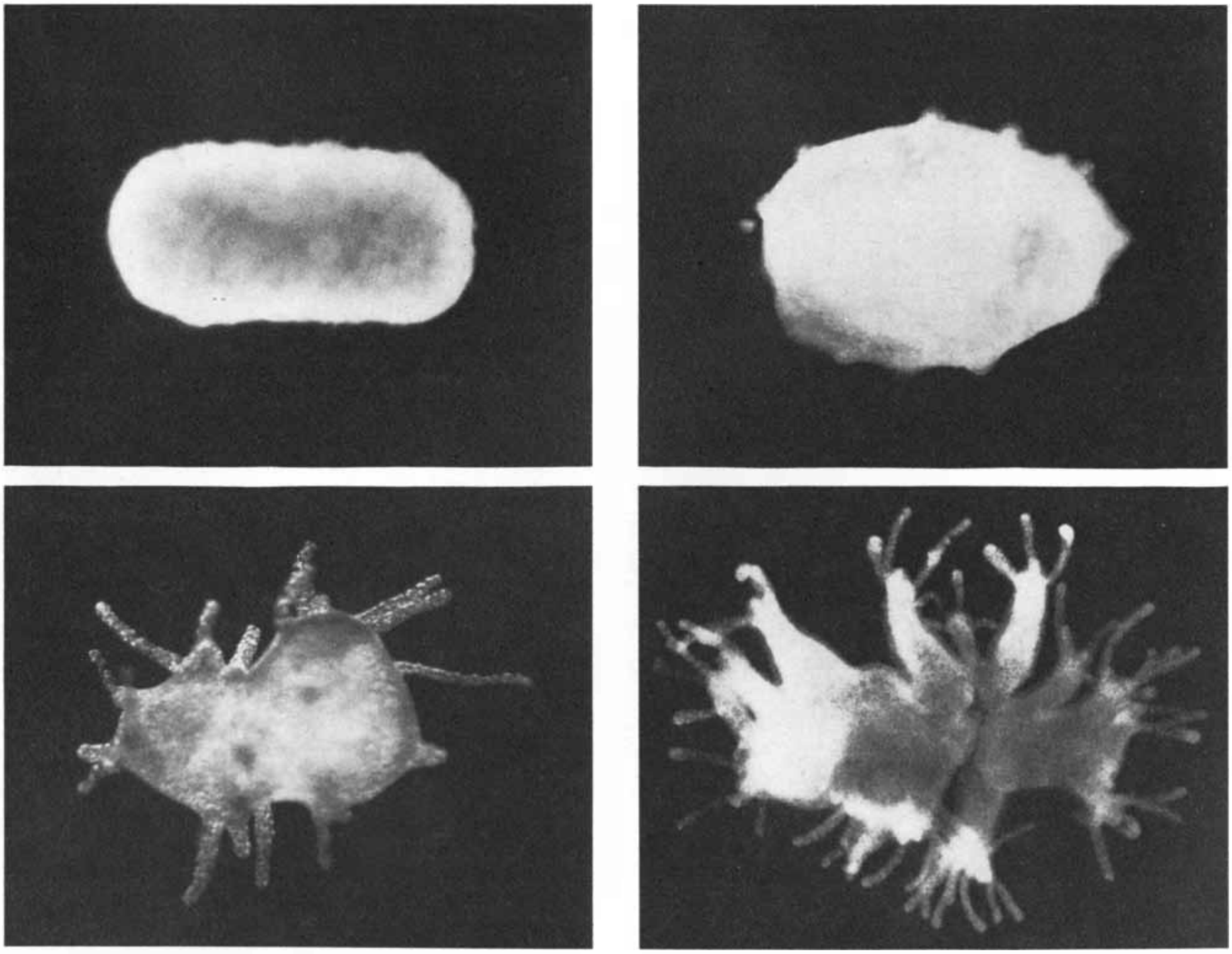
A comprehensive understanding of morphogenesis will take a long time and will draw on observations made with many kinds of organisms. A particularly suitable organism, in my view, is the freshwater coelenterate known as hydra. About half a centimeter long, hydras have a simple tissue structure that nevertheless exhibits a wide range of intracellular and intercellular phenomena, each in a relatively simple form. Hydras have few cell types and a simple three-dimensional structure [*see upper illustration on page 46*]. In addition, they have a strong capacity for regenerating lost parts, so that one can study morphogenesis not only in developing animals but also in regenerating ones. Hydras have an asymmetric polar structure, with a head (the hypostome and the tentacles) at one end and a foot (the peduncle and the basal disk) at the other. The gastric column in between has a zone where buds are formed and split off, reproducing the animal asexually. Hydras are also able to reproduce sexually from fertilized eggs.

Hydras consist of about 100,000 cells of roughly a dozen different types. The cells are marshaled in two layers, ectoderm and endoderm, both of which contain epithelio-muscular cells. Interspersed among these cells are interstitial cells that reproduce themselves and also serve as the precursors of differentiated cells, mainly nerve cells and nematocytes. The nervous system is a simple net that is densest in the head area. Nematocytes are stinging cells that migrate to the tentacles, where they impale and kill the animal's prey. Their mechanical and chemical properties put them among the most complex cells that are known to biology [*see illustration on page 53*].

Hydra has a long history as a model system for investigations in developmental biology. In 1744 Abraham Trembley observed that sections of the gastric column can give rise to a new animal. Not believing what he had seen, he repeated his observation and verified that within a few days a piece of tissue grew



HYDRAS are simple freshwater polyps, about five millimeters long, that commonly reproduce asexually by budding. These two micrographs made in the author's laboratory by Gerald Hansmann show normal hydras with two developing buds. The animals have tentacles and a mouthlike structure (the hypostome) at one end and a foot (the peduncle and the basal disk) at the other. In between is the gastric column with a budding zone at the lower end. Each of these two animals has one young bud and one older bud.

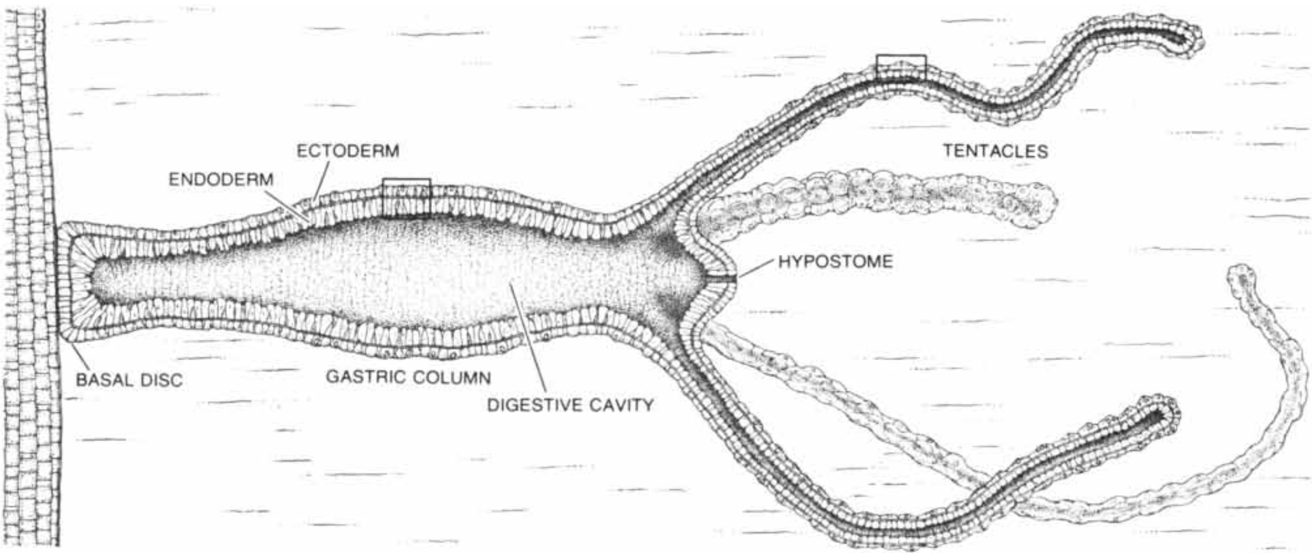


REGENERATING ORGANISM can be obtained from isolated hydra cells. At the top left centrifuged hydra cells form a clump in which interior (endoderm) and exterior (ectoderm) cells are randomly mixed. About a day later (*upper right*) ectoderm and endoderm cells have sorted themselves out, forming a hollow spheroidal structure with two cell layers. After another day tentacle buds have started to appear. Hypostomes with tentacles continue to grow (*lower left*) and in many cases create a multiheaded monster (*lower right*). Eventually the parts separate to form normal animals. Regeneration of hydras was first observed in 1744.

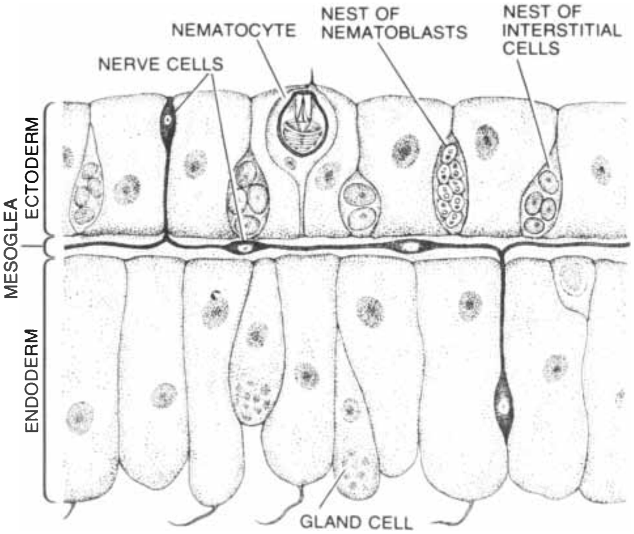
into a new specimen, complete with a head and a foot. Trembley's discovery of the formation of an entire organism from its parts, and of the organism's apparent immortality, created a sensation not only in the scientific community of the day but also in literary and philosophical circles. Since that time some 2,000 investigations conducted with hydras have been published.

As early as 1755 August Johann Rösel von Rosenhof put together even smaller pieces of hydra and observed their regeneration into complete animals. In 1935 the Russian biologist M. P. Aisupiet pressed hydra tissue through a fine mesh and obtained regenerated animals from the emerging disorganized tissue. More recently my colleagues and I at the Max Planck Institute for Virus Research in Tübingen have obtained complete hydras from dispersed cells, and the same finding has been reported by Japanese workers. As far as we know hydras are the most complex organisms that can be entirely reconstituted and regenerated from cells. The same can be done with other animals, such as sponges, but the level of organization in such animals is lower and they do not have real tissues and organs. For animals more complex than hydras organlike structures have been reconstituted from cells but not the entire animal.

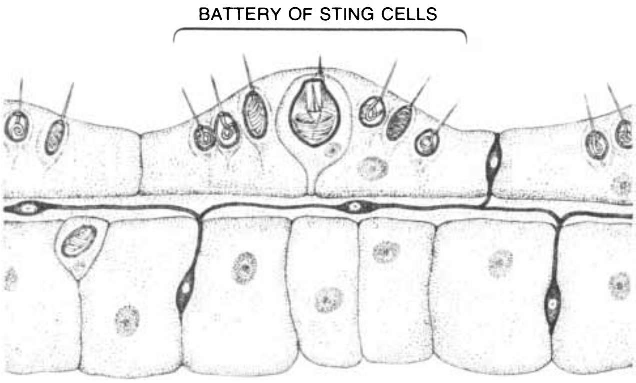
For several years a group in our laboratory has been studying isolated cells and regenerating tissues of hydras as a model for cell differentiation and pattern formation. These studies led to the finding that dissociated hydra cells first form aggregates, or clumps, before regenerating into complete animals. Even though hydras are freshwater animals, if isolated hydra cells are placed in fresh water, they disintegrate. We found, however, that isolated cells remain alive and tend to aggregate if they are placed in certain



SIMPLE ORGANIZATION OF HYDRA is depicted in this schematic cross section. It is this simplicity and the fact that hydras are made up of barely a dozen types of cell that makes these animals a particularly useful model system for the study of morphogenesis.



CROSS SECTION OF GASTRIC REGION shows the arrangement of endoderm and ectoderm. Embedded in the ectodermal layer are interstitial stem cells that not only reproduce themselves but also differentiate into nerve cells and nematoblasts (immature nematocytes).



CROSS SECTION OF TENTACLE shows batteries of nematocytes, or sting cells. Produced in the body column, nematocytes are extremely complex cells that migrate primarily to the tentacles, where their function is to impale and kill animals on which hydras prey.

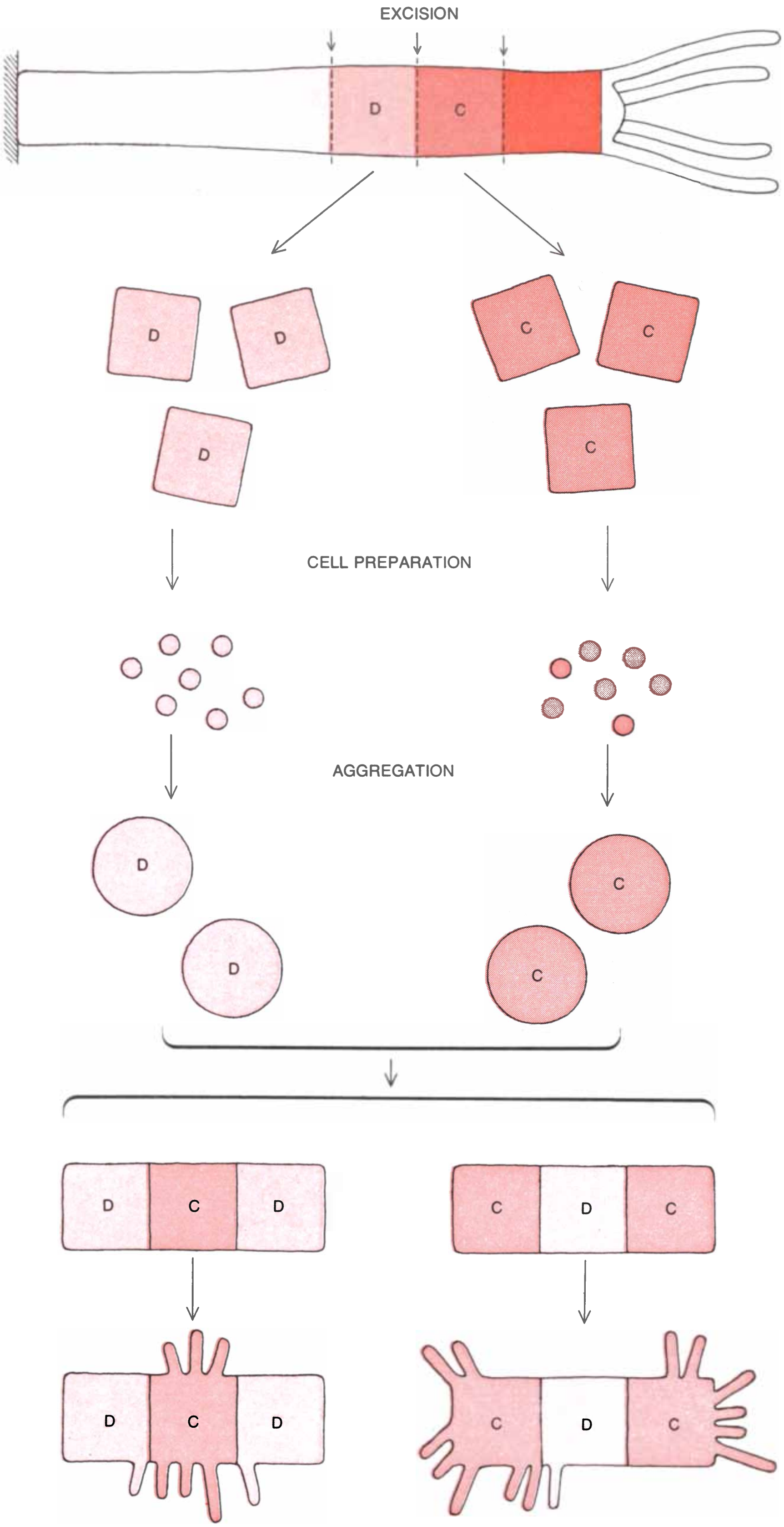
solutions of high salt concentration in which whole animals slowly disintegrate.

On the basis of these observations we are able to obtain regenerated animals in the following way. Hydras are soaked in a suitable medium of high salt concentration and then are disrupted mechanically by passing them repeatedly through a pipette until most of the tissue has disintegrated into separate cells. Remaining clumps of cells are removed by sedimentation or by passing the cell suspension through filters of the appropriate mesh size. The cell suspension, consisting predominantly of single cells, is allowed to aggregate in a dish containing a suitable medium. An alternative procedure is to spin the cells in a centrifuge until they form clumps of defined size, which are then placed in the medium. The cell clumps form a mosaic in which various cell types occupy random positions.

The medium is now diluted gradually to a reduced salt concentration in which isolated cells would disintegrate but in which hydras stay alive. In the course of a day ectodermal and endodermal cells sort themselves out. Eventually a two-layered hollow sphere forms, with ectodermal cells comprising the outer layer and endodermal cells the inner one [*see micrograph at upper right in bottom illustration on page 45*]. The structure resembles an intermediate stage of the developing hydra embryo and is very similar to a regenerating section of the body column after the wounds caused by excision have closed.

In the course of the second day tentacles appear, and soon thereafter developing hypostomes are visible. Often a monstrous many-headed animal is formed [*see micrographs at lower left and lower right in bottom illustration on page 45*]. Eventually, as the monster feeds and grows, the parts separate into normal hydras. The heads of the monster form at random positions, but they tend to maintain a certain distance from one another. If cells are initially derived not from whole animals but from head areas alone, head structures are predominantly formed; cell preparations from foot areas show a predominance of footlike structures. Cells taken exclusively from the gastric column give rise to complete animals with a head and a foot.

The crucial point in the interpretation of such experiments is to prove that single cells really participate in forming new animals. With current techniques it is impossible to remove all clumps containing more than one cell from the preparation; indeed, new cell clumps form continuously in the cell suspension. In



GRADED PROPERTY TERMED POLARITY determines the site of new head regions in hydras regenerated from separated-cell preparations. In a typical experiment cells are obtained from two gastric regions, one region (*C*) closer to the head than the other (*D*). Cell aggregates from the two regions are then arranged in different sequences and observed. It is found that heads (indicated by tentacles) develop preferentially from *C* regions.

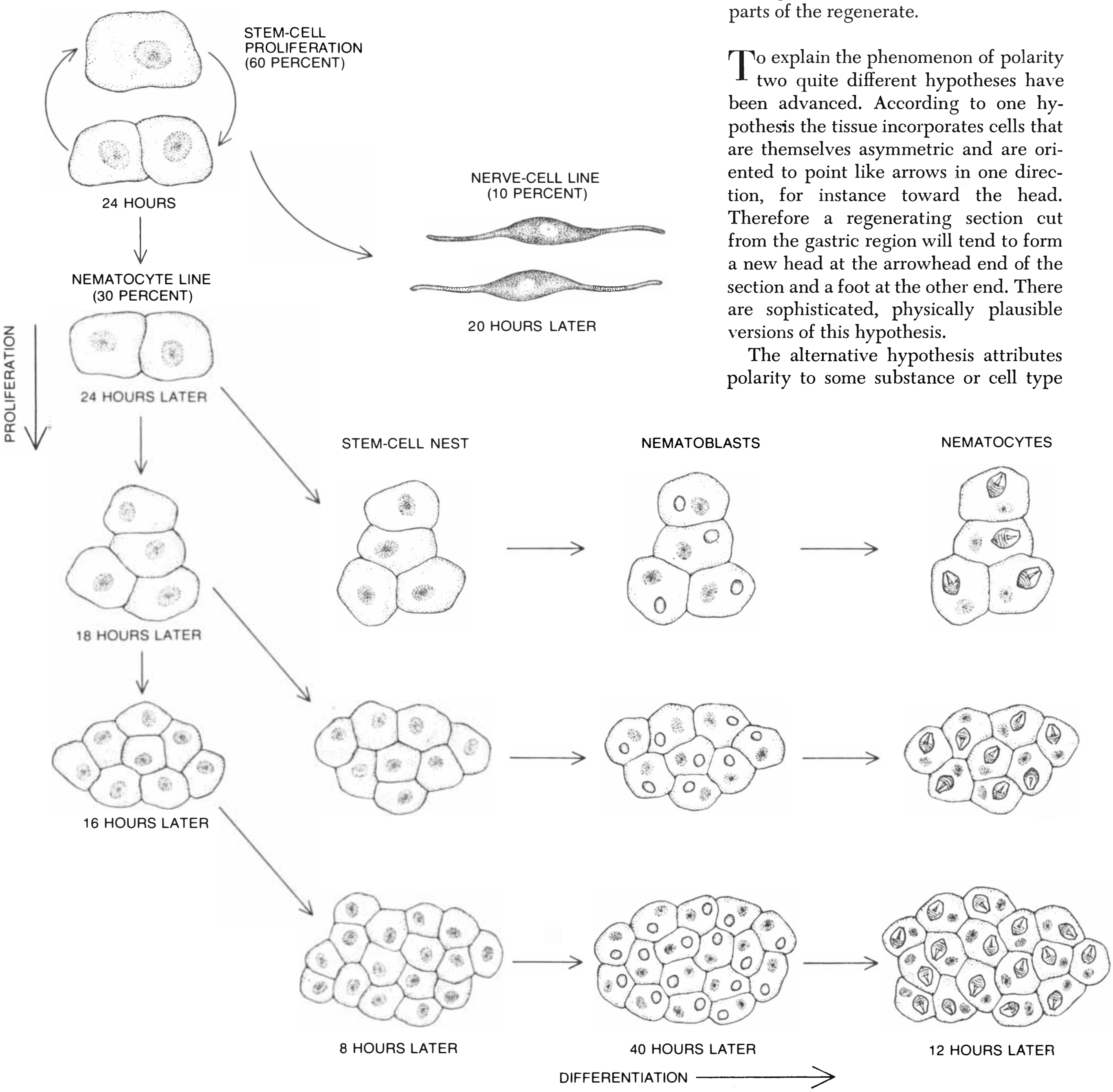
addition, regenerating hydras, whether they are derived from cells or from parts of animals, always lose some of their cells until they can be fed. It is by quantitatively comparing the number of single cells in the cell preparation and the number of surviving cells in the regenerated animals that we have been able to demonstrate that the regenerates are formed predominantly from single cells.

Apart from demonstrating the capacity of a cell system for self-organization, the aggregates from hydra cells provide an assay for the role of cells and cell types in morphogenesis. Therefore it is possible to follow the path of labeled or stained cells in the course of development. In addition, with cell aggregates one can construct artificial spatial arrangements of cells of various types, origins and pretreatments. A relatively simple example is the following.

Regenerating tissue sections cut from the gastric column of hydras show a directional property called polarity. Such sections regenerate a head at the edge closest to the original head. The decision on where to form a new head is made a few hours after cutting, long before any headlike structure is visible. If the area that will become the head of the regenerated animal is cut out and transplanted to the gastric region of another hydra, it may induce the formation of a second head in the host tissue. Visible patterns are thus preceded by prepatterns that distinguish a future head area from other parts of the regenerate.

To explain the phenomenon of polarity two quite different hypotheses have been advanced. According to one hypothesis the tissue incorporates cells that are themselves asymmetric and are oriented to point like arrows in one direction, for instance toward the head. Therefore a regenerating section cut from the gastric region will tend to form a new head at the arrowhead end of the section and a foot at the other end. There are sophisticated, physically plausible versions of this hypothesis.

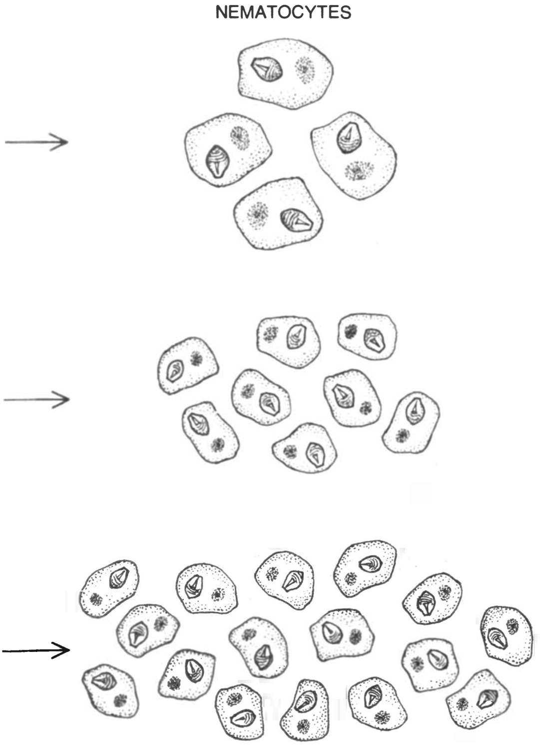
The alternative hypothesis attributes polarity to some substance or cell type



CELL DIFFERENTIATION produces nerve cells and nematocytes out of stem cells, which are one type or possibly several types of interstitial cell found singly or in "nests" of two. Each day about 60 percent of the stem cells reproduce their own kind. Another 10 percent differentiate into nerve cells and 30 percent end up as nematocytes. In the first stages, lasting two to four days, the stem cells destined to produce nematocytes simply proliferate, typically forming nests of four, eight or 16 cells. Following the last division the cells in the nests start differentiating. After eight hours the nematoblast stage is reached, showing the vacuole of the develop-

that forms a gradient between the head and the foot. Within a section cut from the gastric column the region closest to the animal's original head would contain a slightly higher level of the hypothetical substance or cell type than would the region farthest from the head. As a result of a competitive process extending over the regenerating section the future head would be initiated in the region with the "head start," that is, the slightly higher concentration of critical substance or cell type.

To examine these two hypotheses we have prepared cell suspensions from different parts of the body column, formed aggregates and immediately thereafter grafted them together in various serial orders. In this way we can see whether or not cells conserve some position-dependent property in spite of disaggregation and disorientation. We found that if aggregates made from cell preparations derived from head areas (*H*) and aggregates made from cell preparations derived from gastric-column areas (*G*) are grafted to form the array *H-G-H*, new heads will form predominantly in the two *H* regions. Similarly, serial aggregates containing head-derived cells in



ing nematocysts. After 40 hours the nematoblasts have differentiated into nematocytes containing a mature nematocyst capsule. The nests finally break up, enabling nematocytes to migrate individually to tentacles.

the central area (*G-H-G*) will tend to form heads in the middle.

This does not mean, however, that regenerating serial aggregates will invariably form structures corresponding to those in the tissues from which the aggregates were derived. I have mentioned that cells of the gastric column regenerate complete animals on reaggregation. On the other hand, if cell preparations are made from different parts of the gastric column, one from an area close to the head and another from an area farther from the head, heads form preferentially in the aggregates containing cells obtained closest to the head [*see illustration on page 47*].

An analysis of these experiments and similar ones leads us to believe that polarity is not caused by cellular asymmetry but results rather from the graded distribution of some substance or cell type. Before an aggregate is formed, all the cells have been completely disoriented. If asymmetric (polar) cells exist, they would have random orientations in the aggregates. Conceivably they could later tend to align themselves parallel to each other, but the resulting preferential direction of the cells within the aggregate would again be random. In any case, if polarity were to result from cell orientation, the location of the head-forming areas in sequential aggregates should not depend on the areas from which the cells have been derived. Our experiments show, however, that there is a strong dependence.

On the other hand, if different regions of the animal differ in their cellular constituents or in their cell types in a graded manner from head to foot, such differences would be maintained when cells are isolated and reaggregated, irrespective of the disorientation of the cells. Within the aggregates, then, one should find some regions richer in the hypothetical substances (or cells) and some regions leaner. A competitive process extending over the aggregate may cause the relatively rich sections preferentially to form heads. This model agrees with the experimental findings.

The chemical- or cellular-gradient model is also supported by another kind of experiment. Suppose the gastric column is cut into many thin sections at right angles to its long axis, without dispersion into cells, and that each section is reversed and immediately put back with its original neighbors. At which end will the new head appear? According to cell-orientation models the new head should form close to the former foot, and according to chemical- or cell-gradient models the new head should form close to the former head. The latter prediction turns out to be the right one. To be sure, more complicated models can be devised in which a stable gradient is assumed to direct a rapid orientation of cells so that some effects of cell orientation could still be involved in the explanation of the experimental results. Even then, however, the general conclusion remains that the stability of polarity in the tissues of hydra is due not to cell orientation but rather to a graded distribution of some physical property, most probably of a special substance or type of cell.

Any specialized cell in a higher organism, for example a nerve cell, arises during embryonic development from a sequence of determining steps. This sequential determination is somewhat similar to specifying a postal address by state, city, street and number. The lineages of the cells in such an organism make up a branching scheme starting from one cell type (the fertilized egg cell) and ending up with all the cell types of the organism. At each stage in the branching scheme a decision must evidently be made between production of more of the same kind of cell and further specialization along one or several pathways.

Such a process requires accurate quantitative regulation. Growing, budding and regenerating hydras offer fairly simple model systems for studying these elementary processes. Well-fed budding hydras double in number every three and a half days. The body column is in a steady state of cell proliferation and differentiation. It produces new cells continuously, partly to "repair" tentacles and foot but mostly to form the new buds.

Certain cell types, the nerve cells and the nematocytes, do not reproduce but are formed from stem cells: precursor cells that are capable both of maintaining their own production and of giving rise to differentiated cells. In nematocyte production stem cells in the interstices between cells of another type double several times to form little nests, so that nematocytes eventually emerge in groups of four, eight, 16 and 32 cells. Thus a large number of immature nematocytes are found in growing hydras. Insights into the sequence of events and the regulatory processes involved can be obtained by drawing up a balance sheet that accounts for the continuous production of various cell types.

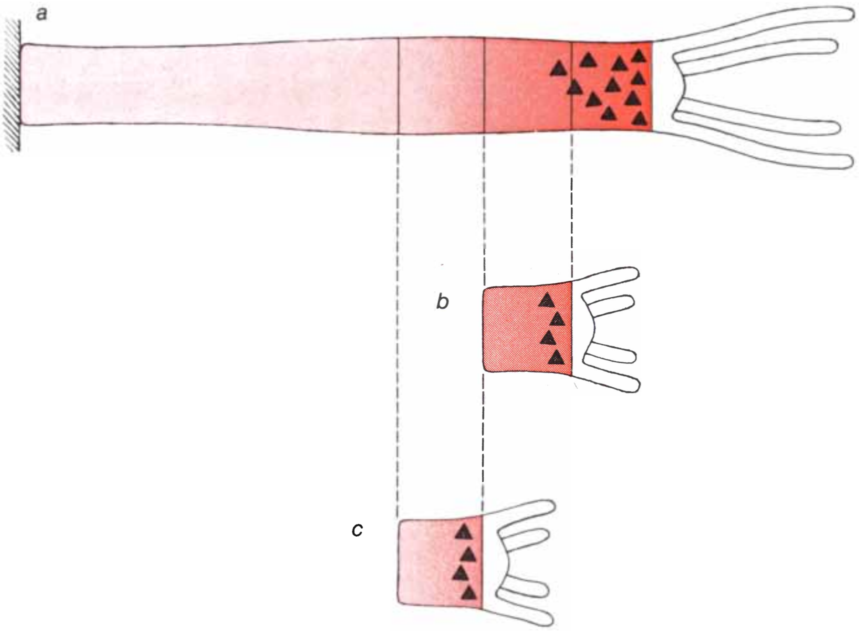
In our laboratory Charles David and Richard Campbell have examined the scheduling of proliferation and differentiation for nerve cells, nematocytes and epithelial cells. By administering a labeled precursor of DNA (radioactive thymidine) to hydras for various time intervals, they were able to measure the time required for the tracer to be incorporated into the various cell types. If incorporation is delayed for a given cell type, one can assume that the type arises from an intermediate precursor. One can then deduce how long the intermediate stage has lasted. For proliferating cells the kinetics of labeling, in conjunction with other measurements, make it possible to reconstruct details of the cell cycle, that is, the stages within one round of replication that add up to the doubling time of the cell. From such measurements the schedule of cell proliferation and differentiation can be reconstructed. If the reconstruction is valid, the numbers of cells of various types present in the tissue must be inversely related to their rate of turnover. This is found to be the case.

With such methods we have established that epithelial cells mainly reproduce their own cell type. The nerve cells are reproduced from interstitial cells in about a day. The production of nematocytes takes much longer [*see illustration on preceding two pages*]. After several days of successive divisions forming nests there is a phase of eight hours during which the cells still look like interstitial cells. Thereafter inclusions appear, and a long phase of 40 hours follows. Eventually, at a well-defined time, the complex intracellular organs characteristic of nematocytes are formed. It appears as if during the 48 hours after the last cell division the requisite building material is synthesized, after which the intracellular structures are marshaled by a fairly fast self-assembly process. Some hours later the nests of nematocytes break up and most of the cells migrate individually to the tentacles, probably attracted by a chemotactic substance.

It can be shown that the number of interstitial cells accounts quantitatively for the production of nerve cells and nematocytes in the animal. Among them are the true stem cells, which are capable both of reproducing themselves and producing differentiated cells, mainly or exclusively nerve cells and nematocytes. Because of this dual task the stem cells have to replicate once a day, much faster than the doubling time of three and a half days for hydras themselves.

The balance sheet of cell proliferation and differentiation in the steady state indicates that several processes have to be quantitatively regulated. Regulation prevents stem cells from becoming either depleted or overproduced and assures that the various differentiated cell types arise in defined proportions. The ratio of stem cells that produce nerve cells to those that produce nematocytes is about 1 : 3 on the average but only 1 : 6 in the gastric column, where as in the head area it is mainly nerve cells that are produced. Presumably these regional differences are directed by chemical signals that vary quantitatively between the head and the gastric column. One can also imagine that chemical signals are involved in the movement of cells, for example the migration of nerve-cell precursors to the head area.

If we follow cell differentiation in a regenerating section rather than in a mature animal, a striking feature is the rapid production of new nerve cells in the developing head. They appear as soon as a day after cutting, even though it takes nearly a day to produce nerve cells from stem cells. The prepattern that determines the location of the new head in regenerating hydras is known to be established within a few hours after cutting; one of its earliest effects seems to be to initiate production of nerve cells for the future head.



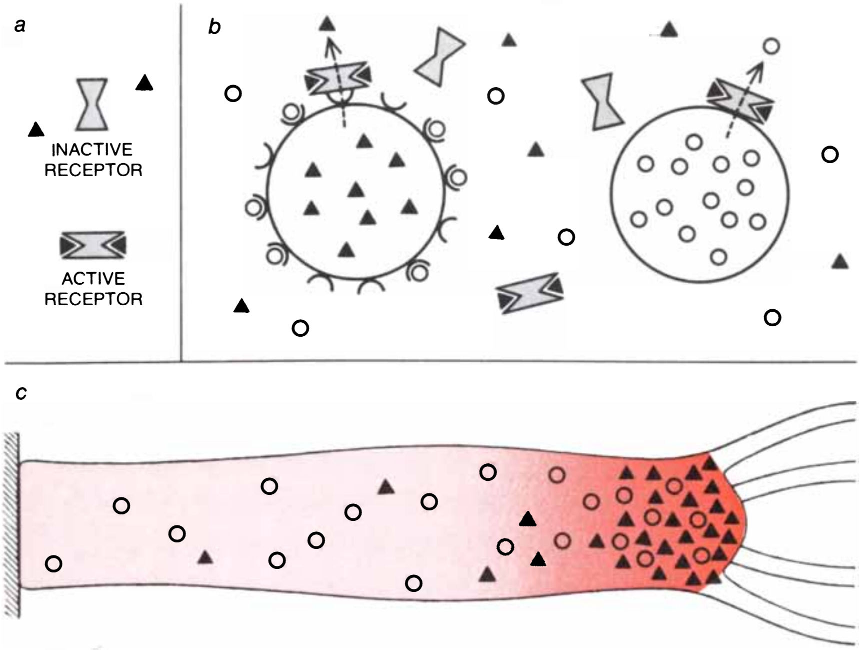
BASIS OF POLARITY, which directs location of parts in regenerating hydras, may be the graded distribution of cell types or some cellular component, indicated here by a color gradient. The gradient, but not its absolute concentration, determines which end of a hydra section will become activated (*black triangles*). The activator produced at that end is believed to act as a prepattern for the formation of head. This hypothesis explains how adjacent sections excised (*b, c*) will each develop a head at the end nearest the original one.

The evidence from observations of the aggregation and regeneration in hydras suggests that the invisible spatial prepatterns that precede the patterns of development are embodied in chemical substances. The differentiation and movement of cells are presumably caused by the formation of gradients of such substances in the animal. Over the past 30 years workers in a number of laboratories have been searching for these hypothetical "morphogens." No one, to my knowledge, has yet succeeded in chemically identifying a morphogen, if we do not count hormones and other long-range regulators. If one is skeptical, one may even doubt their existence. After all, any physical parameter that varies in spatial distribution, for instance some electrical property of cell membranes, would serve as well as the graded distribution of a chemical substance. Nevertheless, I am satisfied that the problem has a molecular answer, that morphogens exist and that the difficulty of identifying them can be attributed to the lack of an effective direct assay.

The morphogenetic effects of applied chemical agents need not be, and often are not, related to natural morphogenesis in the animal. If, however, a specific organic substance that occurs naturally in the animal were found to affect morphogenetic processes at low concentrations, it is probable that the same substance is also involved in natural morphogenesis. Nevertheless to prove that the substance is so involved is not easy.

Several morphogenetic effects have been reported for hydra extracts. In our own laboratory Hildegard Schaller and Stefan Berking have isolated from hydra tissue activating and inhibiting substances that influence budding and regeneration of the animal. They are organic substances whose molecular weights are lower than those of enzymes but are in the range of those of most hormones. Their distribution is graded within the animal, with the highest concentration being found in the head region. Moreover, the substances are active in concentrations as low as 10^{-10} gram per milliliter. This implies that little energy is required to maintain a graded distribution in spite of diffusion and molecular decay.

Perhaps the most interesting features of these substances are that most of them are bound to particulate structures within the cells and that they are active in experimental preparations at a concentration much lower than the average concentrations normally found in the animal. This suggests that the substances are bound to subcellular structures in an inactive form and are released in small amounts to exert their effects. Assuming that the substances are involved in natural morphogenesis, it is not their synthesis but their release that gives rise to activating or inhibiting effects and the spatial pattern of such effects. We do not suggest that all regulatory substances involved in morphogenesis are of this type. Some regulatory effects may be the result of genes being turned on or off; others may be due to the enzymatic synthesis of morphogenetic substances. Nevertheless, the release of preexisting morphogens from subcellular structures is an attractive possibility, since the release can be a fast process requiring no synthetic activity.



EXAMPLE OF MOLECULAR MODEL that leads to pattern formation by lateral inhibition, as proposed by the author, postulates the existence of activator molecules (*triangles*), inhibitor molecules (*circles*) and a receptor protein that is activated by association with two activator molecules (*a*). Activator and inhibitor molecules are stored in separate sacs but can be released by the special protein in its activated form (*b*). The inhibitor can prevent the release of activator by occupying release sites on the activator storage sac but cannot block the release of inhibitor. The inhibitor molecules diffuse more freely and thus have a wider range between release and degradation than the activator molecules do. The resulting distribution of activator in regenerating hydras (*c*) directs head formation.

If the substances we have identified, or other substances still unknown, can be shown to be true morphogens, they should eventually help us to understand biological development. The gap between empirical facts and their theoretical interpretation, however, is still very wide. At this stage it is useful to ask what kind of physical, molecular phenomena might account for the morphogenetic process.

It is well known that spatial order can be generated by the direct interaction of atoms and molecules. For example, atoms and molecules can form crystals. By an apparently analogous process large organic molecules are able to form membranes, fibers, tubes, cell components such as ribosomes and even entire virus particles. Cells interacting by direct contact can also give rise to specific structures. The double layer of ectodermal and endodermal cells found in hydras is constructed from disoriented aggregates of hydra cells by a sorting process that leads to an energetically favorable configuration. In the same way asymmetric tissue structures can be formed by the interaction of asymmetric cells.

The experimental evidence indicates that in addition there are many cellular patterns that do not result from direct cell contact. It appears that instead the ultimate pattern is preceded by gradients of morphogens that generate prepatterns extending over a distance of many cell diameters. The prepatterns then induce the formation of the visible patterns.

A "countdown" mechanism for forming a graded distribution of morphogens has been proposed by Lewis Wolpert of the Middlesex Hospital in London to account for outward-growing structures such as the buds of developing limbs. Such buds have a small growth zone near the tip where cells proliferate relatively fast. If at any given stage that zone produces a substance in an amount slightly less than the amount that was present in the immediately preceding stage, a gradient is created directly in the course of outward growth. That gradient may then give rise to the development of specific regions.

There is another widely occurring type of pattern formation for which a regenerating section of hydra is a typical example. Here again the formation of the visible pattern is explained by an invisible prepattern, but the prepattern is formed *de novo* in a tissue that is initially almost homogeneous. Such apparent creation of highly structured inhomogeneity out of near-homogeneity is an impressive phenomenon of development that was once frequently cited by vitalists: those who maintained that living organisms cannot be explained solely in terms of physics and chemistry. The formation of this type of prepattern is strongly influenced by communication between parts of the tissue extending over many cell diameters. For example, in a regenerating hydra section cut from the gastric column, the same piece of tissue can develop a head or a foot, depending on whether the adjacent tissue is derived from an area closer to the head of the original animal or more distant from it. Further, the size of a part, say the head area in a hydra, can be regulated in proportion to the size of the whole. In the same way a piece of certain insect eggs can give rise to a complete miniature insect in which all parts are reduced in scale.

There is also evidence for inhibitory effects extending over many cell diameters. Areas activated to give rise to a structure can inhibit the formation of further structures in that vicinity. Thus pieces of gastric column transplanted from one hydra to another can induce a second head only at a certain minimum distance from the first head. If several of the buds that will give rise to heads develop, they are spaced out on opposite sides of the body column. Inhibition can also account for the regular spacing of other biological structures, such as the leaves of plants.

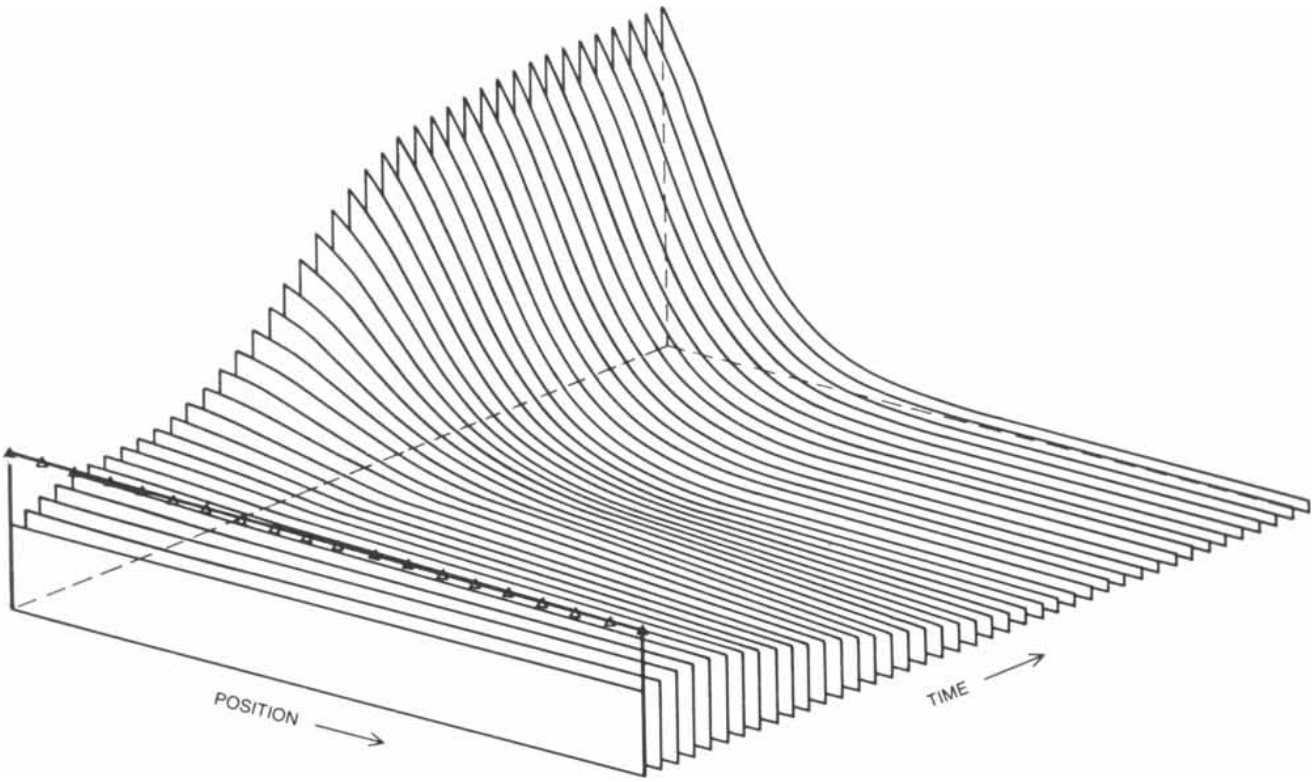
It is not difficult to visualize that a prepattern, once it is formed, is able to direct cell differentiation and morphogenesis. Prepatterns are not, however, the explanation of pattern formation but a central feature to be explained. How are they reproducibly formed, starting from nearly uniform initial conditions? Obviously if one is to explain a pattern by a prepattern, it is not satisfactory to explain the prepattern by some hidden pre-prepattern.

Although the orientation of a pattern can be determined by some asymmetric property, such as the polarity, of the tissue in which it is formed, the shape of the pattern itself must be generated anew in each generation. To explain a prepattern by a pre-prepattern would be to repeat the mistake made by the early biologists who thought that the egg harbors a miniature version of the adult organism. The miniature version would then have to contain microversions for further generations. In order to provide for only a few generations the tiniest version of a complete animal would have to occupy a volume smaller than an atom.

To obtain a physical understanding of pattern formation it is necessary to explain the spatial distribution of morphogens strictly on the basis of the physical interactions and movements of molecules. (It is understood, of course, that the synthesis of morphogens and the response of cells to morphogens in the course of cell differentiation and morphogenesis follow instructions incorporated in the organism's DNA.) With the aid of a computer one should be able to simulate processes of pattern formation by feeding into the machine nothing more than equations and parameters based on molecular physics.

How might one conceive spatial patterns to be formed? To begin with, we see that structure is the rule rather than the exception not only in the living world but also in the nonliving one. The laws of physics are such that a uniform distribution of matter tends to be unstable. Space is filled not with a featureless gas but with planets, stars and galaxies. It is true, of course, that in the test tube most chemical reactions lead to a uniform distribution of soluble substances. About 20 years ago, however, A. M. Turing showed that even in a liquid it is possible for patterns to arise if the liquid contains two substances that act on each other and on their own production, that is, by cross-catalysis and autocatalysis. Somewhat related mechanisms generate in a liquid the beautiful spiral waves studied by Arthur T. Winfree of Purdue University [see "Rotating Chemical Reactions," by Arthur T. Winfree; SCIENTIFIC AMERICAN, June].

The properties of patterns generated by catalytic effects do not generally resemble those observed in morphogenesis. Nevertheless one may inquire whether or not certain types of catalytic reaction known in molecular biology might lead to the kind of biological patterns I have described for hydra. What are the basic features that a theory should account for? For example, one observes that a particular tissue that is initially



COMPUTER SIMULATION of the author's model, described in the illustration on the preceding page, shows how the system can produce a steep, stable gradient of activator molecules with the passage of time. At the outset all the particles postulated by the model (free activator, free inhibitor, receptor proteins and storage sacs) are almost uniformly distributed (symbolized by triangles).

almost homogeneous develops reproducibly a pattern with strikingly different parts; the activation, say, of head formation extends into part of the total area but is also confined to that part. Furthermore, an activated area inhibits the formation of additional activated areas in the vicinity.

To account for these properties my colleague Hans Meinhardt and I have proposed that a requirement is an activator and an inhibitor capable of diffusing through tissues. The activator catalyzes or at least stimulates its own production and the production of the inhibitor. The inhibitor interferes with activation. The inhibitor diffuses more rapidly and thus has a longer range than the activator. We found by mathematical analysis that the same inhibitory effect that is needed to explain that an activated area inhibits the formation of other activated areas nearby is also able to explain (although much less obviously) the formation of striking patterns starting from nearly homogeneous initial conditions, if the catalytic effects of activators and inhibitors are of certain general types. The patterns thus formed have further general features that are found in, among other organisms, hydra. The size of the activated area can be regulated in proportion to the total size. Moreover, those slightly asymmetric distributions of cells or substances that are the cause of tissue polarity affect the orientation but not the general shape of the pattern. Because of the crucial role of long-range inhibition we call this the lateral-inhibition type of pattern formation.

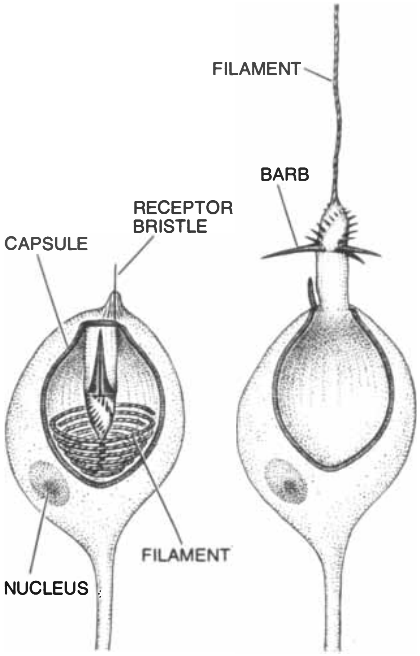
The general principle of pattern formation with lateral inhibition is consistent with many different molecular mechanisms and does not yield the chemistry of the reactions. It is remarkable, however, that relatively simple hypothetical models with features well known in molecular biology can give rise to such patterns, as will be shown by the following example. Activators and inhibitors are assumed to be contained separately in small storage sacs [*see illustration on page 51*]. They are released through the membrane of the sac through the action of a special receptor protein that is itself activated when it combines with two molecules of the activator. (Such "allosteric" shifts of proteins from an inactive configuration to an active one are well known in molecular biology.) When the activated receptor protein occupies a particular site on the membrane of an activator or inhibitor storage sac, activator molecules or inhibitor molecules are released through the membrane. (This mechanism is also known.) Once released, the activator and inhibitor are in the active state and are free to diffuse through the surrounding tissue. The inhibitor molecules become attached to the active sites on the membranes of sacs containing activator and there inhibit the release of activator molecules. (The inhibitor does not, however, inhibit its own release.) The short range of the free activator in the tissues and the longer range of the inhibitor are determined by their inherent speed of diffusion and by the rate of their enzymatic destruction.

If one assigns plausible values to these various parameters, one can put them into a computer and see what kinds of patterns the model can produce so as to demonstrate that patterns of the desired properties are indeed generated. The reasons for pattern formation can be described in qualitative nonmathematical terms as follows. In the model the properties of the receptor protein are chosen so that activator molecules, once they are released, can enhance further release. Thus a local head start of activator concentration above the average value can trigger further activation. The simultaneous increase in the release of inhibitor provides a supply of inhibitor molecules that diffuse rapidly and limit total activator production in a wider area. An increase of activator in one region can thus occur only at the expense of a decrease in other regions, provided that both are within range of the inhibitor. The process confines activation to a part of the total area. Furthermore, the formation of secondary centers of activation is inhibited in the vicinity of existing or developing centers of activation. Eventually pattern formation comes to an end and the pattern becomes stable, either because the diffusion of the activator prevents further confinement or because in the activated area nearly all the receptor-protein molecules have occupied membrane sites so that the activator concentration cannot increase further.

It should be emphasized that the pattern-forming properties of the model do not depend on its special chemical features but rather on its more abstract catalytic features. Many other models are capable of producing similar patterns as long as they have the right features. For example, the inhibitory effect could be caused not by an inhibitor but indirectly by the depletion of a substance required for activation. The production of activators and inhibitors could be regulated by synthesis instead of by release from storage sacs. Various regulating steps could also be guided by the actions of genes in the cell nucleus rather than by molecular events in the rest of the cell. Only biochemical investigation of the actual process can decide among the many conceivable mechanisms.

We have also considered properties of biological systems other than hydras, and we find that our activator-inhibitor model with lateral inhibition can account for symmetric, asymmetric and periodic patterns, for patterns within individual cells as well as for patterns within tissues, for oscillating patterns as well as stable ones and for patterns in two and three dimensions. In two dimensions if the total area exceeds the range of both activator and inhibitor, we obtain nearly periodic patterns of a "bristle" type: a regular spacing of similar structures. On a cylindrical surface we can obtain a regular spacing of activated areas analogous to the spacing of leaves in plants and buds in hydras.

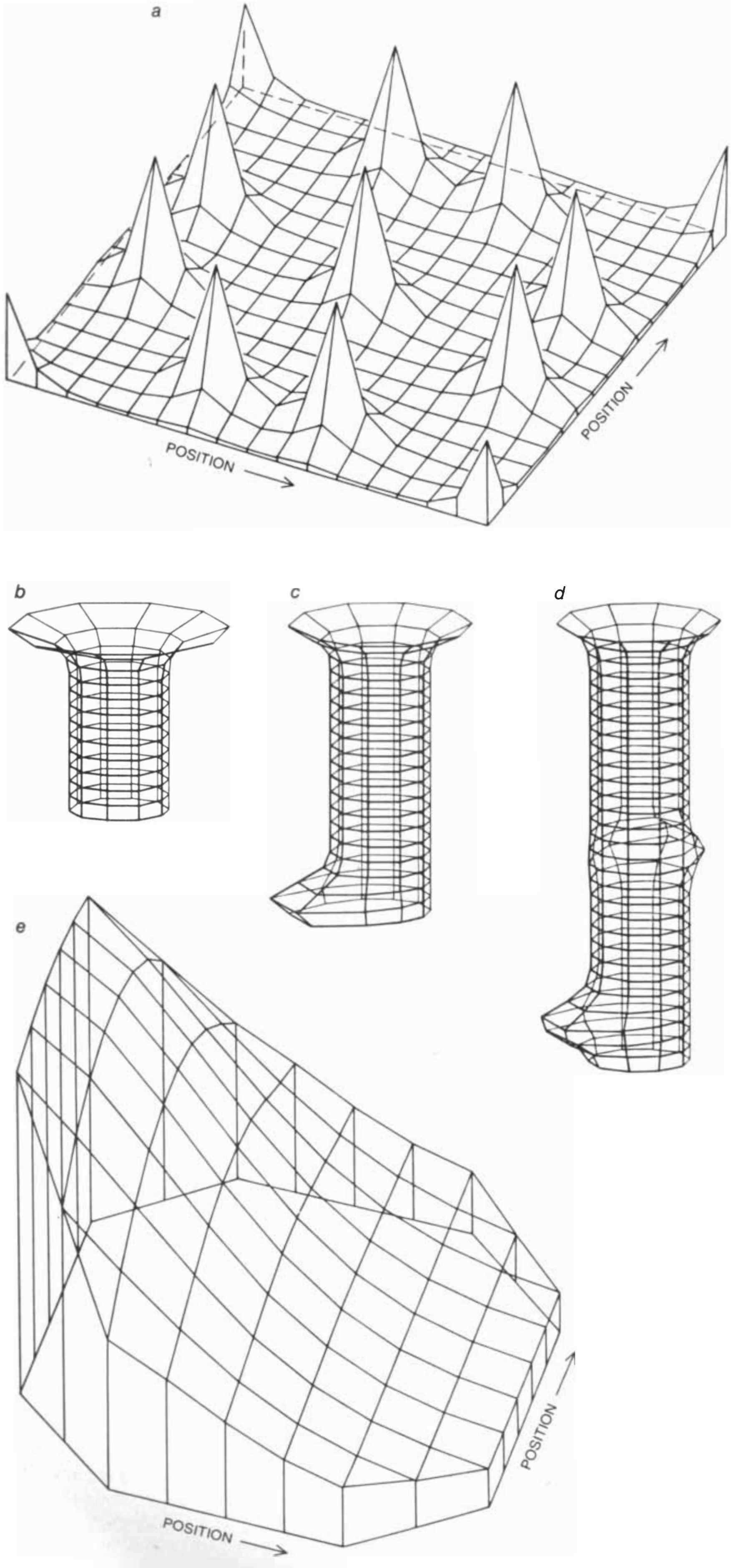
If the total area is comparable to the range of the activator, a sloping gradient can be formed. If cells respond differently to different concentrations within the gradient, one can obtain an aperiodic



NEMATOCYTES have a highly complex structure and function. In hydras there are several types. Before these stinging cells are discharged they contain a long coiled filament. When the receptor bristle is touched, an impulse is relayed to a capsule, which expels the filament (*right*). The filament releases a stinging substance, and the barb at its base helps to impale the prey, which is eventually carried to the animal's mouth.

structure with a specific spatial order of different parts in one of the dimensions, either dividing the tissue into different sections (comparable to the head, the thorax and the abdomen in a developing insect) or producing a graded distribution of subcellular structures or cell types (such as nerve cells in hydras). A second pattern-forming system of the same type could induce pattern formation in a second dimension. After growth, subpatterns may form in parts of the original pattern, giving rise to a still more complex spatial order. Pattern formation of this type may occur not only in tissues but also within single cells. For the sake of economy the genetic material may provide the instructions for only a limited set of pattern-forming systems in each organism and may arrange for each system to be utilized more than once at different stages and locations in the course of development.

I should like to emphasize that mathematical results cannot substitute for molecular biology; only biochemistry can reveal whether and in which cases pattern formation is actually of the type I have described, and what kinds of molecules are involved in it. On the other hand, these theoretical considerations suggest certain kinds of molecular mechanisms that can be subjected to experimental test. They explain how biological patterns can be formed by physical processes, giving rise to features that would not be obvious by merely looking at the chemical properties of the molecules involved. Thus one hopes that the physical understanding of pattern formation in living organisms will advance by the marriage of results on the mathematics of interacting systems with laboratory findings in molecular and cellular biology.



COMPUTER DRAWINGS show how the author's activator-inhibitor model can generate a variety of patterns. On a nearly uniform sheet the model can produce activated regions with a nearly equal spacing that results, for example, in the formation of bristles (*a*). Starting from a cylindrical surface the model is capable of producing a pattern (*b, c, d*) resembling the budding patterns of hydras. Hydra buds appear alternately on opposite sides of the cylinder. A sloping gradient results (*e*) if the range of the activator is comparable to the total size of the area. Such a gradient can elicit any aperiodic structure, continuous or discontinuous, for example the sequential arrangement of head, thorax and abdomen in developing insects. A second slope could also be generated to create patterns in a second dimension. Such plots, based on lateral-inhibition concept, give rise to stable configurations.

A little list of topics

Angiography

Audiovisual equipment and supplies
- Slide projectors
  - Auxiliary equipment
- Motion-picture projectors
  - Auxiliary equipment
- Television projectors

Autoradiography

Chemicals
- Biochemicals
- Protein/polypeptide reagents
- Electrophoresis reagents
- Pharmaceutical intermediates in bulk
- Thin-layer chromatography supplies
- Dyes and stains
- Liquid scintillation reagents
- Clinical chemicals

Cinefluorography

Cinephotomicrography and time-lapse photography

Clinical photography

Close-up photography

Copying

Dental radiography

Electron micrography

Films for special medical-photographic techniques

Fundus photography

General photography
- Cameras
- Color films
- Black-and-white films
- Color papers
- Black-and-white papers

Graphic arts photography
- Films and papers
- Printing plates
- Processing equipment

Gross-specimen photography

Holography and optical-data processing

Infrared photography

Instruction and education
- Audiovisual products
- Color microfiche and microfiche readers
- Micropublishing

Instrumentation photography
- Light-beam oscillography
- Cathode-ray tube photography
- High-speed photography
- Black-and-white films
- Color films
- Photographic papers

Mammography

Medical radiography
- Films
- Processing equipment
- Minification equipment
- Processing chemicals
- Accessories

Micropublishing

Motion-picture photography
- Cameras (super 8)
- Films (super 8)
- Films (16 and 35 mm)

Nuclear medicine

Photofluorography

Photomacrography

Photomicrography

Radiation therapy

Radiograph copying

Radiograph subtraction and duplication

Records and administration
- Microfilming
- Computer-output microfilm
- X-ray minification

Thermography

Ultrasonic imaging

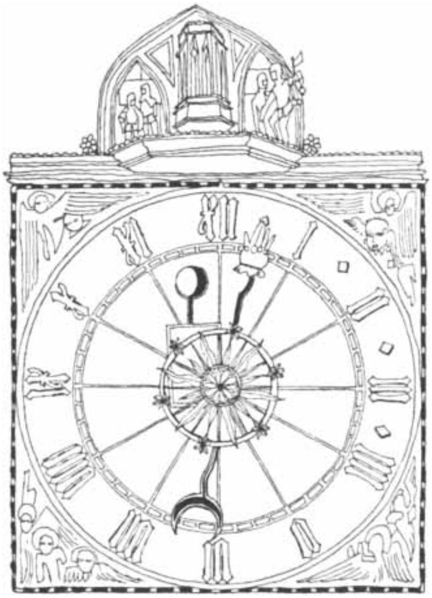
Ultraviolet and fluorescence photography

The health-care professions and biological science find important work for Kodak products and services. Knowledge, though, tends to be fragmentary about what we offer and where to get it, and where we bow out. If you have only hearsay to go on for Kodak products, services, and literature useful in this field, please write to me. A biomedically oriented directory is in preparation. Let me know if you'd like a copy.

RICHARD FULLER

EASTMAN KODAK COMPANY
Kodak Office
Rochester, New York 14650
716 724-4624

SCIENCE AND THE CITIZEN



The Nobel Prizes

The 1974 Nobel prizes in science were awarded to three biologists who discovered new functional structures within the living cell with the aid of the electron microscope and the high-speed centrifuge, to two physicists who developed and fruitfully applied sensitive new techniques for studying radio-emitting astronomical objects, to a chemist who placed the study of synthetic polymers on a firm quantitative basis and to two economists who have helped to clarify complex interrelations among economic, social and institutional phenomena.

The senior member of the group that shared the prize for physiology or medicine, Albert Claude, who is still active at 75, is recognized among his colleagues as one of the principal founders of modern cell biology. A native of Luxembourg, Claude received a medical degree from the University of Liège in 1929 and soon afterward went to the Rockefeller Institute for Medical Research (now Rockefeller University). For the past several years he has been director of the Institut Jules Bordet in Brussels.

In 1937 Claude began using the high-speed centrifuge in an effort to isolate the Rous sarcoma virus, a then mysterious and invisible entity that gave rise to tumors in chickens. (The virus had been discovered at the Rockefeller Institute in 1910 by Peyton Rous, who received a Nobel prize for his discovery 56 years later.) Claude soon isolated and purified the virus and went on to use the centrifuge as a general tool for isolating many active components of mammalian cells. This work led to the isolation of the mitochondria: organelles, or intracellular particles, barely visible in the light microscope that are the principal locus of energy transformation in the cell. Claude also isolated particles that were not visible in the light microscope; he named them microsomes.

Eager to see mitochondria in greater detail and to obtain the first images of virus particles and microsomes, Claude and his collaborator Ernest F. Fullam began in the early 1940's to work with an RCA electron microscope belonging to a nearby industrial laboratory. Although the electron microscope offered a resolution between 50 and 100 times better than the light microscope, it called for entirely new methods of sample preparation. Between 1945 and 1947, with important help from a new collaborator, Keith R. Porter, Claude and Fullam succeeded in obtaining the first electron micrographs of intact chicken-tumor cells in which could be seen, in addition to mitochondria, virus particles, the intracellular structure known as the Golgi apparatus and a "lacelike reticulum" extending throughout the cell's cytoplasm.

At about that time Claude and Porter were joined by George E. Palade, the second member of the group that shared the prize for physiology or medicine. Born in Romania, Palade received a degree in medicine from the University of Bucharest in 1940. He made important advances in the preparation of specimens for electron microscopy and in selecting media for improving the fractionation of cell components in the centrifuge. With Porter, Palade defined the lacelike reticulum revealed by the electron microscope as the endoplasmic reticulum and concluded (with Philip Siekevitz) that the microsomes in Claude's cell fractions were mainly fragments of that structure and not organelles per se.

Palade and his co-workers went on to discover the subcellular particle known as the ribosome, an organelle supported by the endoplasmic reticulum. They showed that the ribosome functions as the cell's protein factory, acting on genetic instructions carried from the DNA in the nucleus by "messenger" RNA. They also helped to clarify the role of the Golgi apparatus, showing that it participates in packaging protein for export from the cell. Since 1973 Palade has been at the Yale University School of Medicine.

Christian de Duve, the third member of the group that shared the prize for physiology or medicine, is a Belgian who maintains laboratories both at Rockefeller University and at the University of Louvain. After obtaining degrees in medicine and chemistry at Louvain in 1941 and 1946 he worked at the Nobel Institute in Stockholm and then at Carl Cori's laboratory at Washington University. Like Claude and Palade, de Duve has specialized in the use of centrifugal fractionation techniques for the isolation and study of subcellular particles. He is best known for his discovery of lysosomes, organelles that he has described as "tiny bags filled with a droplet of a powerful digestive juice capable of breaking down most of the constituents of living matter. [The] lysosomes function in many ways as the digestive system of the cell" [see "The Lysosome," by Christian de Duve; SCIENTIFIC AMERICAN, May, 1963].

The winners of the prize for physics, Martin Ryle and Antony Hewish, have long worked together at the University of Cambridge, where they have vastly extended the sensitivity and resolution of radio telescopes. The resolving power of a telescope depends on its having an aperture, or diameter, that is large in relation to the wavelength of radiation being collected. With light telescopes ratios between aperture and wavelength of several million to one are readily attainable because the wavelength of light is only about half a millionth of a meter. It is clearly impractical to attain the same ratio of aperture to wavelength in the radio region, where wavelengths range from roughly a centimeter upward. A radio-telescope "dish," or parabolic reflector, with an aperture of a million centimeters would be 10 kilometers, or 6.2 miles, in diameter.

Ryle was among the first to demonstrate that a radio telescope of very large aperture can be simulated by combining the radiation collected by several modest-sized dishes mutually adjustable within a distance of several kilometers. Known as the aperture-synthesis technique, this approach yields a resolving power equivalent to that of a single large

DP SCIENCE DIALOG

Notes and observations from IBM which may prove of interest to the scientific community.

The Synagogue of Khirbet Shema Restored

About 90 miles north of Jerusalem lies the village of Khirbet Shema, which until recently was buried under layers of soil. The village was destroyed sometime around 417 A.D., probably by an earthquake which struck the upper Galilee about that time. Among the ruins is an ancient synagogue, which is now being restored with the help of an IBM computer. Co-heading the team of archeologists is Dr. James F. Strange, who teaches biblical archeology at the University of South Florida.

According to Dr. Strange, the ruins of the synagogue show it to be unique, both for the area and time of construction, estimated to have been around the fourth century A.D. "There were several factors which led us to believe the building was quite unlike any other synagogue in that part of Palestine," he says. "The foundation walls were nearly two feet thick and an entrance staircase had stairs ten feet wide. There also were a pair of massive door posts cut from solid rock, eight roof-support columns almost two feet thick and a number of underground chambers."

Preliminary research on selecting a site began several years ago but the actual excavation at Khirbet Shema was started in 1970. The team's main interest was to excavate the synagogue in its context, uncovering adjacent buildings and surrounding parts of the village. By the end of the first summer's work they had recovered the plan of most of the village, an area of more than six acres. Restoration began in 1971 and was completed during the summer of 1973.

The team, which has included up to 80 people along with three architects, first drew tentative plans of the original building and prepared detailed calculations, measurements and instructions for actual reconstruction of pedestals, columns, capitals and doorposts. The final piece of the puzzle lies in finding the type of roof which covered the structure.

To date no remnants of the roof have been found at the site, nor have any other positive clues appeared. So

(continued on third page)

(Continued from first page)

dish filling the entire area between the individual small dishes.

With a succession of improved instruments Ryle and his colleagues have produced major surveys (known as the Cambridge catalogues) of celestial radio sources, most of them too distant and too faint to be detected with light telescopes [see "Radio Galaxies," by Martin Ryle; SCIENTIFIC AMERICAN, September, 1956]. Subsequently some of the most intense radio sources were optically identified as quasars, remote starlike objects that emit many times more energy than entire galaxies.

Ryle's colleague Hewish was awarded half of the prize for physics for his role in the discovery in 1967 of pulsars, radio-emitting objects that seemingly flicker on and off with extreme precision at intervals ranging from a thirtieth of a second to a few seconds [see "Pulsars," by Antony Hewish; SCIENTIFIC AMERICAN, October, 1968]. It is believed that pulsars are neutron stars, objects consisting largely of neutrons compressed to a density of millions of tons per cubic centimeter. Scarcely 10 kilometers in diameter, pulsars evidently are rotating at the frequency indicated by their clocklike signals, which are believed by many physicists to arise from electrons (or positrons) streaming outward in magnetic fields millions of times stronger than any produced in laboratories on the earth. Of some 120 pulsars discovered at radio frequencies, only one has been identified with a blinking visible object. The discovery of pulsars has stimulated the investigation of the properties of matter under extreme conditions, including the study of black holes, where matter is so dense and gravitational fields so intense that nothing, including light, can escape their grip.

Paul J. Flory, winner of the prize for chemistry, is little known outside the fraternity of polymer chemists, but among them he is recognized for bringing order into a field that seemed ready to collapse into a chaotic jumble of unrelated measurements. After receiving his Ph.D. at Ohio State University in 1934, Flory worked for several years at E. I. du Pont de Nemours & Company with Wallace Hume Carothers, the inventor of nylon. After a succession of industrial and academic positions he accepted his present appointment at Stanford University in 1961.

Trying to compare different polymers with each other, say nylon and one of the scores of synthetic rubbers, may seem like trying to compare apples and oranges. They are constructed from different monomers; the monomers may be linked together by a totally different chemistry; the resulting polymers have different lengths, viscosities, solubilities and so on. Flory was able to show that if the polymer is dissolved in a suitable solvent there exists a temperature—the theta (or Flory) temperature—at which the polymer is in a kind of ideal state. At that temperature, which varies for different polymers and solvents, the measured properties of the polymer can be usefully compared with those of other polymers at their theta temperature. Flory has also shown how to compute a "universal constant," from measurements of viscosity, light dispersion, sedimentation rate and the like, that summarizes all the properties of a polymer solution.

The prize in economics went to two men "whose research has reached beyond pure economic science and [whose names] have always been on the list of proposed prize winners" since the Central Bank of Sweden established the award in 1968. Friedrich von Hayek and Gunnar Myrdal found themselves for the first time in their lives associated "for their pioneering work in the theory of money and economic fluctuations and for their penetrating analysis of the interdependence of economic, social and institutional phenomena." From their youthful attempts to comprehend the macroeconomic manifestations of the Great Depression of the market economies in the 1930's, the two prize winners had proceeded on divergent paths into analysis and interpretation of "the factors and linkages which economists usually take for granted or neglect." Von Hayek has stood against all extra-economic impulses and policies that would perturb the system's classically imputed tendency to resolve its inequities and conflicts in the harmony of equilibrium; in recent years he has looked to the perfection of the flow of information in the system to reduce the frictions that exact the human costs (for example unemployment) of economic fluctuation. Myrdal, a lifelong and deeply involved observer of social orders as disparate as those of the U.S. and of India, has posited the vicious circle ("cumulative circular causation") as the natural mode of the economic process. Against the universally observed tendency of the rich to grow richer and the poor to grow poorer he has helped to call the welfare state into action in Europe and America and in the motivation of the economic development of the underdeveloped countries [see "The Transfer of Technology to Underdeveloped Countries," by Gunnar Myrdal, SCIENTIFIC AMERICAN, September, and 'Political Factors in Economic Assistance," by Gunnar Myrdal, SCIENTIFIC AMERICAN, April, 1972].

Imbalance of Terror

The prospects are slim for any significant new agreement between the U.S. and the U.S.S.R. on the limitation of offensive strategic weapons, according to a report issued by the Stockholm International Peace Research Institute (SIPRI). The reason is that such agreements seem to be attainable only when the two sides have achieved some form of parity in destructive capability. At this point the U.S. is far ahead of the U.S.S.R. in a crucial area: the ability to destroy the other side's missile force. Missile-improvement programs now contemplated in the U.S. would greatly enlarge that lead, according to SIPRI. The resulting aggravation of the imbalance of terror would not only reduce the possibility of agreement but also increase the likelihood of nuclear war.

The current U.S. advantage is obscured, the report points out, by the terms in which the U.S. political debate on armaments is conducted: aggregate numbers and megatonnage of missiles. Those properties of nuclear weapons may be pertinent to a strategy of deterrence, in which each side threatens to destroy the other's population and industrial capacity. A year ago, however, Secretary of Defense James R. Schlesinger announced that this country intended to develop a counterforce capability, or the ability to destroy Russian land-based missiles in their silos. The destruction of a missile in a hardened silo requires a combination of high yield and accuracy, and of the two, accuracy is by far the more important. "For example, a weapon with a yield 10 times greater than a Minuteman III warhead is, given the same accuracy, five times more effective in destroying a silo, but a weapon that has the same yield as a Minuteman III warhead, but which is 10 times more accurate, is 100 times more lethal to a silo."

The report undertakes a detailed mathematical analysis of the two sides' nuclear arsenals based on a factor, designated K, that is a function of yield and accuracy and a measure of anti-silo lethality. It concludes that the total K of the U.S. offensive arsenal has increased sharply as a result of the deployment of multiple warheads and improvement in accuracy, even though the number of missile launchers has remained steady since 1967 at 1,710. The total Russian K has increased more slowly as the U.S.S.R. has emplaced more missiles. At this point the U.S. has a K of about 20,000: a five-

SCIENCE/SCOPE

How a communications satellite can serve isolated areas was demonstrated to delegates at the first International Congress of Electrical and Electronic Communications in Mexico recently. A clear voice link was immediately established with people on Isla de Cedros, an island off the west coast of Baja California, via Western Union's Westar I and a portable earth station, both built by Hughes. In another experiment, three Canadian oil companies successfully completed similar demonstrations, using the same earth station with the Hughes-built Anik satellite to relay voice communications from remote arctic sites to company headquarters in Calgary.

A hologram lens system for a pilot's "head-up" display, currently being developed by Hughes research scientists, uses holography to produce the optical properties of a lens on a transparent plate in the pilot's line of sight. Projecting information via this plate does not block his vision outside the cockpit. The new technique has a larger field of view and lighter weight than display systems using conventional lenses and promises to be highly cost-effective.

The first mosaic map of the continental U.S. ever assembled from satellite photos taken from the same altitude and lighting angle was completed recently by the U.S. Department of Agriculture, which combined 595 photos taken by the Hughes-built multispectral scanner system (MSS) aboard NASA's Earth Resources Technology Satellite (ERTS). The map will aid in assessing the nation's surface water, drainage network, land use, and vegetation. Though NASA originally predicted a lifetime of only one year for ERTS, it began its third year July 23. Its still functioning MSS has now sent more than 200,000 photos back to earth.

A space-qualified three-stage cryogenic refrigerator, designed for super-cooling infrared sensors to increase their sensitivity, has been developed for the U.S. Air Force by Hughes. It is the first three-stage refrigerator of the Vuilleumier type ever built and cools down to -439°F. (absolute zero is -459.6°F.). Because it operates at slow speed and low pressure, it provides high reliability and a long, maintenance-free life. It has already operated for 2,000 hours and will be delivered to the Air Force Flight Dynamics Laboratory at Wright-Patterson Air Force Base for a 5,000-hour operating-life test.

Hughes Space and Communications Group, El Segundo, Calif., is seeking engineering managers to direct the design and development of RF and digital electronics circuits and subsystems and solid-state microwave devices, including parametric, tunnel diode, low-noise transistor, linear and class C power amplifiers, and microwave integrated circuits. Requirements: engineering degree, minimum of 10 years technical experience, U.S. citizenship. Write: Mr. A. St. Jacques, P.O. Box 92919, Los Angeles, CA 90009. An equal opportunity M/F employer.

Two weather-mapping instruments will be built by Santa Barbara Research Center, a Hughes subsidiary, under contract with NASA's Goddard Space Flight Center. The VISSRs (Visible/Infrared Spin-Scan Radiometer) will be the major payload aboard the second and third Geostationary Operational Environmental Satellites, scheduled for launch in 1976 and 1977. A VISSR aboard the first Synchronous Meteorological Satellite is providing excellent day-night weather photos of the entire earth's disc every 30 minutes from a stationary position 22,300 miles above the Amazon.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY

Technology: Population, pollution

Almost since the beginning, men have been predicting the end. Population exploding. Disease unchecked. Natural resources used up. Now technology is being added to the list of disaster-causes. Yet by so doing, prophets of doom are laying blame today on the hope for tomorrow.

The problems are always easier to recognize than the solutions.
The classic example was in 1798. With the world's population having risen slowly to one billion, English economist Thomas Robert Malthus wrote, *An Essay on the Principle of Population.* He declared that the world would be unable to support many more people; that only war, famine and disease would hold population in check.

Since then the world's population has increased four-fold. Malthus failed to foresee the enormous impact of medical science on saving lives and the ability of technology to support them. His gloomy predictions never came true.

Doom is still being projected today.
A team from a major university used a computer and a world model to investigate accelerating industrialization, exploding population, widespread malnutrition, depletion of resources and deteriorating environments. The team reported "the basic behavior mode of the world system is exponential growth of population and capital, followed by collapse." They see a collapse within a hundred years if present trends continue.

Sound familiar? Malthus by computer. But the same mistake he made recurs. Malthus did not take into account the growth of technology—nor can it be programmed into that computer model.

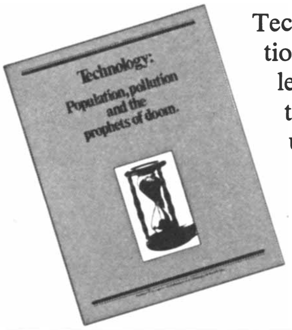
After two million years, is man no longer capable of ingenuity and adaptability?
Man's survival has not been dependent on availability of raw materials or favorable climate, but on his ingenuity. His ability to survive and change in a potentially hostile environment is what sets him apart from other animals.

Are we reproducing ourselves into oblivion? No. As countries develop industrially, birth rates go down and population stabilizes. That's an indirect result of technology. And man has developed population control methods more agreeable than Malthus' war, famine and disease.

What about the prediction of famine? In 1820 the United States farm worker could produce enough to sustain four people. Today he can feed over fifty. And new technologies relating to leaf protein and aquaculture hold a promise for eliminating hunger.

Will we use up the resources we were originally endowed with? The supply is indeed finite,

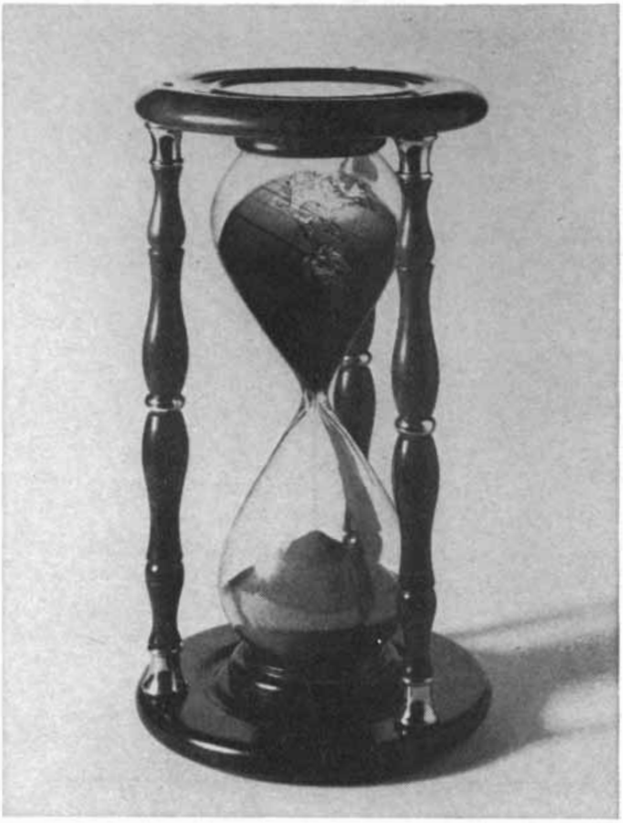
We believe in the promise of technology.



Technology has provided the solutions to many of society's problems. It is also often blamed for them. Some people would have us believe that there's too much technology today — that this very abundance is the basis for so many of our ills.

We believe the solutions to problems lie in the intelligent application of technology—that a misunderstanding of technology's role in man's advancement can only lead to a diminution of his ability to provide solutions.

Therefore, we're sponsoring a series of "white papers" to foster a clearer understanding of technology. Excerpts from the third of these papers are published above. If you'd like a copy of the complete text, please write Gould Inc., "Dialogue on Technology," Dept. S-3, 8550 West Bryn Mawr Avenue, Chicago, Illinois 60631.



fold advantage. The U.S.S.R. is now engaged in replacing its single-warhead missiles with multiple-warhead ones. Presumably it will also seek to improve the accuracy of its warheads, which is now far below that of U.S. weapons. Assuming no change at all in the U.S. arsenal and assuming the greatest degree of improvement that SIPRI can conceive for the U.S.S.R., this country would still have an advantage in total *K* in the early 1980's. Neither side, however, would have a large enough total *K* to destroy all of the other side's missile silos. Neither side, in other words, would have a dependable counterforce capability.

That would be the case if the U.S. stands still. Secretary Schlesinger, however, has proposed improvements in missile yield and accuracy that would increase the total *K* of the U.S. arsenal to more than 110,000 by about 1983—five times the predicted Russian figure and enough to threaten assured destruction. Another development program suggested by Schlesinger would introduce maneuverable reentry vehicles (MARV) for U.S. missiles, vastly improving their accuracy and producing a total *K* of some 4,000,000. "Yet such staggering lethality does not offer any practical superiority," because even if all Russian land-based missiles are destroyed, the U.S.S.R.'s submarine-launched missiles could still devastate U.S. population and industrial centers. In other words, deterrence remains the controlling factor, and the attainment of counterforce capability will have served only to make arms control less likely than ever.

Still Earlier Man

Three fragmentary fossil jaws unearthed on the Ethiopian plateau northeast of Addis Ababa have added at least another million years to the antiquity of the genus *Homo*. The discovery, made during a survey of fossil-rich formations in the Hadar River valley by a joint Ethiopian-French-American expedition, was announced by the leaders of the survey, Maurice Taieb, a geologist from the National Center of Scientific Research (CNRS) in France, and Karl Johanson, an anthropologist from Case Western Reserve University. Up to the time of their find the oldest known human fossil was thought to be a small but manlike skull, ER 1470, found by Richard Leakey in Kenya in 1972 below a 2.6-million-year-old layer of volcanic rock. A layer of basalt overlying the sandstones where the Ethiopian fossils were found dates back some three million years. Leakey estimates that the Kenyan skull is about three million years old. The discoverers of the Ethiopian jaws estimate that they are about four million years old.

The new fossils are one entire upper jaw, complete with teeth, one half of an upper jaw and one half of a lower jaw, each also complete with teeth. So far no additional cranial bones and none of the specimens' other bones have come to light. Both in size and shape the teeth of all three jaws appear to lie within the human range.

Theorists of human evolution, who have not yet fully assessed the impact of Leakey's 1972 discovery, now face an even knottier problem. If members of the human genus flourished as long as four million years ago, then the time when the genus first branched from its ancestral primate stem would necessarily be even earlier. As Taieb and Johanson assert, "All previous theories of the origin of the lineage which leads to modern man must now be totally revised."

Saltwater Space Race

The 22nd challenge for the America's Cup, the trophy held since 1857 by the New York Yacht Club, was contested in September in waters off Newport, R.I. The defending boat, *Courageous*, retained the trophy for the U.S. by defeating the Australian sloop *Southern Cross* in four races straight. In the aftermath of the competition *Scientific American*'s correspondent at Newport made some observations on the technology of the modern racing yacht:

"The '12-meter rule,' which delimits the hull dimensions and sail area of yachts competing for the America's Cup, has now been tested by a generation of marine architects. The results of the last few challenges suggest that within the limits the rule imposes on the basic lines of the hull maximum performance has been very nearly approached.

"In overall form the hull of the winning yacht this year closely resembles that of *Intrepid*, which successfully defended the cup in the previous two match races, in 1967 and 1970. The one boat with radically novel lines was *Mariner*, entered as a contender for the defense. The merits of *Mariner*'s hull were evaluated in an elaborate series of computer simulations and tank tests, but no breakthrough in design was forthcoming. As a result the trials held during the summer to choose a defender lapsed into a contest between *Courageous* and *Intrepid*, which has been revised twice since she was built seven years ago and remains very competitive.

"If the question of optimum hull shape under the 12-meter rule seems to have been settled, if only for the moment, refinements in materials and in gear both above and below decks have become more important. The most obvious expression of the new materials technology is that *Courageous*, *Mariner* and *Southern Cross* all have a hull of aluminum. (A precedent must be mentioned here; the defender of the cup in 1895 was also aluminum-hulled.) The hull plates of the modern aluminum boats are curiously thick; they would do credit to a tugboat. Such conservative construction was specified not by the Race Committee of the New York Yacht Club, but by Lloyd's of London. In spite of this concession to the prudence of the insurance underwriters, the aluminum hulls are quite light. More than three-quarters of the displacement of the new boats is the lead ballast of the keel, giving a very low center of gravity.

"Materials far more exotic than aluminum might have been employed in this year's yachts if they had not been prohibited by a 1971 ruling intended to hold down the cost of the competing yachts. Boats built after 1972 were not allowed to incorporate any materials stronger than steel, lighter than aluminum or heavier than lead. The rule eliminated from new America's Cup racers such developments as uranium ballast (more than half again as dense as lead) and structural members of boron fiber, carbon filament or extruded titanium and beryllium. By 1970 these and other concoctions of the aerospace industry had threatened to transform the sport of yachting into a saltwater space race.

"The new yachts demonstrated that even within the constraints of the 1971 ruling there is room for very sophisticated marine technology. The guys and stays on *Courageous*, for example, are fabricated of a nickel-cobalt-chromium-vanadium alloy with a tensile strength of 240,000 pounds per square inch. The backstay, traditionally a thick cable, is a cold-forged rod about 100 feet long and only a quarter of an inch in diameter. The alloy is virtually immune to corrosion by salt water. Although it is not as light or stiff as the titanium and beryllium rigging on some of the pre-1972 boats, it approaches the limit of what is allowable under the current rule.

"Not all the achievements of yachting technology have to do with materials. The winching systems, for example, are delightfully complex and precise; in their elegant mechanism torque and thrust are balanced throughout. They are cranked by hand (no powered apparatus is allowed) with negligible loss to friction. *Courageous* was also equipped with a

small computer to handle certain navigational calculations. From data supplied by the boat's regular complement of instruments, the computer derived such statistics as true windspeed and direction.

"The race is not always to the lowest Reynolds number, however, nor is new technology always handed down from NASA. Exploring the deck fittings of *Courageous,* I discovered that one of the subtle superiorities of the victorious sloop was a nonskid deck paint that gave a secure footing when wet but was not so abrasive that it might be injurious to crew or sailcloth. On inquiring into the nature of its solid phase, I was told: 'Crushed walnut shells.' "

Second Superfluid

Superfluidity, the ability of a liquid to flow without friction, has until recently been believed to be the unique property of helium 4 (the common isotope of helium) that has been cooled to about two degrees Kelvin, or two degrees Celsius above absolute zero. It had been predicted by some theorists that helium 3 (the rare isotope of helium) would also exhibit superfluid behavior, but the temperature at which superfluidity would be observed was estimated to be several orders of magnitude lower than the transition temperature for helium 4, and such temperatures were unattainable with the apparatus then available. Now, with the aid of new techniques for attaining ultralow temperatures, it has been demonstrated that helium 3 has not one superfluid phase but two.

Writing in *Europhysics News,* a bulletin of the European Physical Society, O.V. Lounasmaa of the Helsinki University of Technology says that a new and exciting era in the study of liquid helium began in 1972 when D. D. Osheroff, Robert C. Richardson and David M. Lee of Cornell University found the two new low-temperature phases of helium 3, which they named the *A* phase and the *B* phase. The *A* phase appears at temperatures between .002 and .0026 degree K. when the applied pressure is between 21.5 and 33 atmospheres. The *B* phase appears at lower temperatures. The point at which normal liquid helium 3 can coexist with the *A*-phase and *B*-phase liquids is .0022 degree K. and 21.5 atmospheres. The *A*-phase liquid disappears at lower temperatures or lower pressures. At zero pressure the transition from the normal liquid to the *B*-phase liquid occurs at .00093 degree.

The first indication of superfluidity in helium 3 came from experiments carried out by Lounasmaa and his colleagues last year. They measured the viscous damping of a vibrating wire in liquid helium 3 and found a sudden decrease in the damping, suggesting superfluidity associated with the *A*-phase and *B*-phase liquids.

Proof that helium 3 is a superfluid came early in this year, when H. Kojima, D. N. Paulson and J. C. Wheatley of the University of California at San Diego published their results that "fourth sound" will propagate in the liquid. Fourth sound is a pressure wave that will travel through a superfluid in a closely packed powder but not through a normal liquid because the viscosity of the normal liquid prevents it from moving or vibrating in the fine channels of the powder.

According to Lounasmaa, the superfluidity in helium 3 is "more closely related to superconductivity in metals than to superfluidity in helium 4." Helium 4 has an even number of electrons, protons and neutrons, and at low temperatures some helium-4 atoms fall into a zero momentum state with zero entropy. These atoms constitute the superfluid component in the liquid. An atom of helium 3 cannot fall to a zero momentum state because of its symmetry properties: it has only one neutron and therefore has a spin. Two helium-3 atoms form a pair, much as two electrons in a superconductor form a pair and drop into the superconducting state. Superfluidity occurs in liquid helium 3 when such pairs are formed. Unlike the pair of electrons in a superconductor, however, which have opposite spins that cancel each other, the spins of the pair of atoms in helium 3 are parallel, giving the pair one unit of spin.

Lounasmaa notes that although many of the properties of superfluid helium 3 have been studied in a little more than two years, many theoretical questions remain. "This rapid progress," he concludes, "shows both the vitality of current research at ultralow temperatures and the importance of these phenomena for the modern theory of collective systems."

High in the Sky

Is there intelligent life elsewhere in the universe? A piece of positive evidence that was facetiously adduced a year or two ago was that whereas methyl alcohol had been discovered in interstellar space, ethyl alcohol (the potable kind) had not. Clearly someone had consumed the stuff.

This positive evidence no longer obtains. In October a group of workers from the University of Maryland, the National Radio Astronomy Observatory (NRAO) and the Commonwealth Scientific and Industrial Research Organization used for the first time a new and highly sensitive spectrometer in conjunction with the 36-foot radio telescope of the NRAO at the Kitt Peak National Observatory. With this instrument they began investigating the dense cloud of gas and dust designated Sagittarius B2, a rich source of most of the known molecules in space, and discovered weak radio emission at the wavelength of 3.3 millimeters. A group from the National Bureau of Standards that followed them in using the same instrument independently found another emission line and tentatively identified it as radiation from ethyl alcohol. Two weeks later a group from the Center for Astrophysics of the Harvard College Observatory and the Smithsonian Astrophysical Observatory detected several more lines and confirmed the identification.

Ethyl alcohol, composed of nine atoms (C_2H_5OH), is one of the largest and most complex molecules now known to exist in interstellar space. Although it is neither unstable nor extremely rare in space, it had not previously been found because its emission is too weak to be detected by most existing radio spectrometers. The substance is spread in a thin vapor throughout Sagittarius B2, which is some 50 light-years in diameter. One calculation shows that, if the ethyl alcohol were condensed, it would come to 10^{28} fifths of a gallon at 80 proof. Potability, however, might be a problem; the alcohol is heavily contaminated with substances such as hydrogen cyanide, formaldehyde and ammonia.

The discovery of ethyl alcohol brings the total number of different kinds of molecule detected in space to 32. In order of their discovery those molecules are: methylidyne, the cyanogen radical, ionized methylidyne, the hydroxyl radical, ammonia, water, formaldehyde, carbon monoxide, hydrogen cyanide, molecular hydrogen, "X-ogen" (an unknown substance), cyanoacetylene, methyl alcohol, formic acid, carbon monosulfide, formamide, silicon monoxide, carbonyl sulfide, methyl cyanide, isocyanic acid, hydrogen isocyanide (only tentatively identified), methylacetylene, acetaldehyde, thioformaldehyde, hydrogen sulfide, methanimine, ethynyl, sulfur monoxide, dimethyl ether, silicon sulfide, N_2H^+ (unnamed) and ethyl alcohol. In addition, two molecules incorporating deuterium, the heavy isotope of hydrogen, are known: deuterated hydrogen cyanide and heavy water.

The Absorption of Light in Photosynthesis

The first steps in photosynthesis are the absorption of light by a pigment molecule and the delivery of the absorbed energy to other molecules capable of entering into chemical reactions

by Govindjee and Rajni Govindjee

When the sun shines on the leaves of a green plant, some of the radiant energy is utilized to promote chemical reactions, with the ultimate result that water and carbon dioxide are converted into oxygen and organic compounds. Photosynthesis has been summarized in this way since the end of the 18th century. The summary is essentially correct, but it describes the process only in terms of what flows into and out of the plant. Today a more detailed and more precise explanation is sought; we want to know what happens inside the illuminated leaf. It is not sufficient to say that light "promotes chemical reactions." Rather, the molecular mechanism by which light is absorbed and by which its energy is utilized must be identified.

In this article we are concerned mainly with the first steps in photosynthesis: the absorption of light by a specific molecule and the transfer of that energy from one molecule to another, as in a bucket brigade, until it is eventually conveyed to those few molecules that participate in chemical reactions. These initial processes are called the "primary events" of photosynthesis. They are physical in nature, and they must be completed before the chemical activities of photosynthesis can begin.

Electron Flow in Photosynthesis

Investigations of the primary events have been hindered by the speed with which the events take place and by their complexity and inaccessibility; many can be observed only in the living cell. Most of the experiments intended to explore their sequence have by necessity been indirect. Many of them have been quite ingenious, however, and they have revealed several important characteristics of the system by which the energy of sunlight is made available to the photosynthetic machinery of the plant.

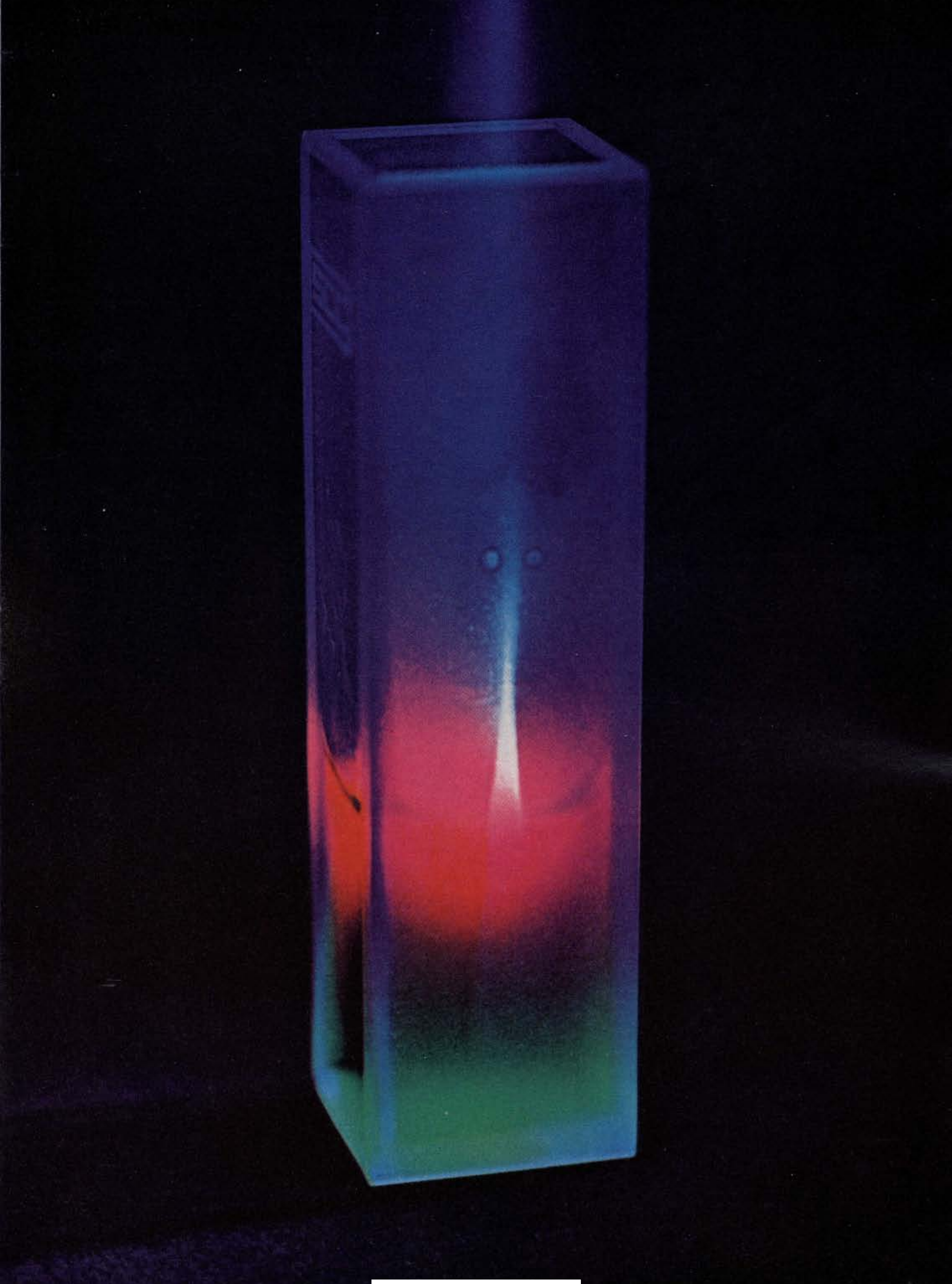
The major organic products of photosynthesis are carbohydrates: substances, such as sugars and starches, whose composition is some multiple of the empirical formula (CH_2O). Because carbohydrates appear superficially to be compounds of carbon and water, it was thought for many years that photosynthesis consisted in splitting carbon dioxide (CO_2), which would allow the oxygen to escape as a diatomic gas (O_2) and free the carbon to combine with water.

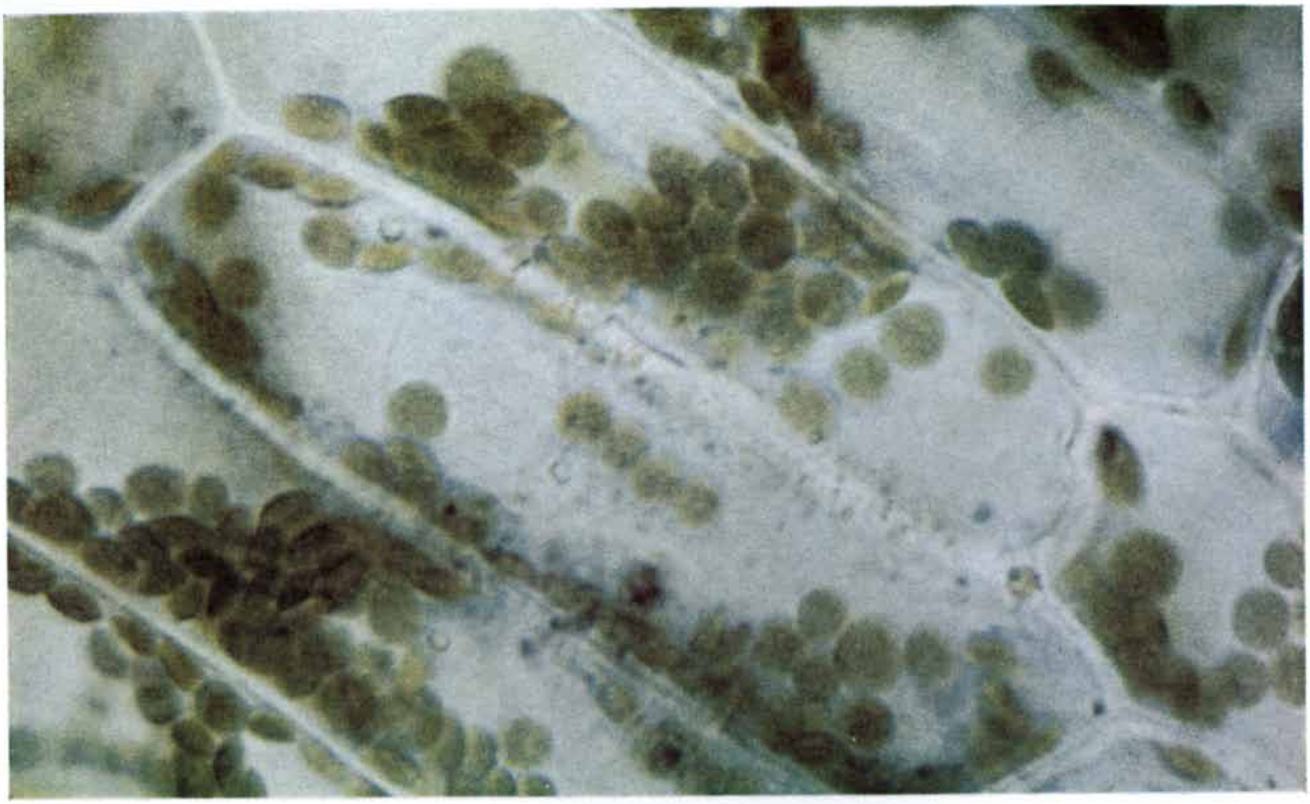
It is now known that this scheme is wrong. Neither carbon dioxide nor water can properly be said to be split or decomposed in photosynthesis. The net effect of the process is instead to transfer hydrogen atoms from water to carbon dioxide; the oxygen evolved comes from the H_2O, not the CO_2. Because the process takes place in water solution it is not necessary to actually move a complete hydrogen atom; if an electron is transferred, a hydrogen nucleus, or proton, can be drawn later from the aqueous medium to complete the atom. Chemical processes of this kind, in which electrons are transferred from one molecule to another, are called oxidation-reduction reactions. The molecule that has lost electrons is said to have been oxidized; the one that has received them is said to have been reduced. Thus in photosynthesis water is oxidized and carbon dioxide is reduced.

Ordinarily, of course, water does not reduce carbon dioxide, and it is not oxidized by it. For the reaction to proceed inside the plant cell, energy must be supplied. The energy requirement of an oxidation-reduction reaction is commonly measured in volts, and the electron transport involved in photosynthesis proceeds against an energy potential of about 1.2 volts.

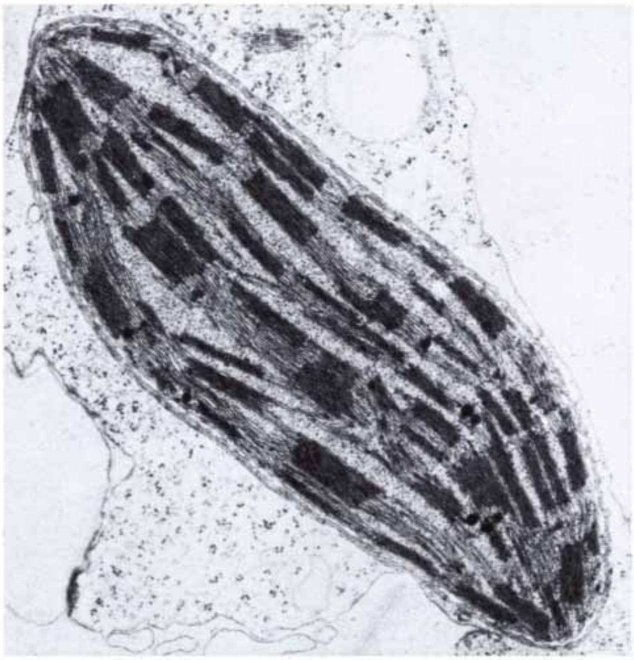
An electrochemical gradient of 1.2 volts represents a rather large barrier, and there is reason to believe that in photosynthetic organisms it is not overcome by a single quantum of light; it appears instead that two quanta are required for the transport of each electron. This hypothesis is supported by numerous recent experiments; moreover, it accords well with an important observation made almost 20 years ago by the late Robert Emerson of the University of Illinois. Emerson found evidence that there are two pigment systems in plants that preferentially absorb light of slightly different wavelength (or color), implying that electron transport takes place in two main stages and involves two photochemical events [see "The Role of Chlorophyll in Photosynthesis," by Eugene I.

FLUORESCENCE of the plant pigment chlorophyll is stimulated by illumination with blue light. The natural green color or the chlorophyll can be seen at the bottom of the container in the photograph on the opposite page. The beam of light, made visible by smoke particles suspended in the air, enters from the top and is absorbed by the pigment, which is in an ether solution. Some of the energy of the absorbed light is dissipated as heat; the rest is reradiated as fluorescence in the red part of the spectrum. Chlorophyll in the leaves of a living plant also fluoresces, but only weakly; most of the energy that in solution is reemitted as red light is applied to the work of photosynthesis in the plant. The photograph was made in the laboratory of Alfred T. Lamme at Columbia University, with chlorophyll extracted from the seaweed *Ulva* by Robert K. Trench of Yale University.

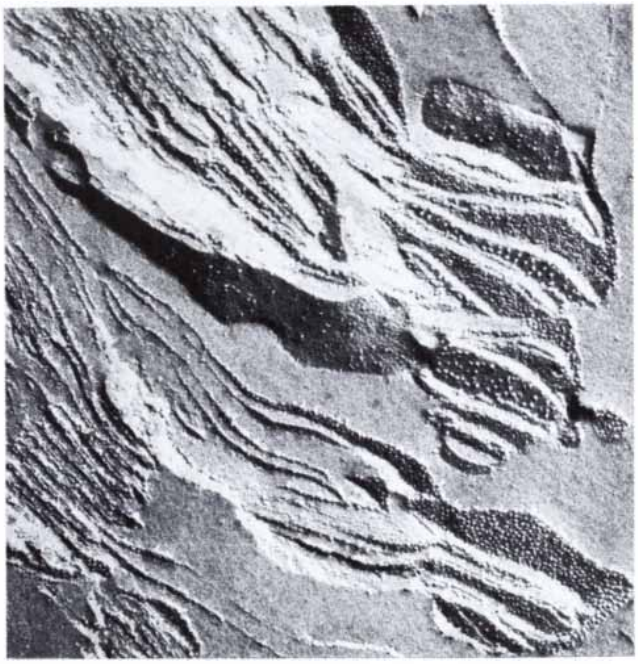




SITE OF PHOTOSYNTHESIS in all higher plants and most algae is the chloroplast, an organelle having an intricate internal structure. In the photograph above, made through an optical microscope, the chloroplasts are the small green bodies, and they alone are responsible for the color of the plant; the cells themselves are translucent. Some algae contain only one chloroplast per cell; higher plants can have as many as 1,000. The photograph was made by D. J. Paolillo of the University of Illinois at Urbana-Champaign.



SINGLE CHLOROPLAST is shown in cross section in an electron micrograph made by Paolillo and R. Chollet at the University of Illinois. The most prominent feature of the chloroplast in this view is the network of lamellae, the membranes on which the photosynthetic pigments are located. The lamellae are organized in dense stacks called grana, where many individual vesicular membranes, the thylakoids, are pressed together. The loosely stacked membranes connecting the grana are called stroma lamellae and the space not occupied by membranes is filled with stroma matrix.



STRUCTURE OF THE LAMELLAE is visible in an electron micrograph made by C. J. Arntzen, also of the University of Illinois. The specimen was prepared by freezing compressed chloroplast fragments, fracturing the frozen pellet and casting a metal replica of the fractured surface. At lower left is a granum; elsewhere individual thylakoid membranes have been pulled apart longitudinally, revealing interior surfaces covered with grains that resemble cobblestones. There are grains of various sizes and each size may be associated with a different function in the process of photosynthesis.

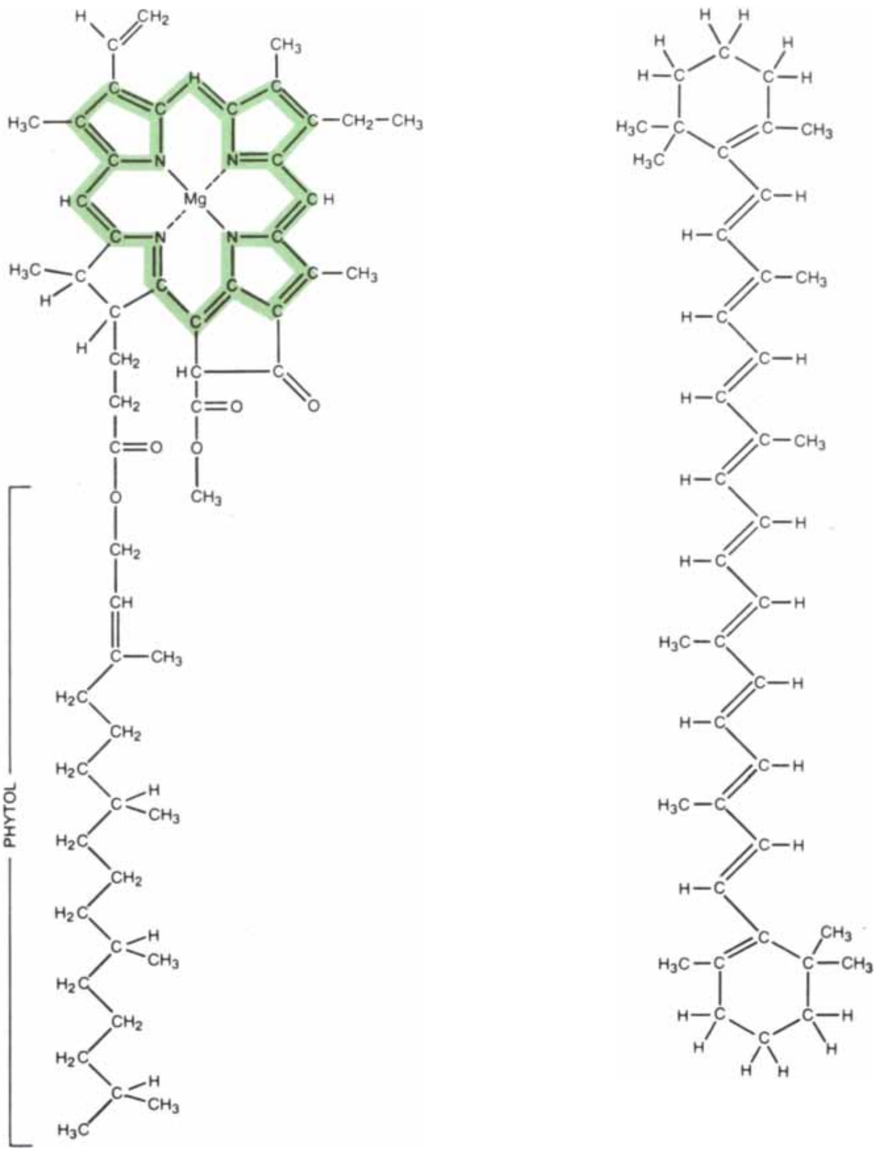
Rabinowitch and Govindjee; SCIENTIFIC AMERICAN, July, 1965].

Not all the molecules associated with the two pigment systems have been identified, nor are all their relations yet clear. Nevertheless, it is possible to draw at least a tentative map of the electron-transport pathway [*see illustration on page 74*]. The scheme is based on a model proposed by Robert Hill and Fay Bendall of the University of Cambridge [see "The Mechanism of Photosynthesis," by R. P. Levine; SCIENTIFIC AMERICAN, December, 1969].

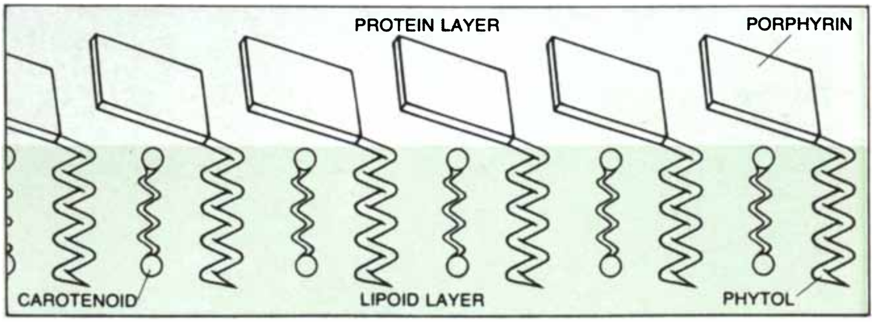
In the modified Hill-Bendall model, light striking the aggregation of molecules designated pigment system II results in the transfer of an electron from a donor called *Z* to an acceptor called *Q*. Recent work on the kinetics of oxygen evolution by Pierre Joliot of the Institut de Biologie Physico-Chimique in Paris and Bessel Kok of the Research Institute for Advanced Studies in Baltimore has shown that a molecule of oxygen is evolved after *Z* has given up four electrons, and thus accumulated four positive charges; *Z* is eventually restored to neutrality by scavenging four electrons from two water molecules. The oxygen atoms from the water form an O_2 molecule and the four protons enter the solution as positive ions.

From reduced *Q* the electrons are transferred, with the electrochemical gradient, to pigment system I. Several kinds of molecule serve as intermediaries, including at least one cytochrome (one of the proteins that also serve as electron-carriers in cellular respiration), an unidentified compound *B*, a plastoquinone and a copper-containing protein, plastocyanin. On absorbing a quantum of light, pigment system I promotes an electron from a "reaction-center" chlorophyll to another acceptor molecule, labeled *X*. The oxidized reaction center is then reduced by an electron that flows from reduced *Q* via the intermediates mentioned above. The reduced *X* then donates the electron to the iron-and-sulfur-containing protein ferredoxin, reducing it. The ferredoxin, with the help of the enzyme ferredoxin-$NADP^+$ reductase, ultimately reduces nicotinamide adenine dinucleotide phosphate ($NADP^+$), a molecule with many roles in metabolism. (In its reduced form it is abbreviated NADPH.) This sequence of events is repeated for each of the four electrons donated by *Z*, so that eight electron transfers, and eight quanta of light, are required for each molecule of oxygen evolved.

Energetically the transfer of electrons from reduced *Q* to the oxidized reaction-



PIGMENT MOLECULES are distinguished by systems of conjugated, or alternating, single and double bonds. When the pigment absorbs light an electron circulating throughout the system of bonds enters an excited state. In chlorophyll (*left*) the conjugated bonds (*colored band*) are in a complex ring called porphyrin. Attached to the ring is a "tail" of phytol, made up of carbon atoms joined mostly by single bonds. Shown is chlorophyll *a*; other forms differ from it only slightly. In the carotenoid pigments (*right*) the conjugated bonds are located in a straight chain of carbon atoms that has a cyclic ring at each end.



ARRANGEMENT OF PIGMENTS in the lamellar membrane may be governed by the physical characteristics of the molecules. In this speculative model, the phytol tail, which is hydrophobic, or repellent to water, projects into the hydrophobic lipoid layer. The porphyrin ring, a hydrophilic, or water-loving, group, associates with the hydrophilic protein layer. Carotenoids are hydrophobic and are probably found in the lipoid portion.

center molecule is a downhill process and is coupled with the production of adenosine triphosphate (ATP), the ubiquitous energy-carrying molecule, from adenosine diphosphate (ADP) and inorganic phosphate. The existence of a cyclic flow of electrons around system I has also been documented. The electrons from reduced *X*, instead of reducing $NADP^+$, are cycled back to the reaction center; when this cyclic flow goes via plastoquinone it is coupled with the production of ATP. (Other sites have been implicated in ATP production, but their role is not yet confirmed.) The importance of cyclic electron flow varies from plant to plant, but in most cases its contribution to photosynthesis is rather small. It is ATP and NADPH that effect the reduction of carbon dioxide and mediate its introduction into a carbohydrate cycle [see "The Path of Carbon in Photosynthesis," by J. A. Bassham; SCIENTIFIC AMERICAN, June, 1962].

Pigment Molecules

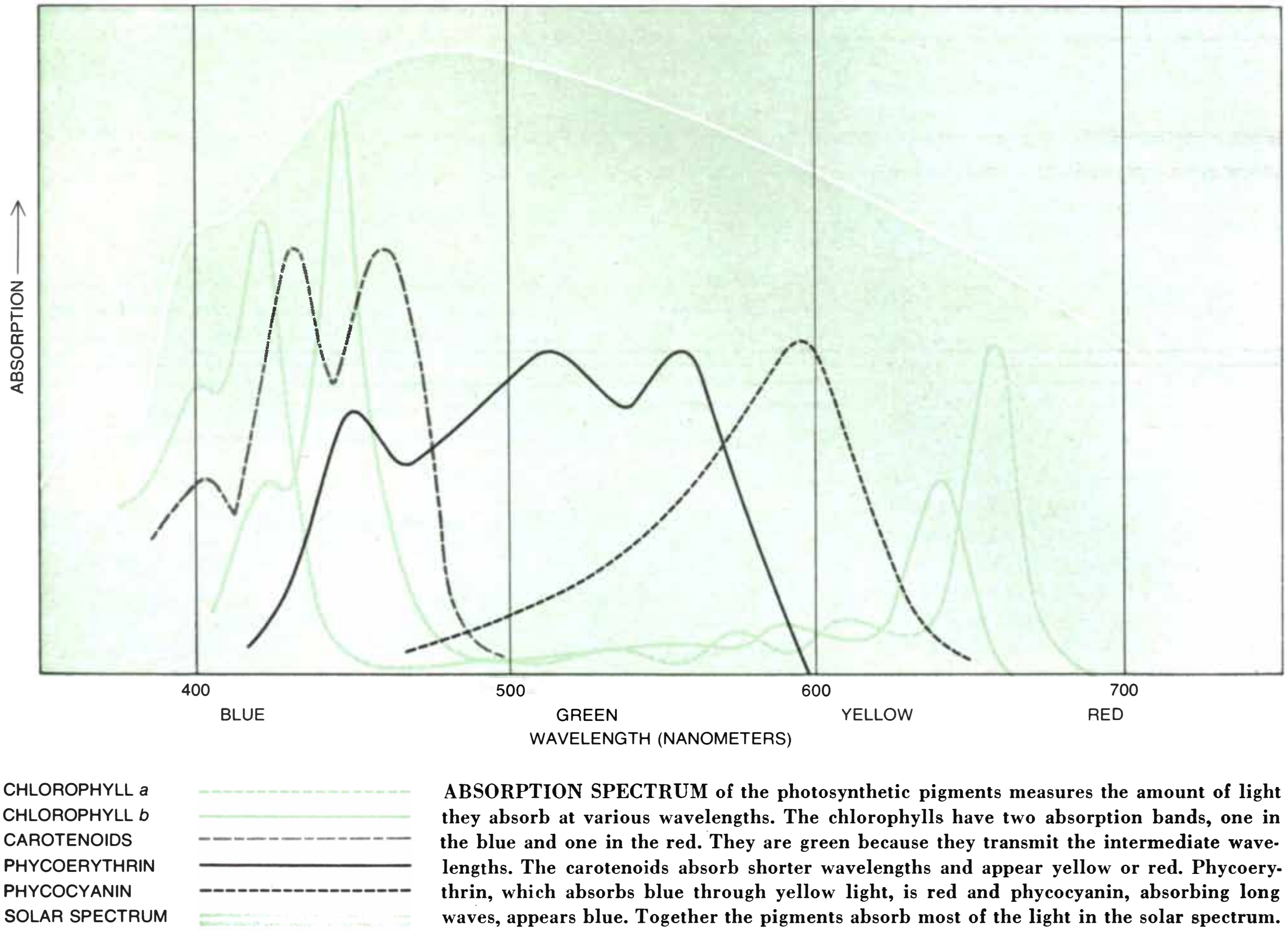
The photosynthetic apparatus of the higher plants is organized inside chloroplasts, the cellular organelles that give plants their characteristic green color. The chloroplasts are complex structures, separated from the cytoplasm of the cell by a membrane and apparently having some autonomy: each has a bit of the genetic material DNA and is able to synthesize some of the proteins it requires independent of the cell nucleus.

Inside the chloroplast is an elaborately folded network of membranes termed lamellae. In chloroplasts from most plants the lamellae form structures called grana, which are separated by a material known as stroma. The grana appear to be dense stacks of membranous sacs, which are the basic units of the lamellae. These sacs, called thylakoids, contain lipids, proteins and pigments. It is on and perhaps between the thylakoids that the business of photosynthesis is transacted.

A further level of structural organization can be inferred in the chloroplast, even if it cannot be resolved with certainty in the electron microscope. Experiments performed more than 40 years ago by Emerson and William A. Arnold, who is now at the Oak Ridge National Laboratory, suggested that a minimum of 2,400 chlorophyll molecules are required to evolve one molecule of oxygen. An aggregate of this size, they proposed, should be considered the ultimate photosynthetic unit. We now know that eight photochemically driven electron transfers are required to generate one molecule of oxygen, so that a smaller unit is plausible, one containing 300 pigment molecules. This is the modern photosynthetic unit, the smallest unit capable of photochemical action.

Embedded in the thylakoid membranes are the pigment molecules that initiate the process of photosynthesis. Pigments are substances that by definition strongly absorb visible light. Most absorb only in certain regions of the spectrum and transmit light of all other wavelengths; as a consequence they appear colored. For example, chlorophyll, the most important plant pigment, absorbs both the longer and the shorter waves in the visible spectrum: red and orange and blue and violet. The transmitted wavelengths, chiefly the yellow and green in the middle of the visible spectrum, combine to yield the green of grass and trees [*see illustration below*].

The majority of organic molecules absorb most strongly in the ultraviolet; the various pigments of the chloroplast ab-



ABSORPTION SPECTRUM of the photosynthetic pigments measures the amount of light they absorb at various wavelengths. The chlorophylls have two absorption bands, one in the blue and one in the red. They are green because they transmit the intermediate wavelengths. The carotenoids absorb shorter wavelengths and appear yellow or red. Phycoerythrin, which absorbs blue through yellow light, is red and phycocyanin, absorbing long waves, appears blue. Together the pigments absorb most of the light in the solar spectrum.

sorb at the longer wavelengths of visible light because they have chains or rings of carbon atoms connected by "conjugated," or alternating, single and double bonds. In chlorophyll the system of conjugated bonds is located in a ring, and it encompasses nitrogen atoms as well as carbon atoms [*see top illustration on page 71*]. The ring structure is one of the class of compounds called porphyrins, which are found widely in both plants and animals. In chlorophyll the central cavity of the ring is occupied by a magnesium atom; the function of the magnesium is not yet fully understood. Porphyrin rings in the blood protein hemoglobin and in the cytochromes contain an atom of iron instead of magnesium.

Attached to the porphyrin ring in chlorophyll is a long hydrocarbon "tail," the phytol chain. It consists of carbon atoms linked together, but only one of the bonds is double; for this reason the phytol chain does not appear to play an important part in determining the chlorophyll absorption spectrum. Its function may be to anchor the molecule in the thylakoid membrane. The phytol chain is hydrophobic, that is, it repels water but has an affinity for oils and fats; the porphyrin ring, on the other hand, is hydrophilic, being drawn to water. It has therefore been proposed that the phytol portion of the molecule may project into the lipid layer of the thylakoid membrane, securing the porphyrin at a particular location in the protein part of the membrane [*see bottom illustration on page 71*].

Three major varieties of chlorophyll are distinguished by small differences in structure in one region of the porphyrin ring. Chlorophyll *a*, the most abundant form, found in all higher plants and in all algae, has a methyl group ($-CH_3$) at that position. In chlorophyll *b*, which is found only in the higher plants and in the green algae, the methyl group is replaced by an aldehyde group ($-CHO$). Finally, in bacteriochlorophyll, the form of the pigment found in photosynthetic bacteria, a single bond is substituted for one of the double bonds in the conjugated ring, and two hydrogen atoms are added.

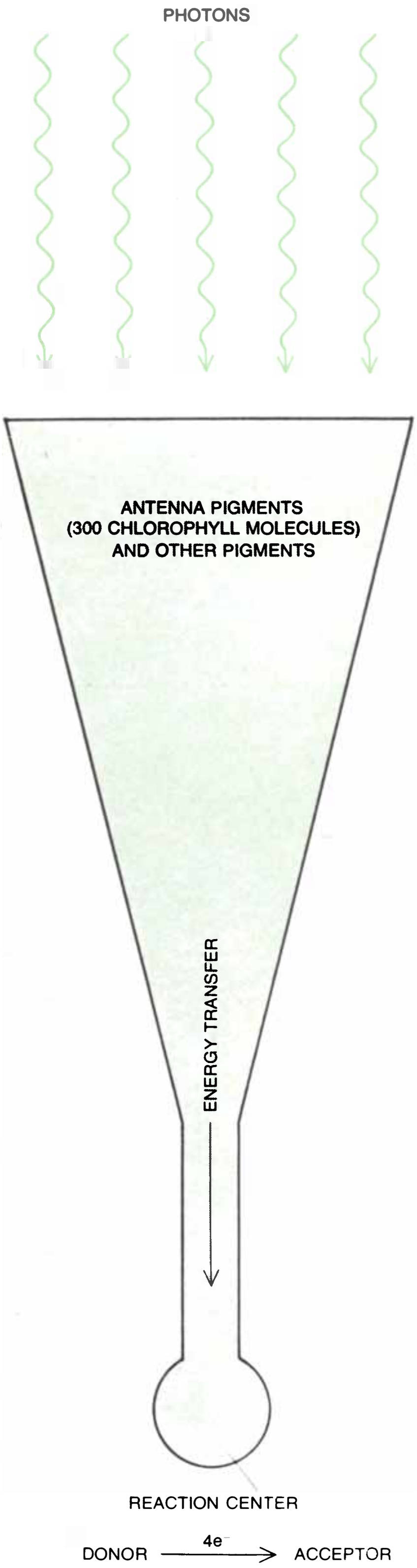
In addition to these structurally distinct molecules, various spectral forms of chlorophyll have been detected in living cells: they can be distinguished only by differences in their absorption spectra. C. Stacy French of the Carnegie Institution of Washington's Department of Plant Biology has detected forms of chlorophyll *a* whose maximum absorption of red light is at wavelengths of 660, 670, 680, 685, 690 and 695 to 720 nanometers. The variations are probably produced by chlorophyll molecules in different environments; they may be aggregated differently with other chlorophyll molecules, or they may be associated with proteins.

Although the chlorophylls are undoubtedly the most important plant pigments, they are not the only ones present and they are not the only ones that participate in light absorption. The carotenoids, found in bacteria, algae and higher plants, absorb mainly blue wavelengths and are yellow, orange or red in color. (Their name derives from the same root as "carrot," and they give the carrot its orange color.) They consist of long chains of carbon atoms linked by conjugated single and double bonds and bearing a six-carbon ring at each end [*see top illustration on page 71*]. The carotenoids are divided into two classes: the carotenes, which are hydrocarbons and therefore consist of hydrogen and carbon only, and the carotenols, which are alcohols and ketones and contain oxygen as well as carbon and hydrogen. The final category of pigments is the phycobilins, named for their resemblance to pigments in the bile of animals. They include the red phycoerythrins and the blue phycocyanins and are found only in the blue-green and red algae. The molecules of both types have an open ring related to the porphyrin structure of the chlorophylls, but in the phycobilins the ring is bound to a protein component.

Light Absorption

When a photon is absorbed by an atom or molecule, its effect is to change the configuration of the electronic charge associated with the valence, or outer, electrons surrounding the atomic nucleus (or nuclei). Because the new configuration has more energy than the "ground," or lowest, state, the atom or molecule is said to be in an excited state. The transition from the ground state to an excited state can take place only under certain conditions prescribed by the laws of quantum mechanics. The electrons can occupy only specified, distinct energy states; regions between states are forbidden to them. Moreover, the energy of the absorbed photon must exactly match the energy of the transition. Since the energy of a photon is inversely proportional to its wavelength, only certain wavelengths can be absorbed by a particular atom or molecule [*see top illustration on page 75*].

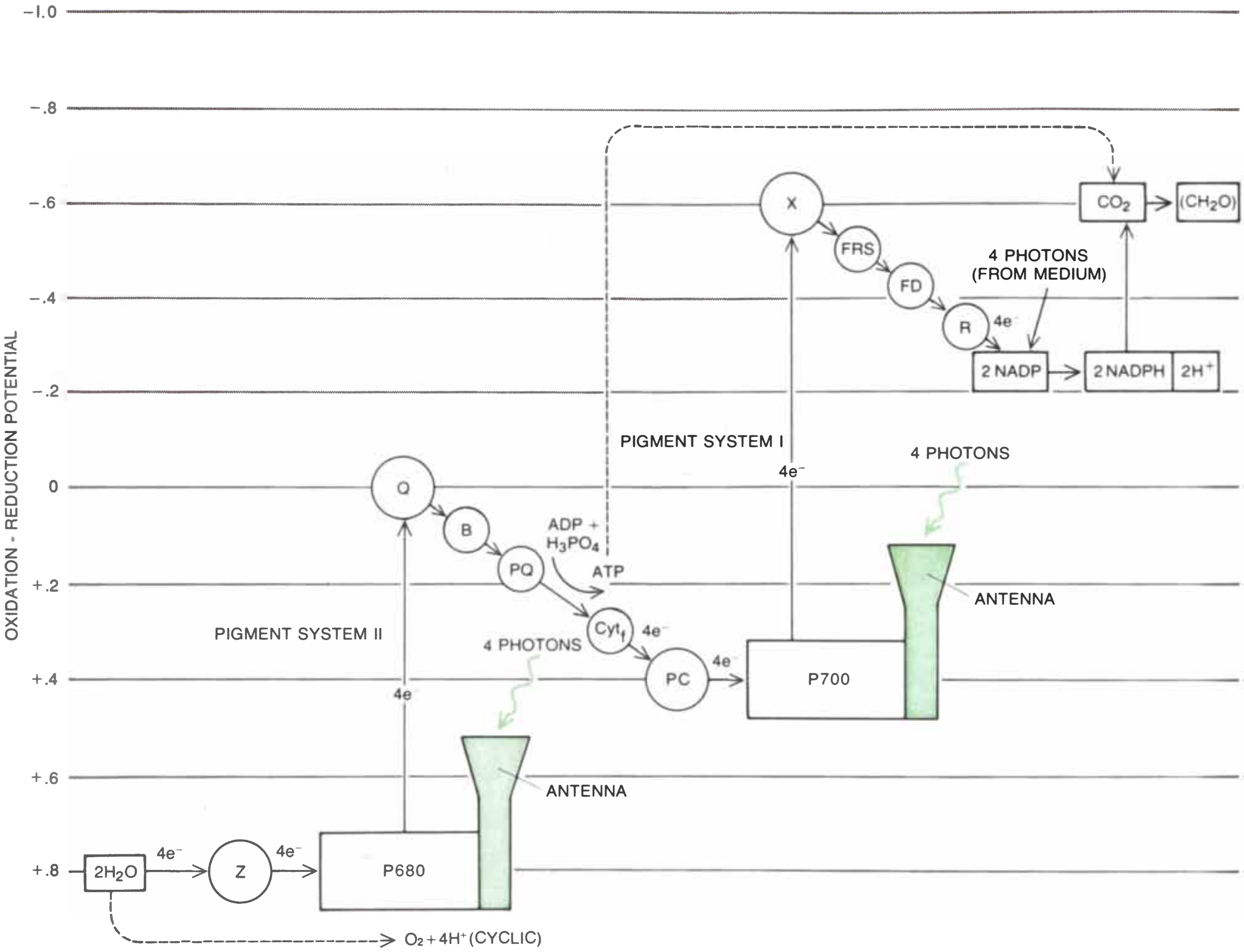
In atoms these restrictions are quite



PHOTOSYNTHETIC UNIT is thought to consist of about 300 chlorophyll molecules; in the higher plants and algae that seems to be the minimum number necessary to bring about the transport of a single electron. The unit includes the bulk pigments, which serve as an energy-gathering "antenna," and a reaction center, where energy is trapped and utilized to promote chemical reactions.

confining, since the permitted energy levels are few and transitions between them are large. As a consequence the absorption spectra of atoms usually consist of a relatively few narrow lines. In molecules, however, and particularly in pigment molecules, a number of factors tend to broaden the lines into bands. For one thing, the conjugated bonds in pigments bring about a great proliferation of allowed states because each double bond adds a pair of electrons shared by the conjugated system as a whole. The effect of a large system of conjugated bonds is therefore to decrease the transition gap between the ground state and the first, or lowest, excited state and to create many additional excited states just above the first. It is this effect that is responsible for shifting the absorption spectrum of pigment molecules from the ultraviolet region of the spectrum into the visible region; because of the smaller transition to the first excited state pigments are able to absorb photons of lower energy and hence greater wavelength.

Two other factors also increase the number of possible quantum states in molecules: the vibrational and the rotational energy of the molecules. These motions are also confined to discrete energy levels, but the levels are much more closely spaced, so that the ground state and each of the excited states broaden into a manifold of substates. In approximate terms the transition between the ground state and the first excited state in chlorophyll *a* represents an energy difference of between one and two electron



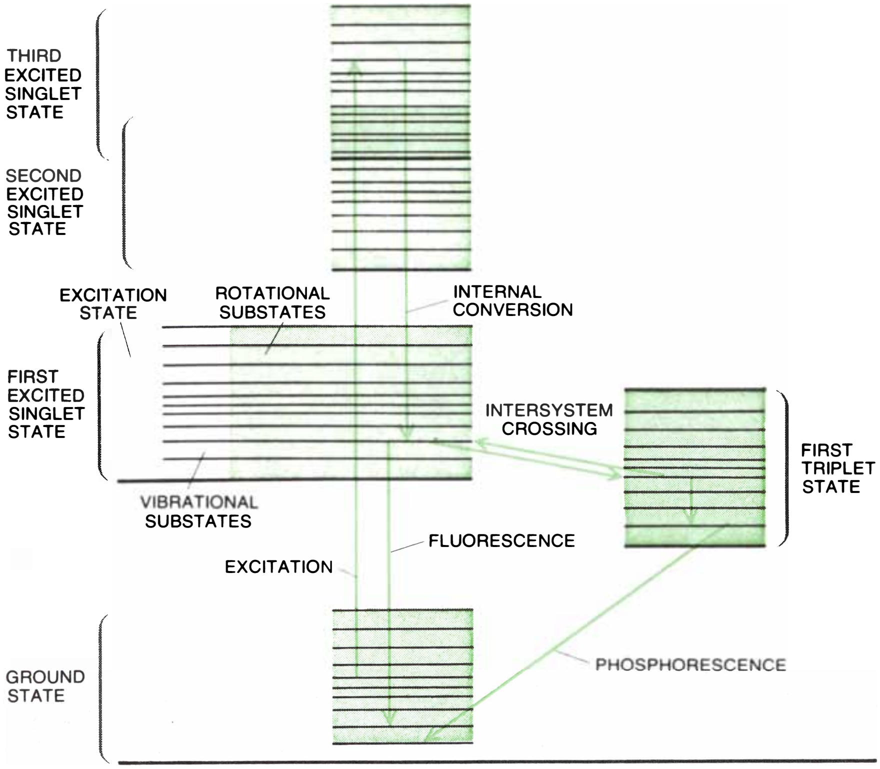
FLOW OF ELECTRONS from water to carbon dioxide proceeds against an electrochemical gradient of 1.2 volts and requires two photochemical events. Four electrons must be transferred, one at a time, to liberate a molecule of oxygen and reduce a molecule of CO_2 to carbohydrate. The process begins with the absorption of a photon by the antenna of pigment system II. The energy of excitation is conveyed to a chlorophyll molecule in the reaction center of the photosynthetic unit; the molecule is designated P680 because one of the bands in its absorption spectrum is at 680 nanometers. The excited P680 transfers an electron to the acceptor *Q*, and subsequently recovers an electron from the donor *Z*. After *Z* has given up four electrons it regains them by oxidizing two molecules of water. From *Q* the electron is passed through a series of carrier molecules, including *B*, which has not been identified, plastoquinone (*PQ*) and cytochrome *f* (Cyt_f), to plastocyanin (*PC*). Plastocyanin injects the electron into pigment system I. The reaction-center chlorophyll of pigment system I, designated P700, is excited through its own antenna pigments, and promotes an electron to the acceptor *X*. Finally, the electron is passed through ferredoxin reducing substance (*FRS*), ferredoxin (*FD*) and the enzyme ferredoxin-$NADP^+$ reductase (*R*) to nicotinamide adenine dinucleotide phosphate ($NADP^+$), which is thereby reduced to NADPH. NADPH is the primary product of these reactions. In addition, during two of the electron transfers adenosine triphosphate (ATP) is generated from adenosine diphosphate (ADP) and inorganic phosphate. One site for ATP generation is between plastoquinone and cytochrome *f*; the other is associated with system II but has not been located. The NADPH and the ATP drive the process by which CO_2 is incorporated into carbohydrates (multiples of the unit CH_2O). There is also a cyclic system of photosynthesis, in which electrons pass from P700 to *X* and then return through various electron carriers to P700; in this system only ATP is produced.

volts. (An electron volt is the energy acquired by an electron when it is accelerated through a potential of one volt.) The vibrational substates are separated by about .1 electron volt and the rotational substates by about .01 electron volt.

When a photon of appropriate energy strikes a molecule of chlorophyll *a*, the pigment enters an excited state almost instantaneously, within 10^{-15} second. This is an extremely brief interval; in a vacuum light travels only about three ten-thousandths of a millimeter in 10^{-15} second.

There are two main excited states available in chlorophyll *a*. The absorption of red light (with a wavelength of about 680 nanometers) raises the molecule to the lowest and most important of the levels, called the first excited singlet state. Blue light (with a wavelength of about 440 nanometers) promotes the molecule to the third excited singlet state. (Transition to the second excited singlet state, between the first and the third, is weak.) The upward transition may begin from any of the various vibrational and rotational substates of the ground state and end on any of the substates of the excited state.

Having reached an excited state, the chlorophyll molecule cannot immediately and directly apply its energy to the tasks of biochemistry, nor is useful work the only possible outcome of the absorption of a photon. The excited molecule can give up its energy in any of a number of ways, all of which conform to the same laws of quantum mechanics that govern light absorption.

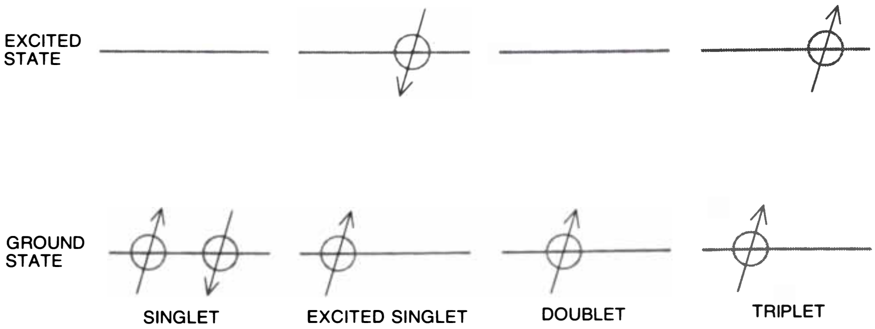


EXCITATION OF AN ELECTRON in a pigment molecule promotes the electron to one of a few discrete energy states; intermediate levels are forbidden. Superimposed on the permitted energy states are substates representing the vibrational and rotational energy of the molecule; the substates also are discrete, but they are closely spaced. When an electron absorbs a photon it is first elevated to an energy level called an excited singlet state. Depending on the energy of the photon, the electron may reach the first or some higher excited singlet state; it quickly subsides to the first, however, by dissipating part of its energy as heat in a process called internal conversion. In the first excited singlet state the molecule can utilize its energy in a chemical reaction, lose energy by internal conversion, reradiate it as fluorescence or enter another excited state, the triplet state. Because the triplet state is longer lived than the excited singlet states, light is emitted by molecules in the triplet state only after a delay and is called phosphorescence instead of fluorescence.

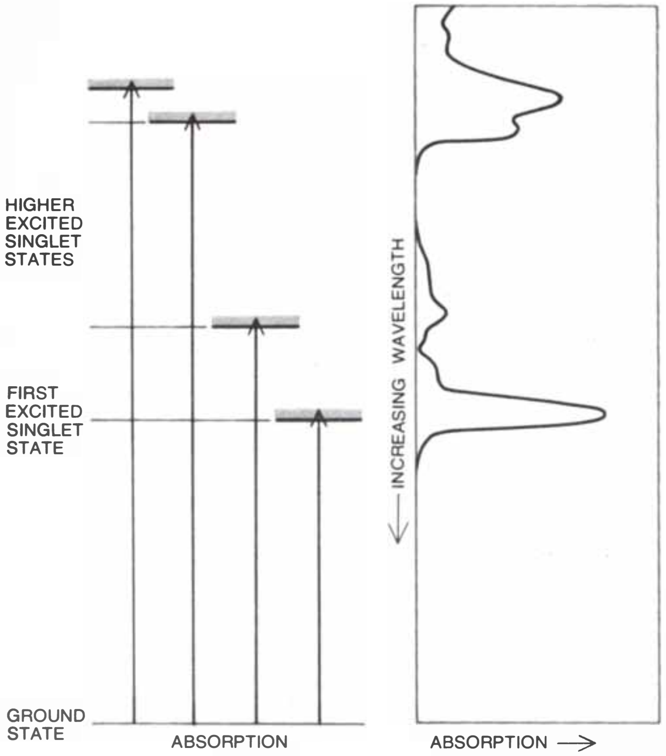
The Fate of Excitation Energy

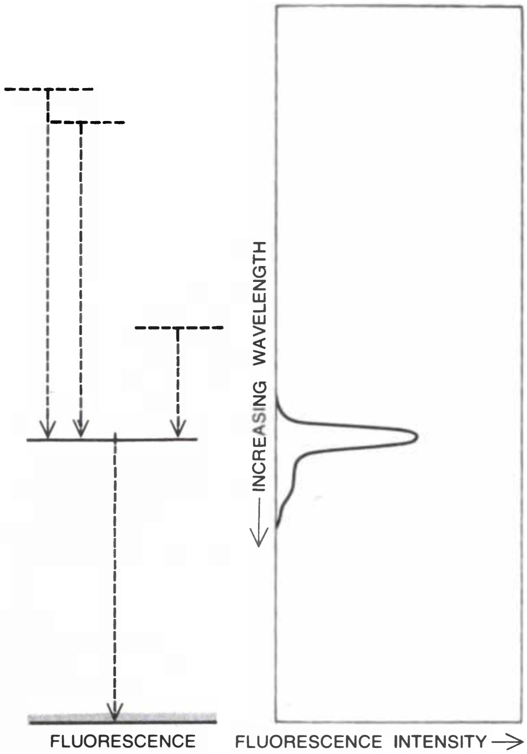
For the purposes of photochemistry the most important energy level in chlorophyll *a* is the first excited singlet state. Molecules elevated to higher excited states return to the first excited singlet state rapidly (in from 10^{-14} to 10^{-13} second), so that they reside in the higher state too briefly to enter into any competing processes. The return to the first excited state is achieved through the small transitions separating the vibrational and rotational sublevels. Each of these transitions is so small that the wavelengths associated with them are not perceived as radiation, only as heat. By the same mechanism a molecule that has been raised to one of the higher sublevels of the first excited state will quickly decay to one of the lower substates. This process is called internal conversion; since it results in the degradation of the received energy to heat, it contributes only to the kinetic motion of the molecules, not to photochemistry.

Because of internal conversion a high-energy photon of blue light is of no more use to a plant than a relatively low-energy red one. The squandering of the energy of short-wavelength light, however, is less extravagant than it may at first seem. Although individual photons of short wavelength are the most energetic, most of the energy of the solar flux is distributed in the yellow and orange and beyond. There are simply greater numbers of photons at longer wave-

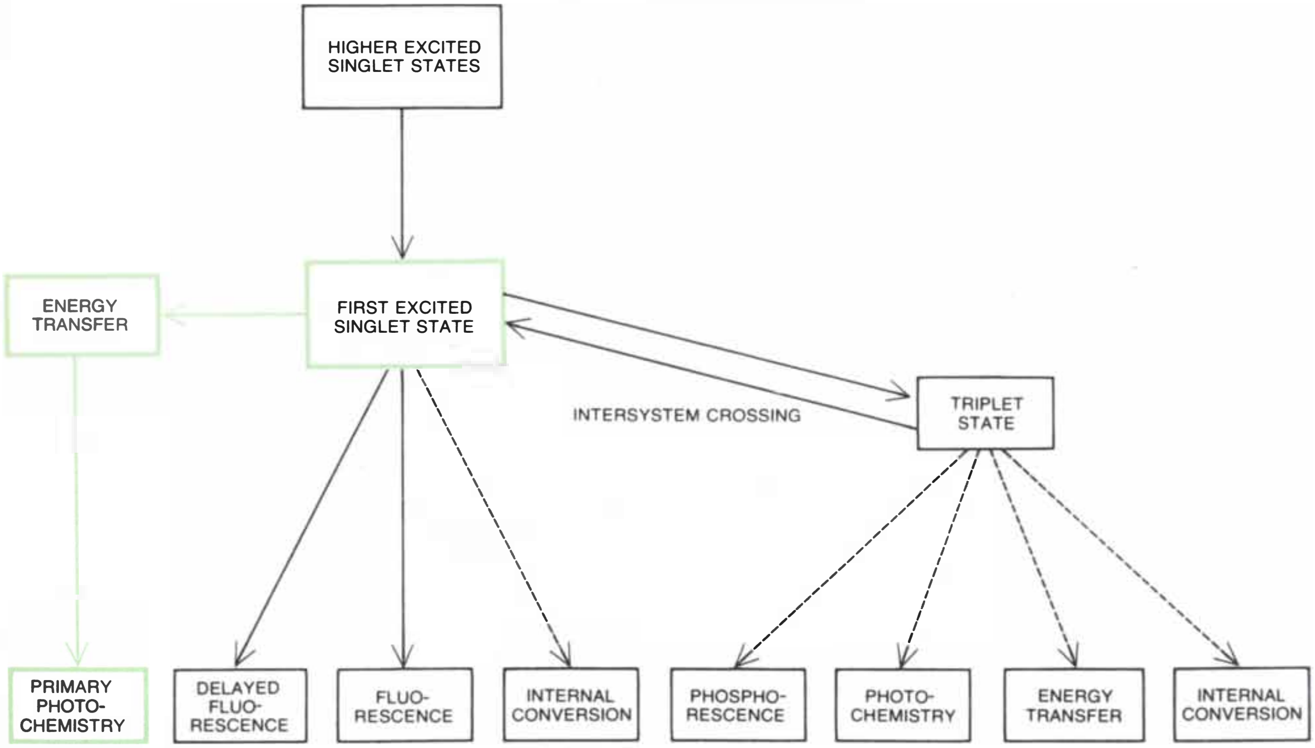


SINGLET AND TRIPLET STATES are defined by the "spin" of the outer electrons in an atom or molecule. In this diagram the arrows represent the spin axis. If the spins are antiparallel, the molecule is in a singlet state; if they are parallel it is in a triplet state. The doublet state requires an unpaired electron. "Singlet," "doublet" and "triplet" refer to the number of ways the electrons can orient themselves with respect to a magnetic field.





ENERGY TRANSITIONS associated with light absorption and fluorescence by chlorophyll are constrained by the requirement that the molecule occupy only discrete energy states. Only photons whose energy corresponds to the energy difference between the ground state and an excited state can be absorbed; since the energy of a photon is inversely related to its wavelength, the selective absorption of photons is reflected in the absorption spectrum of the pigment (*diagram and graph at left*). Even though chlorophyll absorbs light in several parts of the spectrum, it fluoresces only in the red (*diagram and graph at right*). The fluorescence is from the first excited singlet state; molecules in higher states decay to this one by internal conversion. Fluorescence is at longer wavelengths than the lowest-energy absorption band because molecules relax to a low substate of the first excited singlet state before fluorescing.



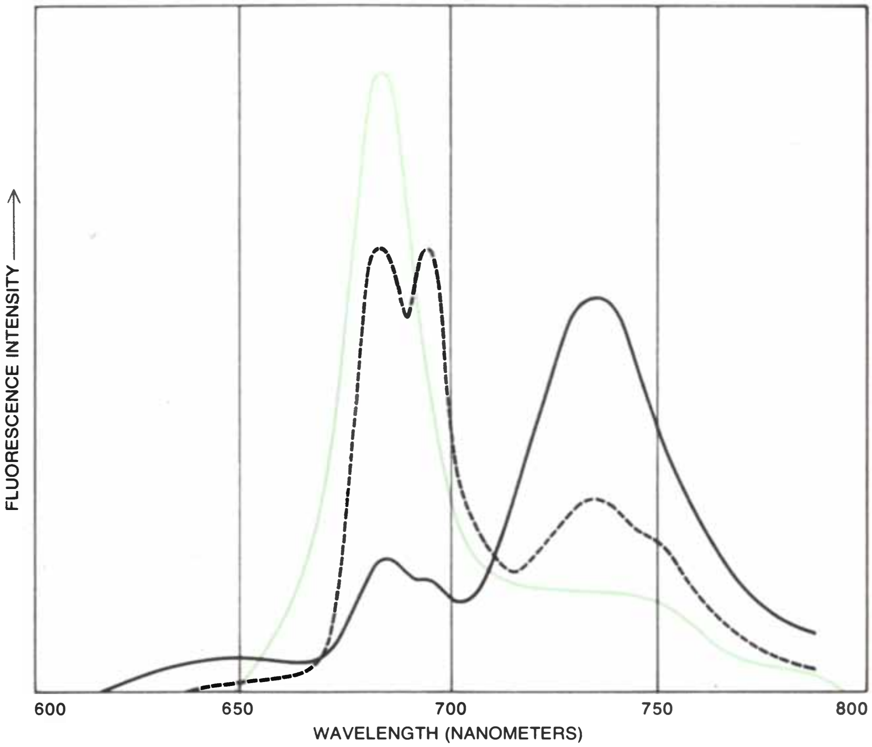
FATES OF EXCITATION ENERGY are of varying probability and of varying importance to the living plant. The pathway indicated in color is the only one known to contribute to photosynthesis. Transitions shown in solid black lines take place in the plant but do not result in useful chemical reactions. Transitions indicated by broken lines have not been observed in living cells.

lengths, and their aggregate energy is much greater.

Once an excited chlorophyll molecule has subsided to one of the lower sublevels of the first excited singlet state, several pathways are open to it. Through further internal conversion it can dissipate its remaining energy of excitation as heat; it can enter another excited state, of a somewhat different character, called the triplet state; it can return to the ground state by emitting a quantum of light as fluorescence; it can transfer its energy to an appropriate neighboring molecule, or it can enter into a chemical reaction, such as those characteristic of pigment systems I and II. The last process is obviously the most important, but it is by no means the only one that takes place.

Singlet and triplet states are distinguished by a quantum number called the total spin number, which has to do with the directions in which the electrons in a molecule "spin." In singlet states the spins of a pair of electrons are antiparallel, or opposite; in triplet states the spins are parallel. Doublet states, which are also found in some photosynthetic pigments, occur when there are unpaired electrons. The triplet state is so named because electron pairs spinning in parallel can align themselves in three ways with an external magnetic field; moreover, they superimpose three additional energy sublevels on the quantum state of the molecule. In the singlet state all orientations with respect to an external field are equivalent, and no additional sublevels are introduced [*see bottom illustration on page 75*].

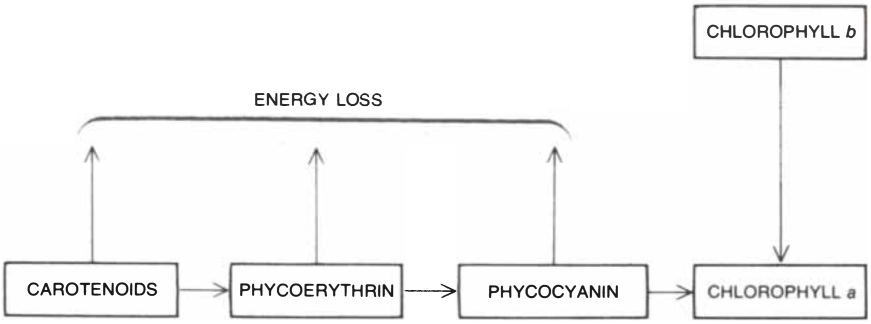
The triplet state is metastable, that is, it is stable for a comparatively long time. Events that can reverse the spin of an electron are relatively uncommon, and transitions between the ground state and the triplet state are therefore rare. The triplet state can be readily entered, however, from an excited singlet state. There is considerable evidence that molecules in the triplet state are present in the chloroplast. An analysis made in our laboratory at the University of Illinois suggested that they are present, but in low concentration; our data suggest that at any given moment less than one chlorophyll molecule in 10 million is in the triplet state. J. S. Leigh, Jr., and L. P. Dutton of the University of Pennsylvania School of Medicine have found evidence for the triplet state of bacteriochlorophyll in photosynthetic bacteria, but only at low temperature and only when the photochemical oxidation-reduction processes driven by the pigments were blocked.



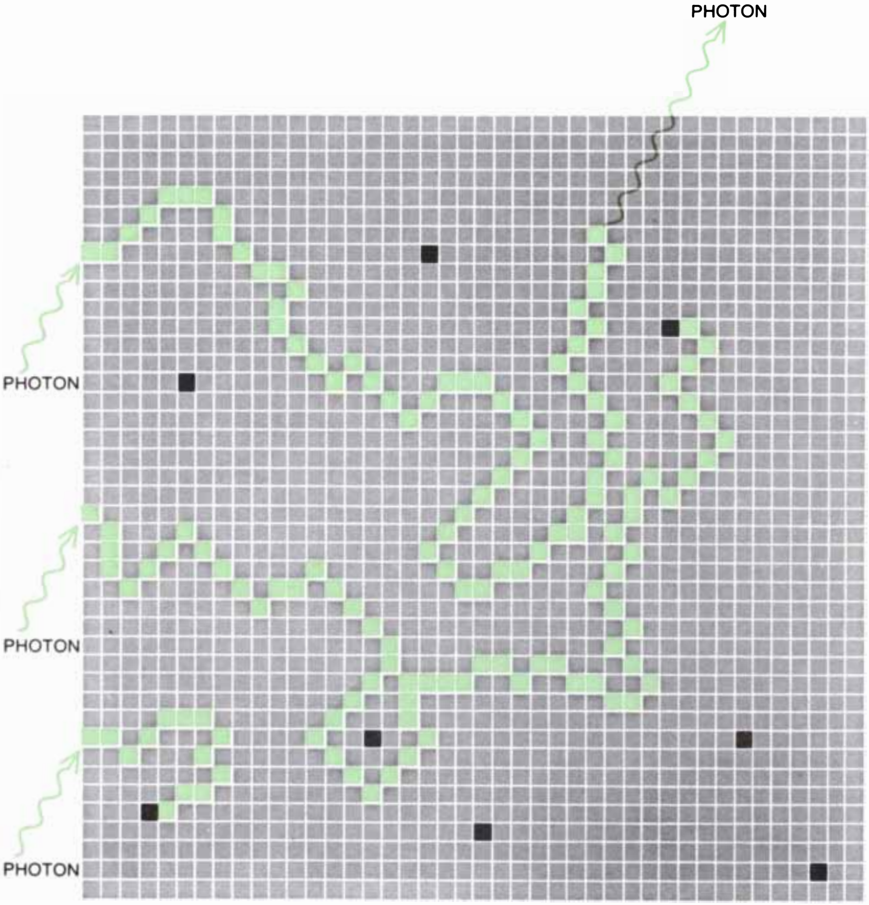
FLUORESCENCE EMISSION SPECTRA can discriminate between pigment molecules serving different functions. Intact chloroplasts at room temperature (*colored line*) have a strong fluorescence peak at about 685 nanometers and a much smaller peak at about 735 nanometers. At low temperature a more complex pattern emerges. Pigment system I (*solid black line*) fluoresces most strongly at 735 nanometers, with minor bands at 684 and 695. In pigment system II (*broken black line*) peaks at 685 and 695 nanometers are more important.

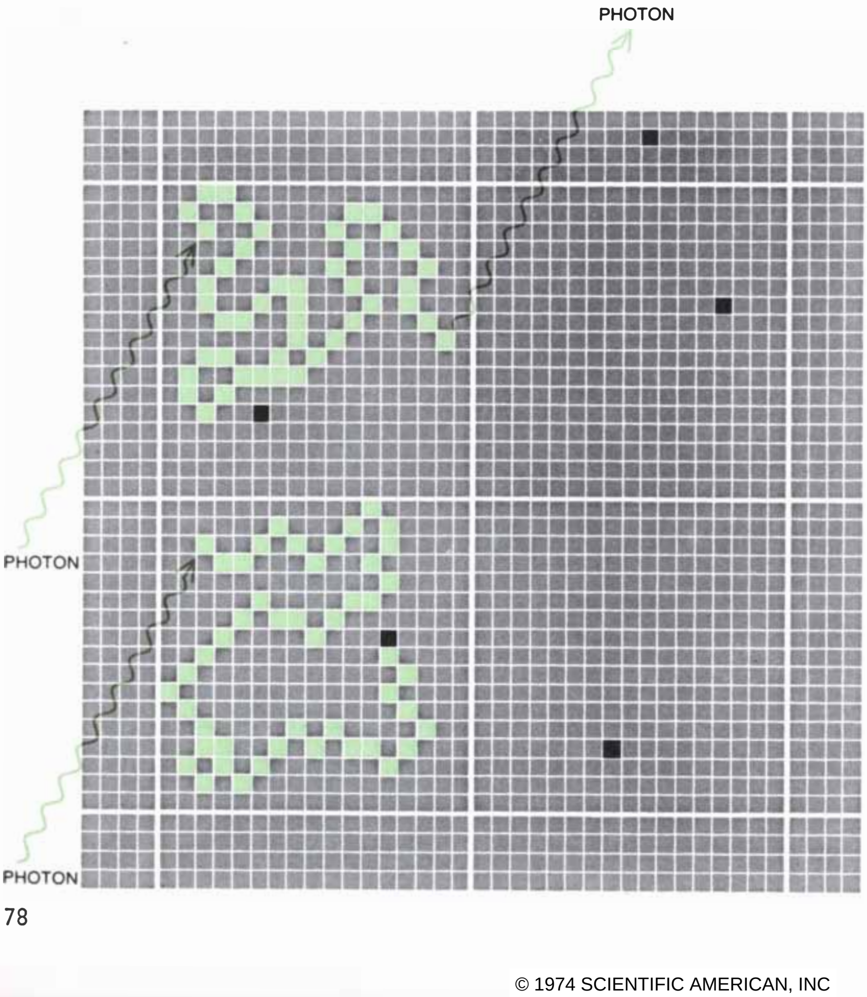
Even if the triplet state is present in illuminated plants, there is no evidence that it forms a part of the main photosynthetic pathway; there are many other ways in which its energy could be dissipated, such as internal conversion and a return directly to the ground state with the emission of light. The latter effect is termed phosphorescence: it is characterized by light emitted after a delay and at a considerably longer wavelength than the light absorbed. A molecule in the triplet state can also return to the first excited singlet state if it absorbs a small quantum of energy; it then has before it all the possible fates of other molecules at that energy level.

Light emitted during a transition from



TRANSFER OF ENERGY from one kind of pigment to another follows an established sequence. The carotenoids, which absorb blue light, pass their energy of excitation on to phycoerythrin. Energy received in this way, as well as the energy of green light absorbed directly by phycoerythrin, is transferred to phycocyanin. Phycocyanin absorbs orange light, and passes the accumulated energy to chlorophyll *a*. In each of these transitions energy is lost by internal conversion, and in some of the transfers energy is also dissipated as fluorescence. Chlorophyll *b* contributes its energy directly to chlorophyll *a* without loss.





a singlet state directly to the ground state is termed fluorescence instead of phosphorescence. Ordinarily there is little delay between absorption and reemission, and the wavelength of the fluorescent light is only slightly longer than that of the light absorbed. The small increase is caused by the loss of energy through internal conversion.

Fluorescence

Fluorescence has provided some of the most valuable techniques available for the investigation of photosynthesis, but for the plant it is an entirely wasteful phenomenon. Ordinarily, however, the magnitude of the effect is rather small. For every 100 photons absorbed, only from three to six are reemitted. For this reason fluorescence in living plants is faint and can rarely be detected by the unaided eye. The fluorescence of chlorophyll in solution, on the other hand, can be quite bright, since the process that would normally drain away most of the energy of excitation—photosynthesis—is disrupted.

Chlorophyll fluorescence is invariably red, even if the exciting light is blue or green or yellow. The fluorescent light is emitted in a transition from the first excited singlet state, and molecules promoted to higher states by more energetic photons merely dissipate part of their energy as heat before fluorescing [*see top illustration on page 76*]. At room temperature the fluorescence spectrum consists of a major band at 685 nanometers and a minor band at about 740 nanometers. When the spectrum is measured at low temperature, bands appear at 685, 695 and 720 nanometers. Pigment system II is primarily responsible for the bands at 685 and 695 nanometers, and

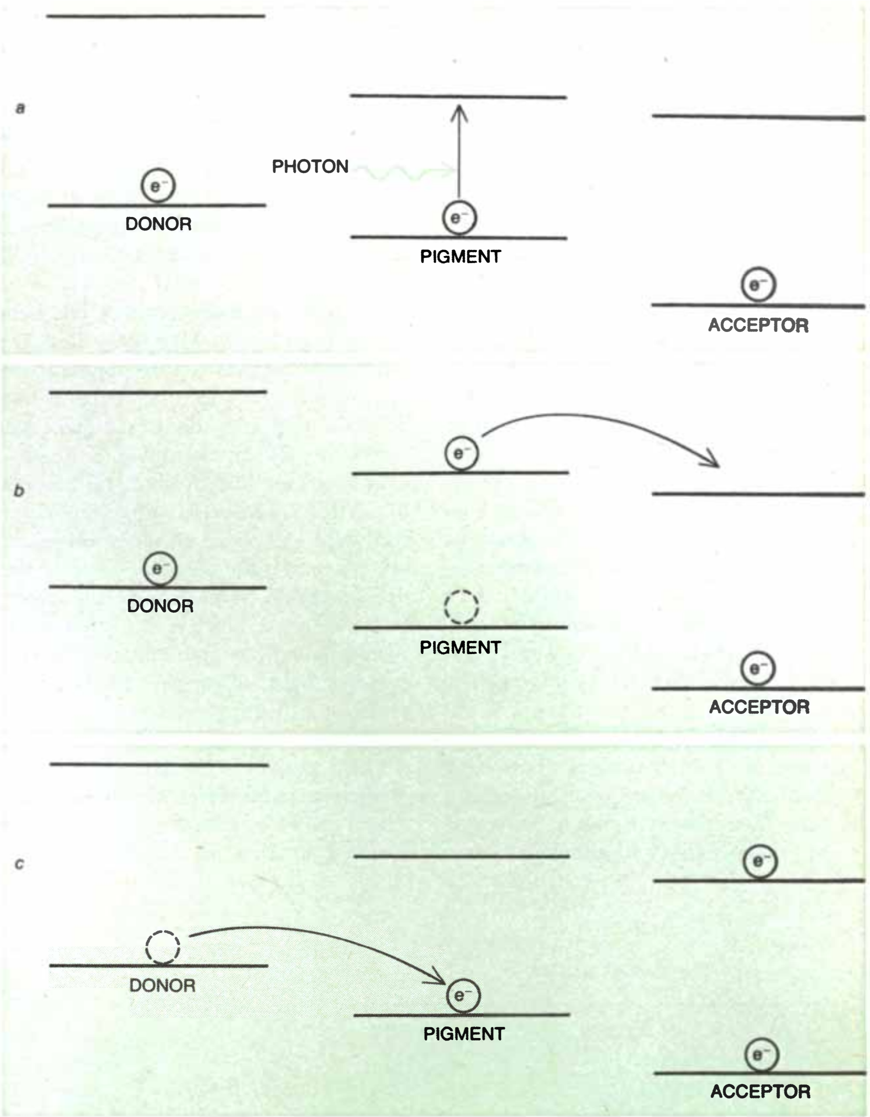
FLOW OF ENERGY through chlorophyll molecules in the thylakoid membrane is thought to be rapid and efficient, but undirected. In the diagrams at left each square represents a pigment molecule; there is one reaction center (*black squares*) for approximately 300 antenna molecules. In the "lake model" (*top*), energy absorbed by any molecule can wander through the entire mass of pigments until it is trapped in a reaction center or until it is reemitted. In the "isolated puddle model" (*bottom*), a single reaction center is committed to each aggregate of about 300 chlorophyll molecules. The energy again meanders randomly, but only within the confines of the 300-molecule unit. The mechanism of energy transfer, which probably involves resonance between nearby molecules, operates with perfect efficiency.

pigment system I for the 720-nanometer band.

In living cells the intensity of chlorophyll *a* fluorescence changes as a function of time. This effect, discovered by the late Hans Kautsky of the University of Heidelberg, consists of a fast change, completed about two seconds after the plant is illuminated, and a slower fluctuation that achieves a steady state only after several minutes. Both temporal patterns could reflect the shifting relation between pigment systems I and II. The fast change is sensitive to the rate of electron flow through the transport system and might be explained as indicating that pigment system II briefly gets ahead of system I, and wastes the surplus energy in fluorescence.

The slow change is thought to be related to temporary alterations in the thylakoid membrane. One hypothesis suggests that if the two pigment systems are spatially close together when the light is first turned on, energy from pigment system II, which is highly fluorescent, might "spill over" into the weakly fluorescent system I, thereby diminishing the overall light emission. If the two systems were then separated, the spillover would be eliminated and the fluorescence would increase. The conformational change in the thylakoid membrane required by this theory might be caused by the movement of ions. An alternative explanation postulates changes in the density or spacing of chlorophyll molecules in pigment system II. Separating the molecules from one another could be expected to decrease internal conversion and thus to increase fluorescence.

The lifetime of an excited state in a fluorescent molecule is an important parameter in photosynthesis, since it gives an indication, albeit an indirect one, of how much time is available for the energy of excitation to reach a chemically active molecule. The lifetime is defined as the time required for fluorescence to decay to $1/e$ of its maximum intensity (e is approximately 2.7). As we have mentioned, the lifetime of the higher singlet states is extremely brief, from 10^{-14} to 10^{-13} second; these states, however, do not enter into photochemistry. The first excited singlet state is much longer lived. Its lifetime was first measured by Seymour S. Brody at the University of Illinois and by Aleksander N. Terenin and his colleagues at Leningrad State University. New measurements in our laboratory of the lifetime of chlorophyll *a* in living cells, made with photosynthesis stopped by inhibitors or by low temperature, or made in light so bright



TRANSPORT OF ELECTRONS "uphill," against an electrochemical gradient, in the reaction center of the chloroplast takes place in three stages. On absorbing a photon (*a*), an electron in a pigment molecule is promoted to an excited state. The electron can then be transferred to the acceptor molecule (*b*); because the pigment is in an excited state this process is "downhill," that is, it is favored by the electrochemical gradient. In the final step (*c*), the pigment regains an electron from a donor molecule, another downhill process.

it saturates the photochemical system, yield a lifetime of about 2×10^{-9} second. Michael Seibert of the General Telephone and Electronics Research Laboratories and R. R. Alfano of the City College of the City University of New York have recently employed a laser and an instrument capable of marking extremely brief intervals in measuring fluorescence lifetimes in chloroplasts. At the resulting low average intensities they discovered two fluorescing species, one with a lifetime of less than 10×10^{-12} second, the other with a lifetime of about 300×10^{-12} second. From the emission spectra they have suggested that the two species are molecules of chlorophyll *a* in different environments, the more rapidly decaying one making up part of pigment system I, the other contributing to pigment system II.

The Transfer of Energy

Most of the pigment molecules in the chloroplast do not take part directly in the chemical processes of photosynthesis. Of the 300 molecules in the basic photosynthetic unit, almost all are thought to serve merely as "antenna" molecules. (The metaphor is not farfetched; after all, the molecules are tuned to receive signals of a particular wavelength.) The antenna pigments transfer the energy they absorb to a reaction center, which might consist of a single specially deployed molecule of chlorophyll *a*. This molecule then introduces the energy into the oxidation-reduction cycle.

If such a system is to operate efficiently, some method of rapidly transferring the energy of excitation from one molecule to another must be provided. That energy is indeed transferred is suggested by a simple observation: green leaves contain chlorophyll *b*, which is fluorescent in isolation, yet in living cells only the fluorescence of chlorophyll *a* can be detected. Evidently the accessory pigments transfer their energy to chlorophyll *a* so efficiently that their own fluorescence is quenched. The effect can be demonstrated directly in monomolecular layers of mixed chlorophyll *a* and *b*; even when the mixture is illuminated at a wavelength that will stimulate only chlorophyll *b*, only chlorophyll *a* fluorescence is observed.

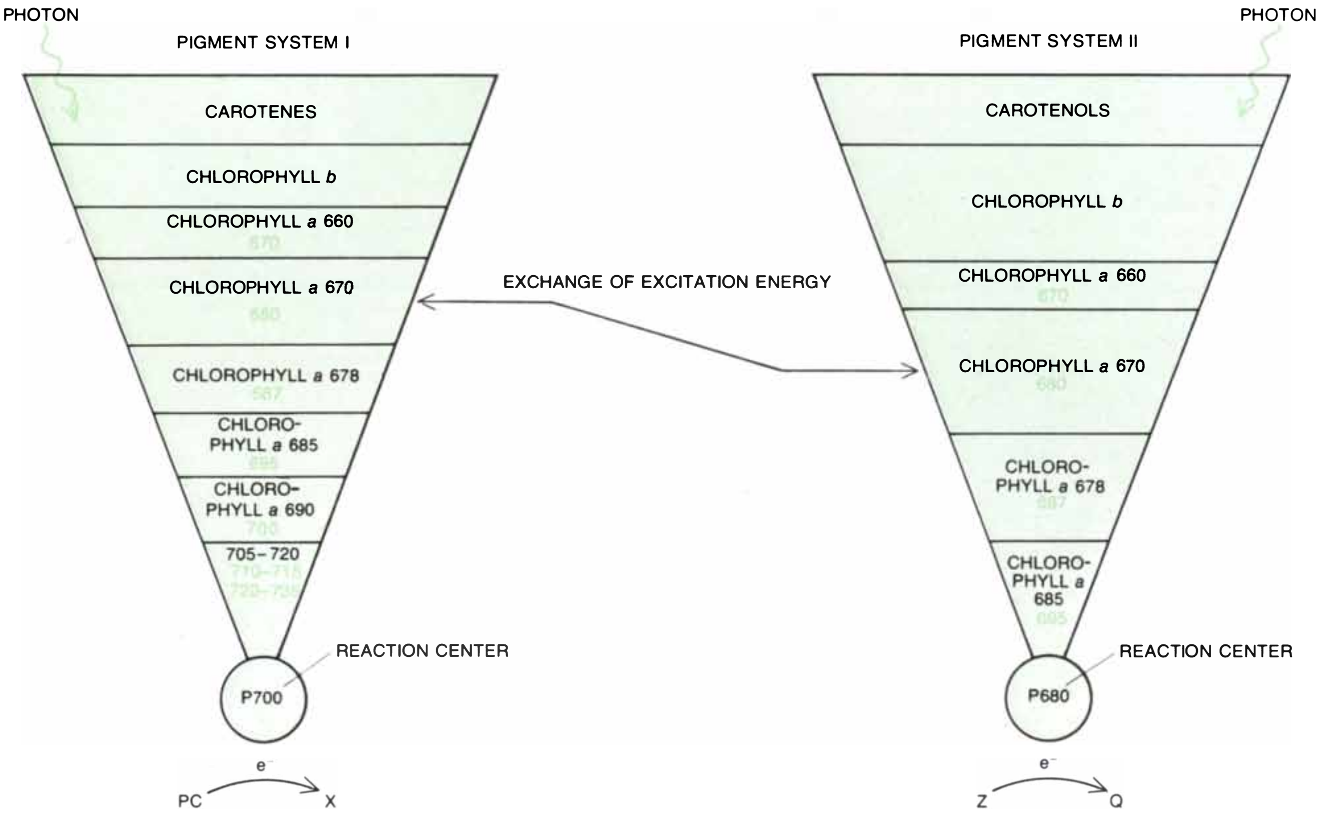
Among the accessory pigments a sequence of energy-carriers has been worked out. It begins with the carotenoids, then passes through phycoerythrin, phycocyanin and allophycocyanin (another phycobilin) to chlorophyll. In plants that have no phycobilins, energy can be transferred directly from the carotenoids to the chlorophylls, but when the phycobilins are present they are invariably utilized in sequence. Each transfer from one kind of pigment to another involves a modest loss of energy as fluorescence or heat [*see bottom illustration on page* 77].

Once the excitation energy has been passed to a chlorophyll *a* molecule it can migrate through the entire population of chlorophyll *a* molecules without loss. That the migration does take place was first shown by Arnold and E. S. Meek at Oak Ridge in 1956. Arnold and Meek illuminated chloroplasts with plane-polarized light and observed the resulting fluorescence. If the exciting energy had been retained by each molecule during its fluorescence lifetime, the polarization would have been maintained in the fluorescent light; it was in fact destroyed entirely, indicating extensive energy migration.

One possible explanation of the mechanism underlying this migration is called the Förster resonance theory, formulated in 1948 by the late Theodor Förster of the Technical University in Stuttgart. The theory holds that a molecule must first subside to the lowest vibrational sublevel of the first excited singlet state before its energy can be transferred; once this requirement is met the molecule is coupled to its neighbors by a resonance fundamentally dependent on their distance and orientation and on the overlap between the fluorescence spectrum of the donor and the absorption spectrum of the acceptor.

Faster mechanisms, in which the molecule need not first relax to its lowest vibrational state, have been proposed. They cannot yet be excluded, but calculations based on measurements of fluorescence lifetime support Förster's theory. An upper limit to the time required for energy transfer has been provided by an experiment devised by Peter M. Rentzepis of Bell Laboratories. Working with bacterial photosynthetic systems that had been stripped of all antenna pigments, so that only the reaction centers remained, he observed an energy exchange in $6 \pm 2 \times 10^{-12}$ second.

The path taken by the migrating excitation energy in the antenna pigments is thought to resemble a "random walk"



DISTRIBUTION OF PIGMENTS in the higher plants is different in the two pigment systems. Among the carotenoids, carotenes dominate in system I and carotenols in system II. Chlorophyll *b* is more abundant in pigment system II. Chlorophyll *a* is the most important pigment in both systems, but each has a characteristic assortment of forms of this pigment, distinguished by their absorption spectra (*black*) and fluorescence spectra (*color*). The reaction center pigments of the two systems also exhibit different spectra.

When You Pay Over $12 For Mumm Champagne, You Should Know Why.

Some sparkling wines call themselves Champagne. But true Champagne comes only from the province of Champagne in northeastern France. And from here comes Mumm, the finest Champagne the world has ever known.

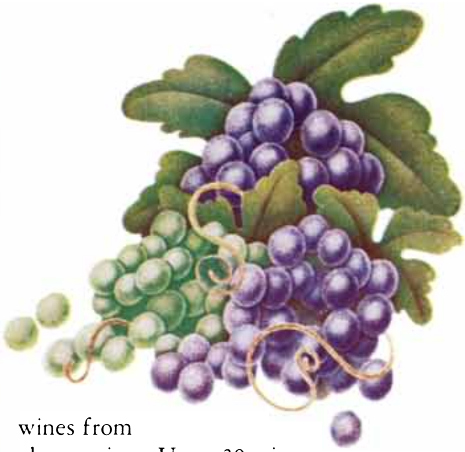
What sets Mumm apart? Very simply, its exceptional quality and unique *style,* or taste. For 150 years, the distinctive taste of Mumm has been consistently achieved. Here's how it's done.

HARVEST

Every single grape that goes into Mumm Champagne is inspected by hand. Skilled sorters mercilessly reject grapes which are imperfect.

THE CUVÉE (Blend)

After the grapes are pressed and the juice has fermented, expert winetasters note the qualities of the best



wines from the province. Up to 30 wines are then precisely blended to achieve the one perfect harmony of taste.

FERMENTATION

The second fermentation must take place very slowly *in the bottle* to produce the light, continuous bubbles of great Champagne.

As the wine approaches maturity, after years of aging, the bottles are turned daily and tilted downward to collect the tiny yeast particles, a natural by-product.

DÉGORGEMENT

A few months before shipping, a bit of the wine in the neck of each bottle is frozen and an expert quickly removes the cork and the yeast which

has solidified next to it. Without losing any of the sparkle, the bottles are recorked and prepared for export throughout the world.

Today you can still enjoy the finest Champagne the world has ever known. Ask for Mumm Extra Dry, Cordon Rouge Brut and incomparable René Lalou at fine wine merchants and better restaurants.

Mumm's The Word

IMPORTED BY BROWNE VINTNERS NEW YORK, N.Y., AND SAN FRANCISCO, CA
TO LEARN MORE ABOUT OUR FINE CHAMPAGNES, WRITE BROWNE VINTNERS, 505 BEACH STREET, SAN FRANCISCO, CA 94133

that continues until the energy quantum happens to encounter an available reaction center, or until it is reradiated by fluorescence. The reaction centers themselves might be dispersed randomly in a monolithic bed of bulk pigment, so that any quantum of energy could wander throughout the mass and in principle could eventually reach any of the reaction centers. Alternatively, each reaction center might be surrounded by, and served exclusively by, a fixed aggregate of pigment molecules. These two hypotheses are called the "lake model" and the "isolated puddle model." The actual situation appears to be a hybrid of the two, calling for a "connected puddle model" [*see illustrations on page 78*].

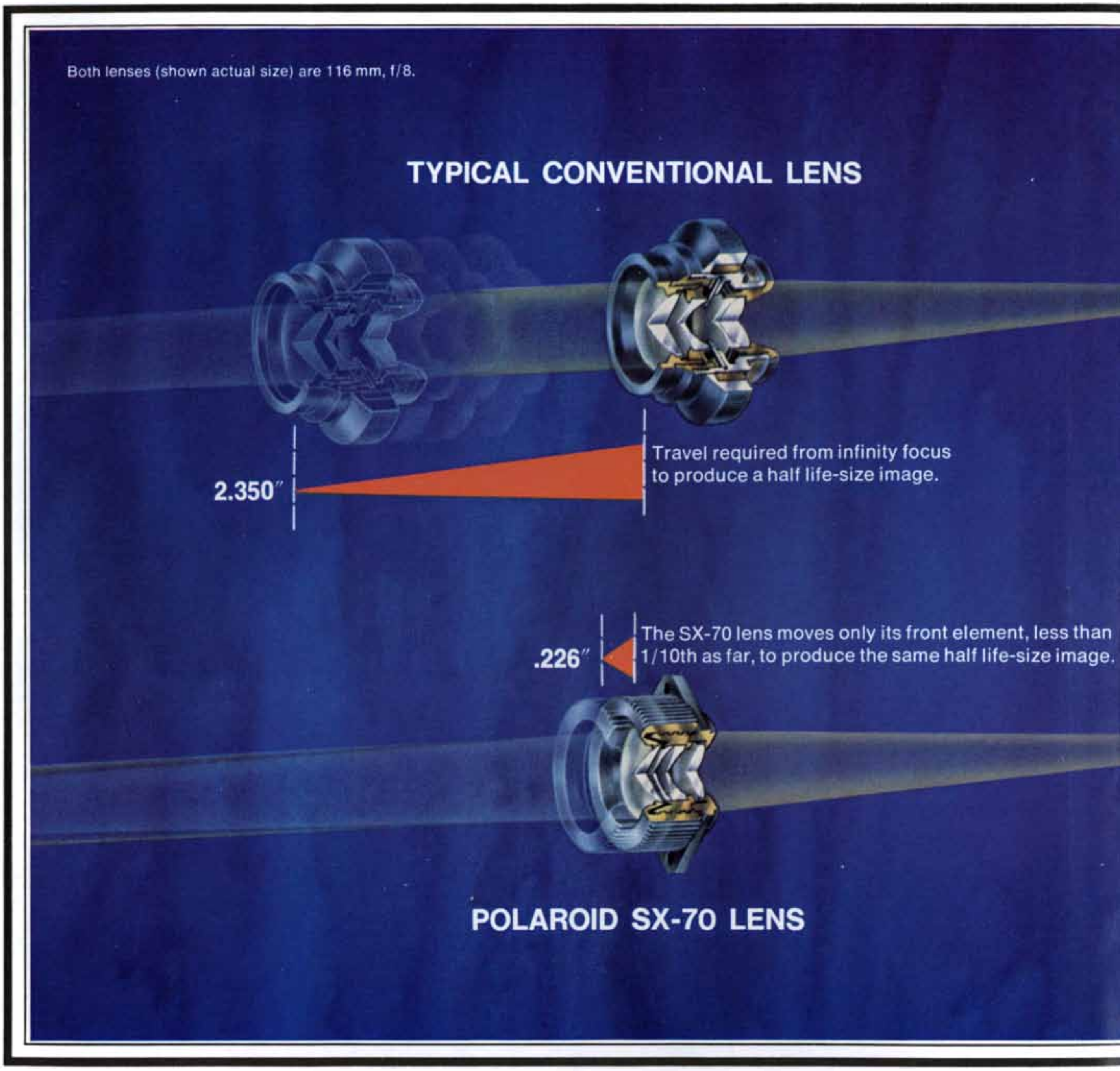
The Reaction Center

The elaborate system of pigmentation in plant cells is intended solely to convey the energy of light quanta to reaction centers where the central events of photosynthesis take place. Learning the nature and the exact business of the reaction center has been an intractable problem in the study of plant metabolism, but some of the molecules active there have now been at least labeled spectrally if not identified chemically. Louis N. M. Duysens of the University of Leiden discovered the first of these molecules when he observed a light-induced change in the absorption spectrum of photosynthetic bacteria at 890 nanometers; the change was later shown to be produced by bacteriochlorophyll in the reaction center. Kok identified a similar change in the absorption of green plants at 700 nanometers; this light-induced absorption change is now known to be caused by the reaction-center molecule of pigment system I, designated P700. Exposure to light has been demonstrated to oxidize P700. The analogous molecule for pigment system II was discovered by G. Doring and Horst T. Witt of the Technical University in West Berlin and is labeled P680 or P690; whether or not it is a direct party to an oxidation-reduction process has not been definitely determined.

These specialized pigment molecules appear to stand at the culmination of the physical, energy-gathering stage of photosynthesis, and at the beginning of the chemical, energy-storing phase. There are a number of ways of determining whether or not they do mark that intersection. The chemical processes they participate in should be the first to take place after a brief flash of light, and they should be in the main pathway of photosynthesis. From experiments on the inhibition of photosynthesis by low temperature we know that the reactions should also continue even at very low temperatures (below 77 degrees Kelvin, or 77 degrees Celsius above absolute zero). These tests have been made and the evidence that has accumulated is encouraging. P700 is in fact oxidized by light at 77 degrees Kelvin (although the pigment cannot be fully reduced at that temperature). The absorption change associated with P680 or P690 has been observed at low temperatures. A series of calculations, beginning with early measurements made by Kok, show that the oxidation of P700 is a highly efficient reaction, and therefore it must be part of the main pathway. (If it was not, it would compete successfully with the main course of events.)

Finally, if P680 and P700 are the active molecules of the reaction center, they must be capable of initiating the sequence of events that leads to the transfer of electrons. They could do so, when they are in the photoexcited state, by donating one of their own electrons to the acceptor molecule, then recovering it when in the ground state from the donor [*see illustration on page 79*]. In the case of pigment system II, for example, P680, after absorbing a photon, would transfer one of its valence electrons to the molecule labeled *Q* and in the process would return to the ground state. It could then accept an electron from the donor molecule Z. This hypothesis implies that the reaction-center pigments bear a positive charge during the electron transfer, and the presence of chlorophyll ions in the chloroplast has indeed been demonstrated by electron-spin-resonance techniques that monitor unpaired electrons.

If this sequence of events is correct, some means must be provided to prevent Z and Q from reacting with each other to annul the charges developed on them. A direct reaction between them is thermodynamically the most favorable of all reactions, and it is known to take place, although only to a minor extent. (It is responsible for the delayed light emission discovered in 1951 by Arnold and Bernard L. Strehler at Oak Ridge.) The back-reaction between Z and Q must not be allowed to dissipate a significant amount of the absorbed energy, however, if photosynthesis is to achieve useful results. What mechanism intervenes to separate them is for now entirely unknown. That is only one of several mysteries, however, that remain to be solved by further study of photosynthesis.



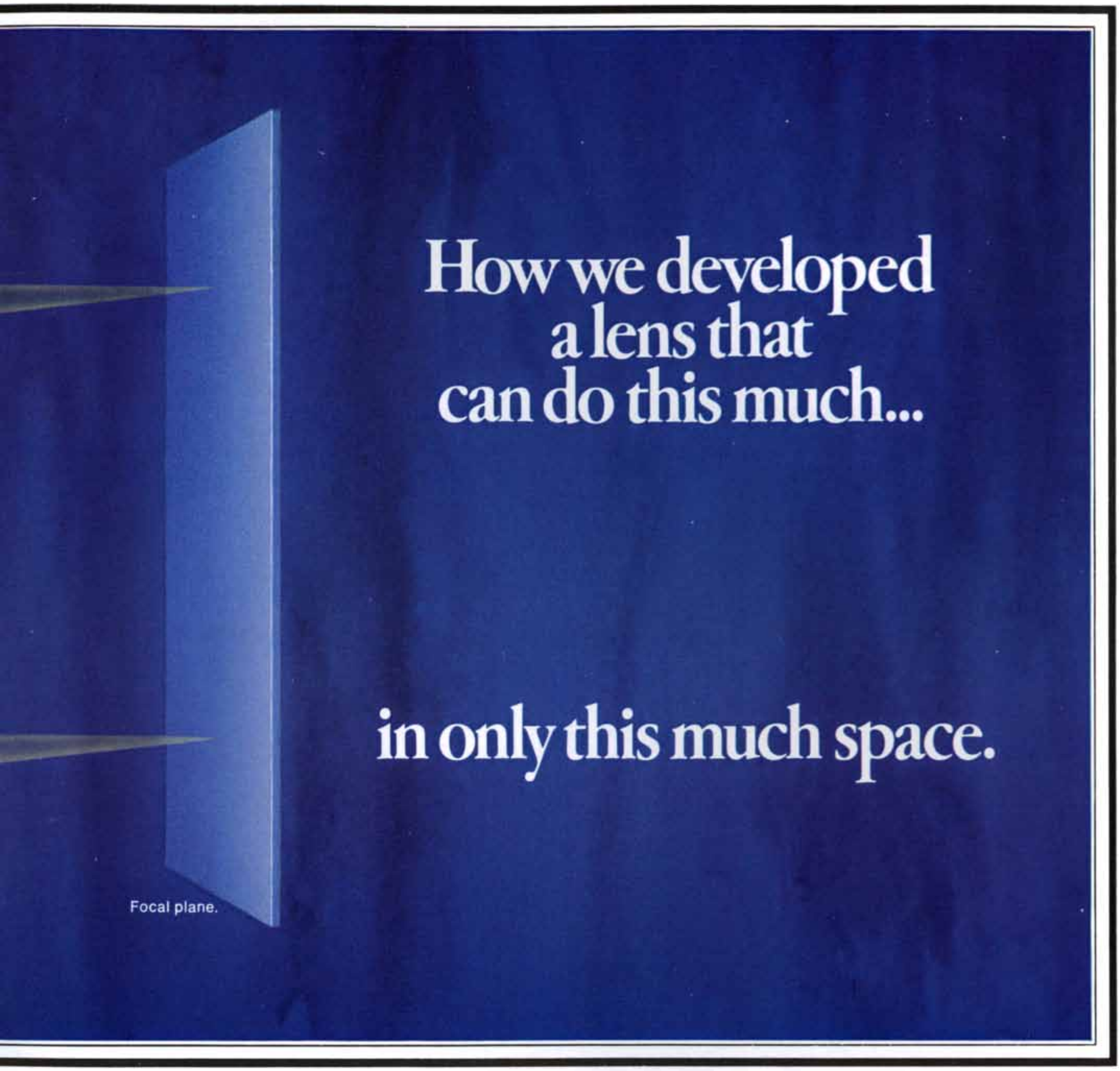
The four-element glass photographic objective of the SX-70 has qualities never before seen in a camera lens. It's remarkably compact (1 inch in diameter, and even less than that in thickness); and, with less than ¼ inch movement, it focuses on subjects as close as 10.4 inches away for remarkable close-ups.

To appreciate this accomplishment, consider a conventional lens like the Tessar design shown above. To produce the same close-up picture that the SX-70 Land camera can make (about ½ life-size), the entire Tessar lens (shown above) would have to move forward 2.350 inches from its infinity-focus position. But the SX-70 lens can make the same picture by moving only its front element a mere .226 inch. (If the Tessar lens were restricted to this .226 inch travel, it would focus no closer than 7 feet.)

Furthermore, since just the front element of the SX-70 lens moves during focusing, the camera does not require the exposure adjustment of 1¼ f-stops (the "bellows factor") which is needed when the entire lens moves away from the film to take such a picture.

When we started on this project, our optical team knew only that the SX-70 would require a lens of unprecedented compactness with about a 5 inch focal length and as extensive a focus-

OID SX-70 CAMERA LENS.



ing range as possible.

During the early stages of the design, we considered three and four-element systems, flirted with aspherics, and even explored approaches in which all of the lens elements moved in relation to each other as the camera was focused.

As we entered into the project, we found that normal computer optical design routines were inadequate. Our stringent space limitations didn't permit the significant variations in element thickness or air spaces that conventional routines require. Furthermore, we wished to optimize the system's performance over its extremely large focusing range, a feat which proved to be difficult at best with existing programs. Ultimately, we devised a whole new computer optimization routine appropriate to the extraordinary design task we had undertaken.

Our lens designers spent well over 1,000 hours using these powerful new computer techniques, exploring in detail some 30 different lens designs, most of which would have performed adequately in all but the most critical situations. Finally we chose a four-element design with front-element focusing whose deceptively simple form belies its unusual performance capability.

The design of the mounting

Continued on next page.

cell for the lens elements was also a formidable task. Although we were quite successful in minimizing the sensitivity of the lens elements to mounting variations, our constant search for compactness, coupled with the unusual focusing range, placed uncommonly stringent demands on the precision of the front element motion.

We discovered we could provide this controlled motion by replacing conventional "V" threads with a special four-start square thread.

This led to a simple 3-piece mounting cell suitable for economical, high volume manufacture. Nevertheless, several of its dimensions are held to tolerances measured in tenths of thousandths of an inch (0.0001″) to assure that the lens system will reach its ultimate performance capability.

Having found an elegant solution to the lens design problem, we turned our attention to assuring that each lens would achieve the outstanding performance specifications that our design permitted. We wanted to avoid some well-known shortcomings of conventional photographic resolution tests.

Look closely at the enlarged resolution target (Figure A). While the right-hand target is considerably "sharper," the resolving power (determined by the smallest bar set in which the bars and spaces can be defined) is higher on the *left*. High resolution does not always produce sharp pictures.

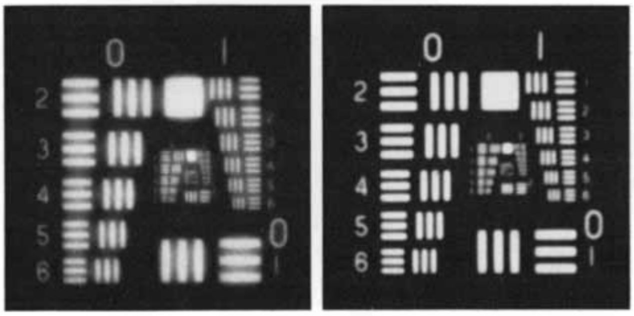


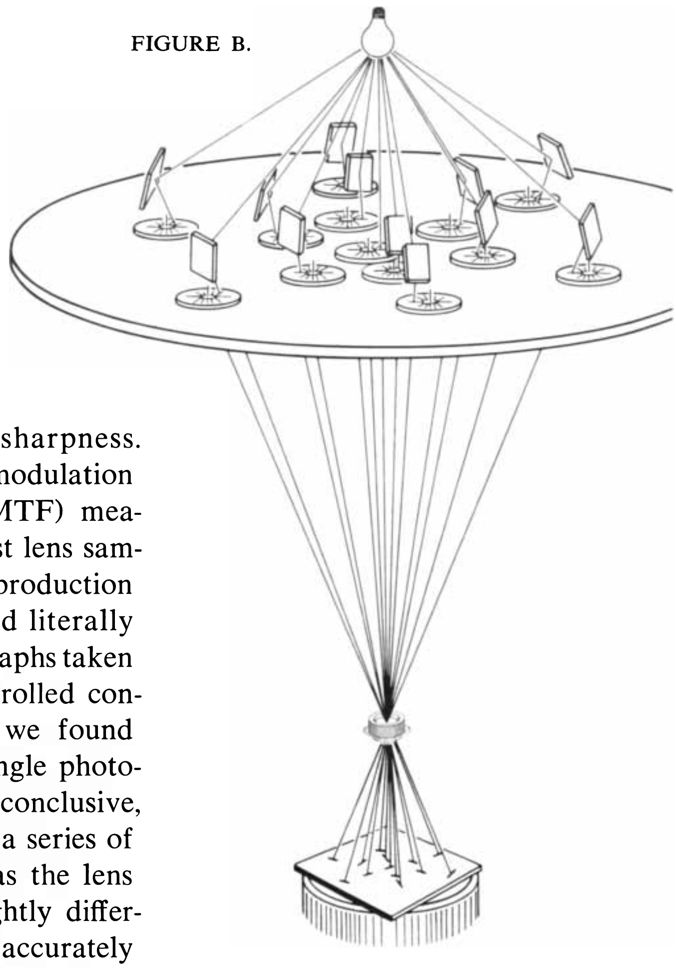
FIGURE A.

In a search for the best measure of the optical performance of this unusual lens system, we considered the role of the human eye in perceiving picture sharpness. We made extensive modulation transfer function (MTF) measurements on the first lens samples obtained from production tooling. We analyzed literally thousands of photographs taken under carefully controlled conditions. And while we found that a study of a single photograph was often inconclusive, comparisons among a series of photographs taken as the lens was focused for slightly different object distances, accurately predicted performance in a wide variety of conditions.

This so-called "through-focus series" required an examination of many portions of the image plane as the lens was focused through optimum. Each series took hours of an experienced photo technician's time. And we knew that, if we were to test each lens to maintain our high standards, we must perform this analysis in a matter of 3 seconds. Clearly, we needed to automate the image analysis process.

To replace the photographer, camera and film, we turned to photo-detectors, electronic signal processing, servomechanisms and computers. An exhaustive study correlating photographs and MTF measurements of representative lenses convinced us that we could derive the required information from accurate single-frequency contrast transfer measurements. To do that we needed a way to keep track of the variation of contrast transfer at many locations in the field as the lens was focused.

FIGURE B.



Our solution uses conventional rotating light choppers and an electronic technique known as "multiplexing." Thirteen rotating slotted wheels are illuminated by a single light bulb (Figure B). The lens being tested focuses the slots of each wheel onto one of 13 slits in its image plane. All of the light passing through the slits is collected onto a single photodetector.

By spinning each wheel at a different speed, the light passing through each of the 13 channels is modulated at a different audio frequency. As the lens is focused, the rise and fall in intensity of each tone is proportional to the contrast transfer at a particular slit location. By applying simple computer logic to these signals, we are, in fact, able to duplicate the results of hours of our through-focus picture analysis in the required 3 seconds.

We perform this complex full-field image quality test and *also* check the scatter and color balance contributions of over 1,000 lenses per hour on automated testing equipment. Every SX-70 lens is produced for us by Corning Glass Works. And every lens must pass this test.

You have seen how our quest for an extraordinary lens of uncompromising quality and versatility led us to new levels of sophistication in optical design, manufacturing and testing. This endeavor has produced a lens capable of doing what no camera lens has ever been asked to do before.

But then, the SX-70 system in its entirety does what no photographic system has ever done before. It does all the things *you* had to do before. And it does them so unobtrusively that, as you watch your photograph materialize before your eyes, you may wonder, briefly, how it got there.

The SX-70 System from Polaroid

THE SOLIDIFICATION OF CASTINGS

Many properties of a cast metal part arise from events that take place as the molten metal solidifies in the mold. Advances in understanding of these events make it possible to control them

by Merton C. Flemings

In modern technology metal is formed in many ways, notably machining, forging, stamping, rolling, extruding and casting. In the broadest sense they all begin with casting, which is defined as shaping metal by first melting it and then allowing it to solidify in a mold. Usually the term casting is applied to objects that are given what is essentially their final shape in this way. For the other methods of metal forming a casting is the primary simple shape, such as an ingot, on which the final form is imposed by one or more of the other methods.

What happens as the metal solidifies in casting is crucial, because many of the properties of the finished product arise from the events of solidification. Until fairly recently, when modern techniques of research and inspection were brought to bear on the subject, exactly how a melt solidified was a mystery. Casting was more of an art than a science, and foundry operators tended to be secretive about the techniques they had found successful. Some features of the solidification process remain poorly understood today. However, the interaction between theory and practice in the field of solidification has been strong in recent years, to the great benefit of both.

Solidification, which is a freezing process just as the transformation of water into ice is, begins at the wall of the mold and proceeds inward. At any given moment during solidification one would usually find a zone of solid metal, a zone of liquid metal and between them a zone where the liquid was being transformed into the solid. In the industry this latter region is termed the mushy zone.

Only when the metal is extremely pure or when special control is exercised over solidification does one find a smooth interface between the solid zone and the liquid zone. If impurities or alloying substances are present, even in small amounts, they tend to be rejected by the metal solidifying at the interface. They then lower the melting point of the liquid next to the growing solid and cause the growing interface to become unstable. As a result the interface becomes jagged. Tiny crystalline protrusions, often shaped somewhat like pine trees, extend into the liquid; they are called dendrites, from the Greek for tree.

In their commonest form dendrites have similar dimensions in all directions. They are therefore "equiaxed," meaning that the distance along each crystal axis is approximately equal to that along the others. In a mold where the thermal gradient from outside to inside is steep many of the dendrites grow in a direction approximately perpendicular to the mold wall; they are known as columnar dendrites [*see illustration on page 90*].

A look into some molds during solidification would not show a solid, a mushy and a liquid zone simultaneously. An example is a nonferrous alloy that freezes over a relatively wide range of temperatures when it is cast in a mold such as sand, which has good insulating properties. In this case thermal gradient within the metal from outside to inside is not steep. Here the fully liquid zone disappears quite early in solidification and the fully solid zone does not appear until solidification is almost complete. During most of the process liquid and solid coexist throughout the casting.

It is well to remember that the solution of important practical problems in the field of solidification has called for conceptualizing on vastly different scales simultaneously, which is more difficult than one might suppose. Processes at the liquid-solid interface must be visualized at the level of angstrom units: 10-billionths of a meter. Inclusions, which are impurity compounds that form during solidification, are on the scale of micrometers: millionths of a meter. The spacing of the arms of the dendrites is usually some fraction of a millimeter. The crystal grains of the metal itself range in size from millimeters to centimeters. Castings and ingots can be on the scale of meters. One cannot understand structural features such as surface finish, shrinkage and segregation (the separation of alloying elements in the course of solidification) without at the same time understanding the processes that are taking place at the microscopic level.

The dendrites that form in the mushy zone in equiaxed dendritic solidification arise from either of two mechanisms. One is heterogeneous nucleation, in which the formation of crystals out of the liquid is initiated by foreign particles. The other is the multiplication of dendrites, a phenomenon that has only recently become known in metallurgy.

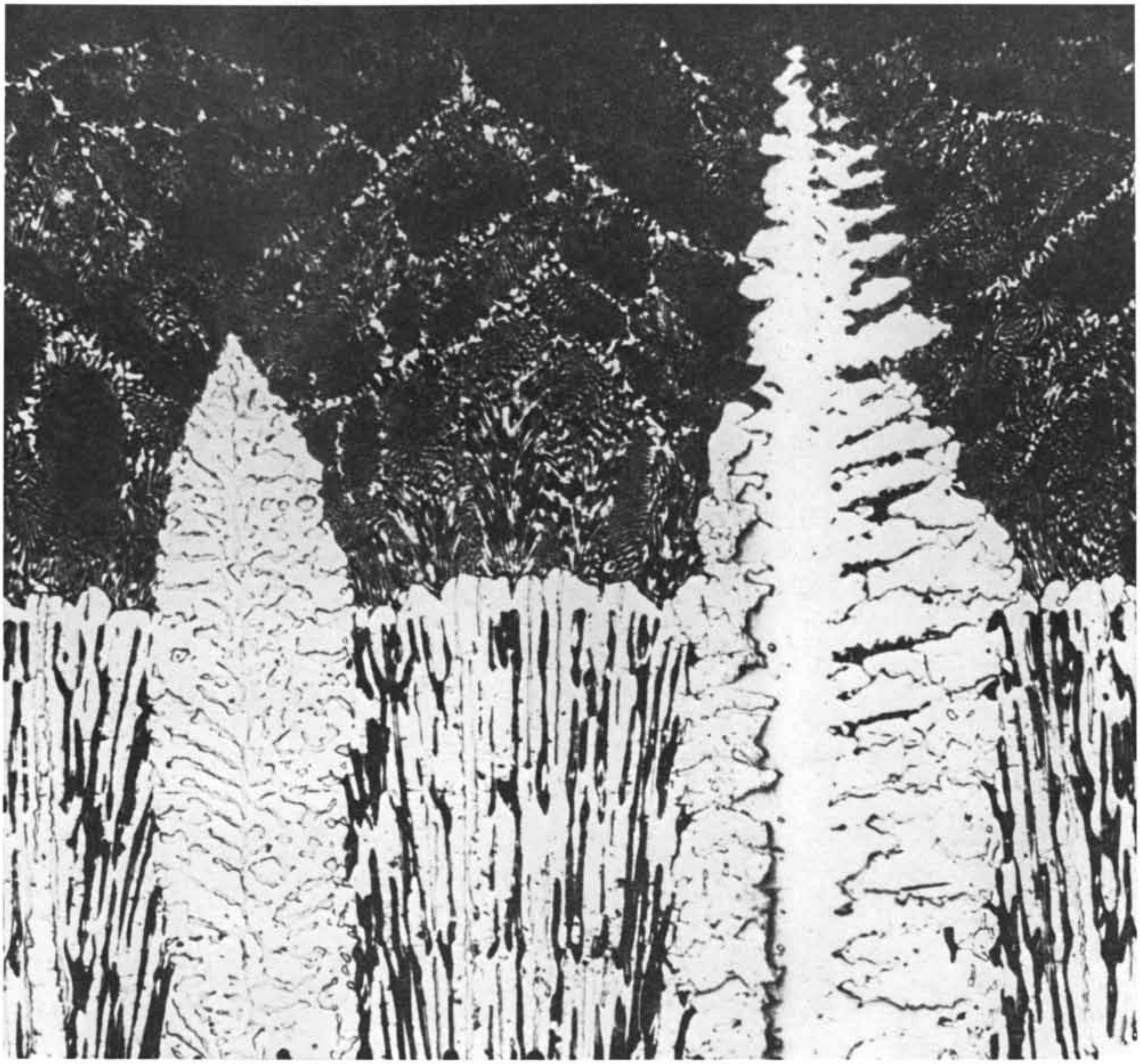
In equiaxed dendritic solidification the orientation of the dendrites is random. If the heterogeneous nuclei present are not too potent, however, and if dendrite multiplication is minimized (by minimizing convection or maintaining a steep temperature gradient or both), an aligned columnar dendritic growth is obtained. Except for the obvious difference in the alignment of the dendrites, the process of solidification is basically the same for both the randomly oriented and the columnar dendrites.

On the other hand, if the alloy is vigorously agitated during at least the first stages of solidification, dendrite multiplication occurs to such an extent that the usual dendritic structure is no longer obtained. Instead, solidification proceeds from spheroidal growth centers. The rheological (flow) properties of this structure are quite different from those of the dendritic structure.

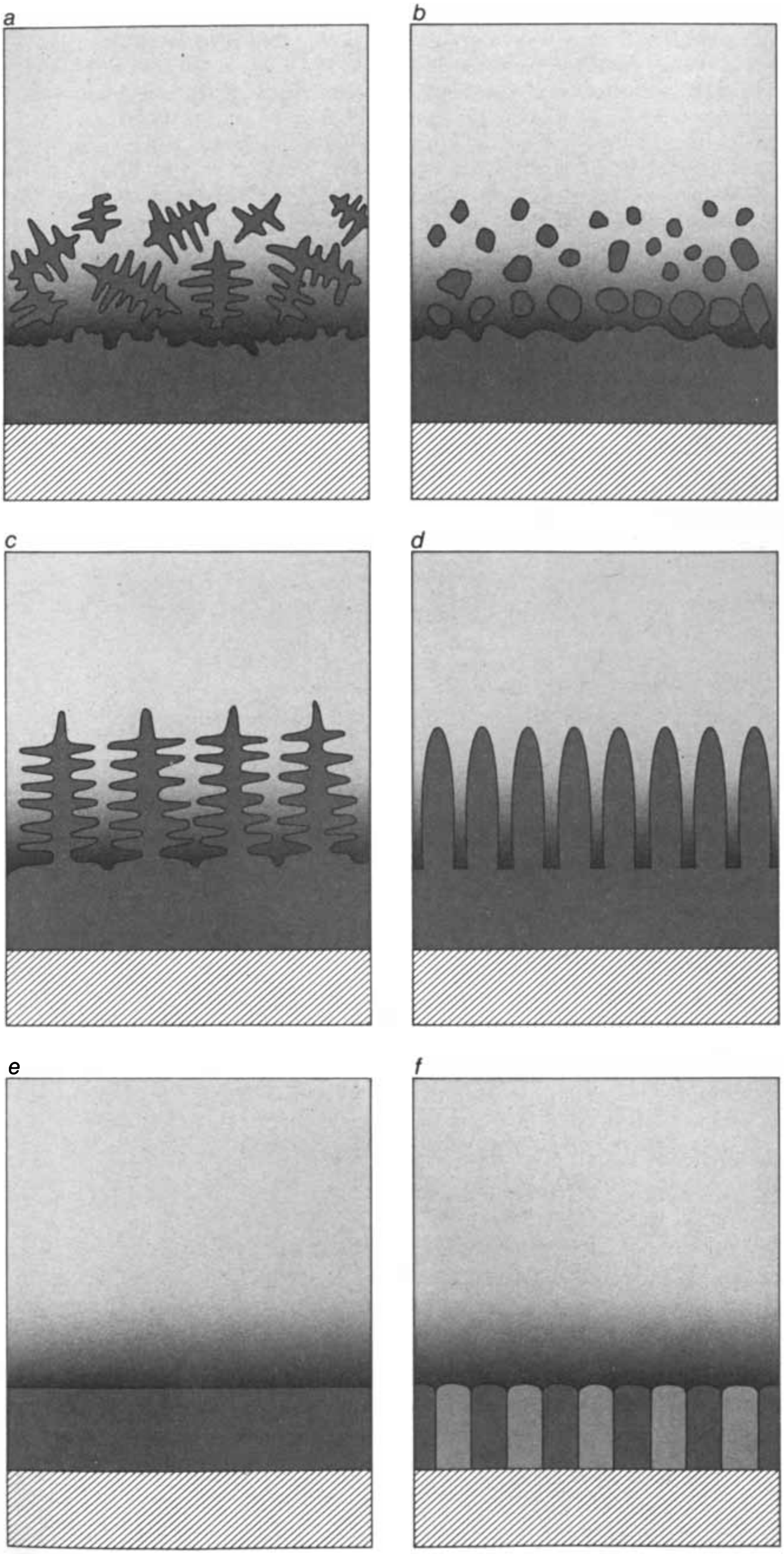
The three structures within a solidifying metal described so far—equiaxed dendritic, columnar dendritic and spheroidal—are readily obtained in normal castings and ingots at the relatively low values of G/R (thermal gradient divided by growth rate) that are customary in those processes. Other structures are obtained when G/R is increased by adding heat at one end of a solidifying alloy while extracting it at the other end. At intermediate values of G/R a cellular structure appears. It too can be either equiaxed or columnar. At high values of G/R one obtains either of two plane-front structures. One is a single-phase structure; it is obtained in conventional single-crystal growth. The other is a two-phase structure; it is obtained in the *in situ* composites of which I shall say more subsequently.

At the scale of the crystal grains in a solidifying metal one finds that each grain is composed of tens or tens of thousands of dendrite arms. All the arms of a grain, however, developed from the same initial growth point, and their crystallographic orientation is approximately the same. It is therefore customary to say that the grain consists of a single dendrite.

Going in the other direction on the size scale, that is, toward structures that require higher magnification to be seen, one next encounters inclusions [*see illustration at bottom right on page 91*]. They are of two types. One type, of which oxysulfides in steel are an example, is trapped by growing dendrites. The other type has a tendency to be pushed by the growing metallic solid into the spaces between the dendrites; examples are silica and alumina, oxides of silicon and aluminum. Inclusions of the second type can form before or during the solidification of the metal. At



SOLIDIFYING ALLOY was photographed at the interface between the liquid and solid phases. The alloy consisted of aluminum, copper and nickel. In this micrograph, in which the enlargement is about 320 diameters, the solid part of the alloy is at bottom and the liquid part, which is dark, is above it. The treelike solid structure at right is a primary dendrite, or grain, and the bladelike structure at left is a two-phase eutectic cell, meaning a structure of two different solid components that is freezing out of the liquid phase at a lower temperature than other components. The material on each side of these protruding structures is a three-phase composite.



FINE STRUCTURES OF SOLIDIFICATION are portrayed at a magnification roughly comparable to the one in the micrograph on the preceding page. In each case the mold wall is at bottom, the solid phase is above it and the liquid phase is at top. The structures are (*a*) equiaxed dendritic, meaning that the dimensions of each dendrite are similar in all directions, (*b*) equiaxed nondendritic, (*c*) columnar dendritic, (*d*) cellular, (*e*) plane-front, single phase and (*f*) plane-front, two phases. Darkening indicates enrichment of liquid by solute.

both stages they are free to move and therefore to collide with one another, forming clusters. Another feature found at this scale is microporosity, extremely small vacant spaces resulting from shrinkage or the evolution of gas during solidification.

At still higher magnification one reaches the atomic scale. One must turn to this scale to form a picture of the character of the liquid-solid interface. A simple model is a flat interface. On one side of the interface, according to this model, all the atoms are in the liquid state. On the other side they are in the solid state, and they lie in precisely the crystallographic positions dictated by the characteristic crystal habit of that substance.

Another model, devised by my colleague J. W. Cahn at the Massachusetts Institute of Technology, depicts a diffuse interface [*see illustration at bottom left on opposite page*]. Here the boundary between liquid and solid is not sharp or smooth. Some of the atoms on the "liquid" side of the line lie in crystallographic positions, as in a solid, and some of the atoms on the "solid" side of the line are not exactly in their proper crystallographic positions.

The second model is probably the more realistic representation of the interface in a solidifying metal. It must be said, however, that this area of solidification is notably lacking in good experimental work. A better understanding of the atomic structure at the interface is needed because it is the processes at this level that determine the preferred crystallographic growth directions of dendrites and whether or not regions called facets (flat, slow-growing crystal faces) appear on dendrites during solidification.

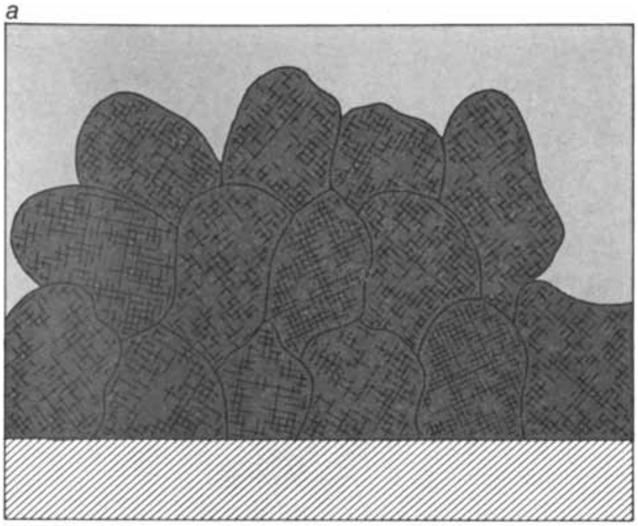
Against this background of general remarks about the structures and processes of solidification, I turn to examples of recent advances in solidification theory and practice and of how the theory and practice have interacted. The examples are presented more or less according to their significance and to the degree of their success in advancing knowledge of solidification. They therefore move up and down the scale of size.

The spacing of the arms of the dendrites in a given alloy is found to depend strongly and solely on the cooling rate. Moreover, it does not vary much from alloy to alloy. The importance of the spacing is that mechanical and other properties of a variety of alloy castings and materials wrought out of cast ingots depend strongly on the spacing. For example, by suitably chilling a sand-cast aluminum alloy to achieve a fine spacing of the dendrite arms it is possible to produce castings of premium quality, which have higher tensile strength and yield strength than ordinary castings. This principle has been employed by a segment of the aluminum-foundry industry for a number of years to produce such premium-quality castings for the aerospace industry. The principle stems from basic work on the relations between solidification and structure that has been done at M.I.T. for more than 20 years.

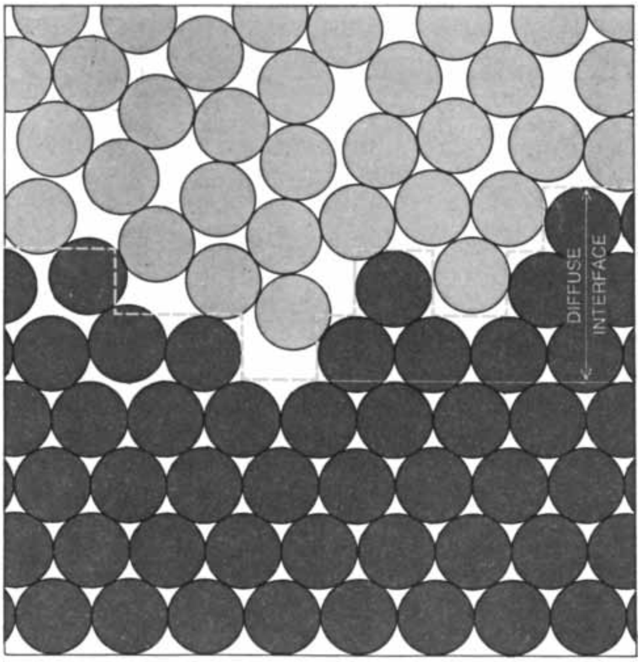
High-performance wrought alloys of high quality are now made by processes (or with process controls) that achieve fine spacing of dendrite arms. Examples include aluminum alloys of high strength and tool steels of exceptional quality. The properties of wrought materials with fine spacing of the dendrite arms are measurably and sometimes dramatically better than the properties of conventional material.

The key to fine spacing of dendrite arms is a high cooling rate. One way of achieving such a rate is to cast smaller ingots. Others are increasing water chilling in direct-chill casting, employing the electroslag remelting process and making billets by compacting atomized liquid droplets.

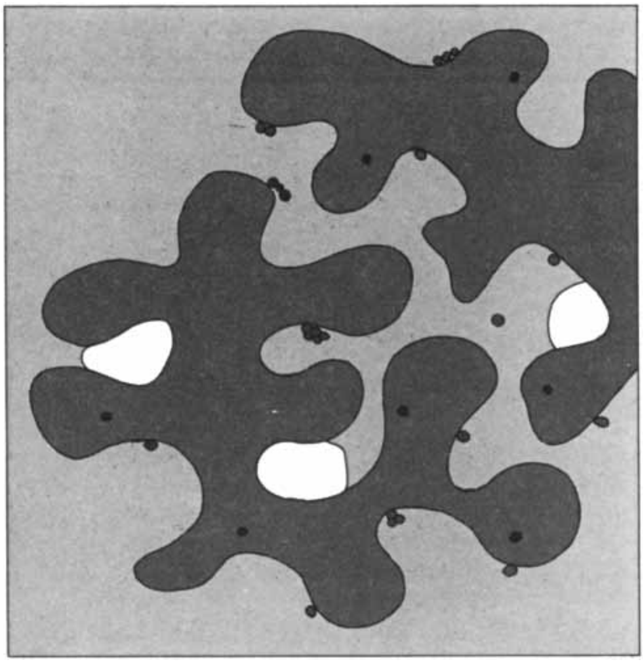
The spacing of the dendrite arms in



GRAIN STRUCTURE in a solidified casting or ingot is shown adjacent to the mold wall. The structures are (*a*) equiaxed dendritic and (*b*) columnar dendritic. Each grain consists of a large number of dendrite arms, perhaps as many as tens of thousands. They are portrayed schematically. The size of the grains, which are individual crystals in the solid, ranges from millimeters to centimeters.



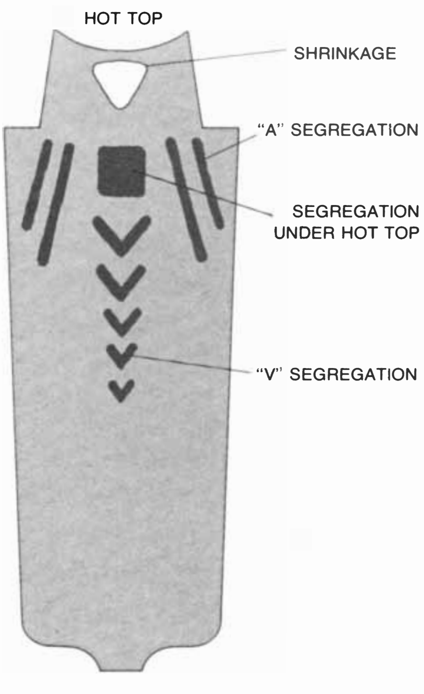
DIFFUSE INTERFACE between the liquid and the solid in a solidifying casting is depicted at the scale of atoms. The broken colored line roughly divides liquid and solid phases, but some of the atoms above it lie in the crystallographic positions of a solid and some below it are not precisely where they would be in a solid.



TINY INCLUSIONS, which are of micrometer size, often appear during solidification. They are particles of impurities. The black circles represent inclusions that were trapped by growing dendrites. Small gray circles represent inclusions that were pushed ahead of the advancing solid. Larger open spaces represent voids.

atomized droplets is very fine indeed. One alloy of current industrial interest made in this way is IN 100, a nickel-base superalloy containing cobalt, molybdenum, titanium, aluminum and other elements. In the cast form this alloy has an ultimate tensile strength of 140,000 pounds per square inch. It has a ductility of only 1 percent, that is, it can be stretched under test only 1 percent of its length before it breaks; accordingly it is difficult or impossible to forge. The compacted version has a tensile strength of 220,000 pounds per square inch, and with a ductility of up to 18 percent it can be readily worked. This material, with its fine dendrite structure and grain size, is ideal for applications at temperatures below about 1,400 degrees Fahrenheit (760 degrees Celsius).

Improvements to be gained by achieving a fine spacing of dendrite arms are not limited to metallic alloys. Marked improvements have been achieved recently in the performance of alumina-zirconia abrasives, notably by the Norton Company with its *ZF* and *ZS* abrasives. The abrasive grit in these products is melted and cast before being milled and incorporated into the grinding wheel. The improvement in performance, which is roughly double that of conventional abrasives, is achieved by controlling the casting process to obtain the fine spacing of the dendrite arms.



LARGE-SCALE FEATURES of solidification appear in a steel ingot weighing several tons. They include shrinkage in the hot top, which is a reservoir of molten metal designed to ensure complete filling of the mold. Below the hot top are various forms of segregation, which is a separation of alloying elements during solidification. They include a rectangular region rich in solutes, *A*-shaped segregations similar to the "freckles" described in the text and segregations of *V* shape. Segregations are defects that good casting practice tries to minimize.

Why is it that the spacing of dendrite arms is so strongly (and apparently so exclusively) dependent on the cooling rate? The practical importance of dendrite-arm spacing has led to considerable work on this fundamental question, partly in the hope of finding some other way to control spacing. The answer to the question has now largely been given, and it proves to be not at all what was expected.

It turns out that when a dendrite is growing in a melt, the arm spacing is at first very fine, even if solidification is proceeding slowly. As solidification at a given location continues, however, only some of the arms continue to grow. Others dissolve, so that during solidification the spacing of dendrite arms becomes progressively larger. Essentially, then, the final spacing of the arms is determined by a coarsening process. (This conclusion stems from work done by T. Z. Kattamis, now at the University of Connecticut, when he was a graduate student in my laboratory.) The less time the foundryman allows this process to continue, the finer the spacing will be.

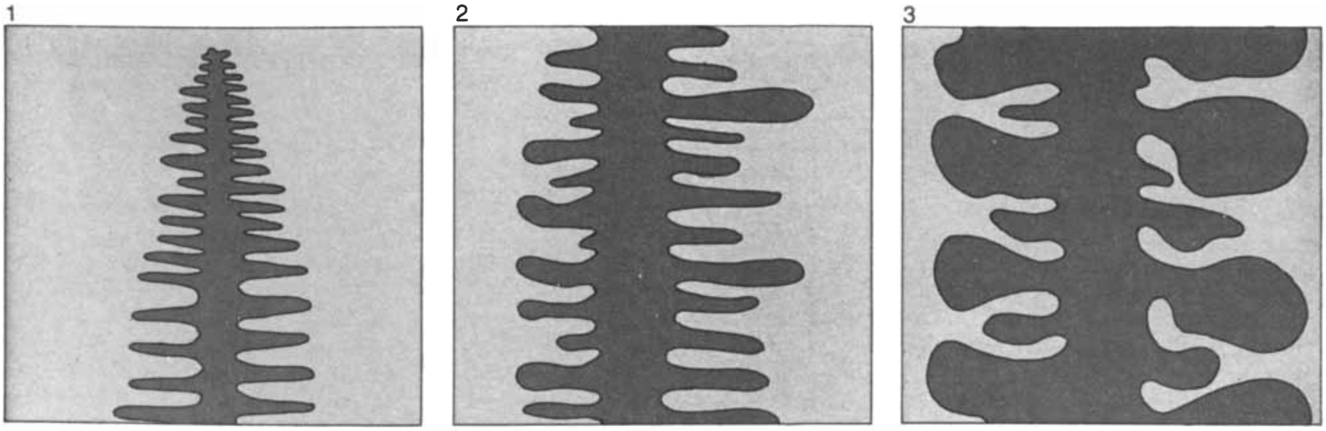
Castings with a fully columnar structure can be made by reducing convection to a low value and by being sure that effective heterogeneous nuclei are not present. These objectives are achieved by extracting heat unidirectionally through one face of the casting. In most metals the orientation of the crystals at the chill face is random, but the grains most favorably oriented for growth gradually crowd out their neighbors. The result is a columnar structure. Ferrous and superalloy castings produced in this way typically have better ductility at room temperature and better strength at high temperature than equiaxed structures. Therefore much effort has been devoted to making blades and vanes for gas turbines with this kind of structure.

By suitable manipulation, as by making the grain grow through a small hole or around several corners, it is possible to obtain a structure that consists of only one crystal grain. The structure therefore has a preferred orientation perpendicular as well as parallel to the axis of growth. The casting has no grain boundaries, although it does have a dendritic structure. These unidirectional structures have even better properties, such as resistance to creep, than the castings with columnar structure.

Columnar structures have other uses. One of them is to attain improved magnetic properties. In one alnico composition, for example, the maximum energy obtainable in a columnar structure is twice what can be obtained in the equiaxed alloy. Many alnico magnets are therefore made with at least a partly columnar structure.

Other articles in this magazine have described the value of composite materials, in which two materials are combined in such a way as to yield properties that are superior to (or radically different from) those of either material alone. Usually the two materials are combined mechanically, as when strong fibers are embedded in a matrix of epoxy resin. Composites can be made in one step by casting, provided that the value of G/R is raised high enough. At intermediate values, as I have mentioned, one achieves a cellular structure. In an alloy that has two phases, or components, after solidifying, casting at a still higher value of G/R causes the cells to become gradually shorter, and also smaller in cross section, until finally they can be pictured as being driven back into the fully solid interface. At this critical value one achieves the plane-front composite solidification that I touched on above. A structure made in this way is called an *in situ* composite. It has been known for many years that plane-front solidification could be obtained with single-phase alloys at sufficiently high G/R. Only through recent work in my laboratory by F. R. Mollard did it become understood that two-phase *in situ* composites could be similarly grown.

Studies on such low-temperature model systems as tin-lead and aluminum-copper-nickel alloys have been of great value in developing an understanding of the fundamentals of solidification in *in situ* composites and in defining the conditions of growth necessary to produce similar structures in high-temperature alloys, which are the focus of current practical interest. On an experimental basis H. Bibring and his colleagues of the French National Office of Aerospace Studies and Research have produced an *in-situ*-composite alloy of tantalum carbide rods in a cobalt matrix. This alloy and others have significantly better properties at high temperatures than conventional superalloys, and so a great effort is being made to develop practical processes to make gas-turbine blades and vanes having this *in-situ*-composite microstructure.

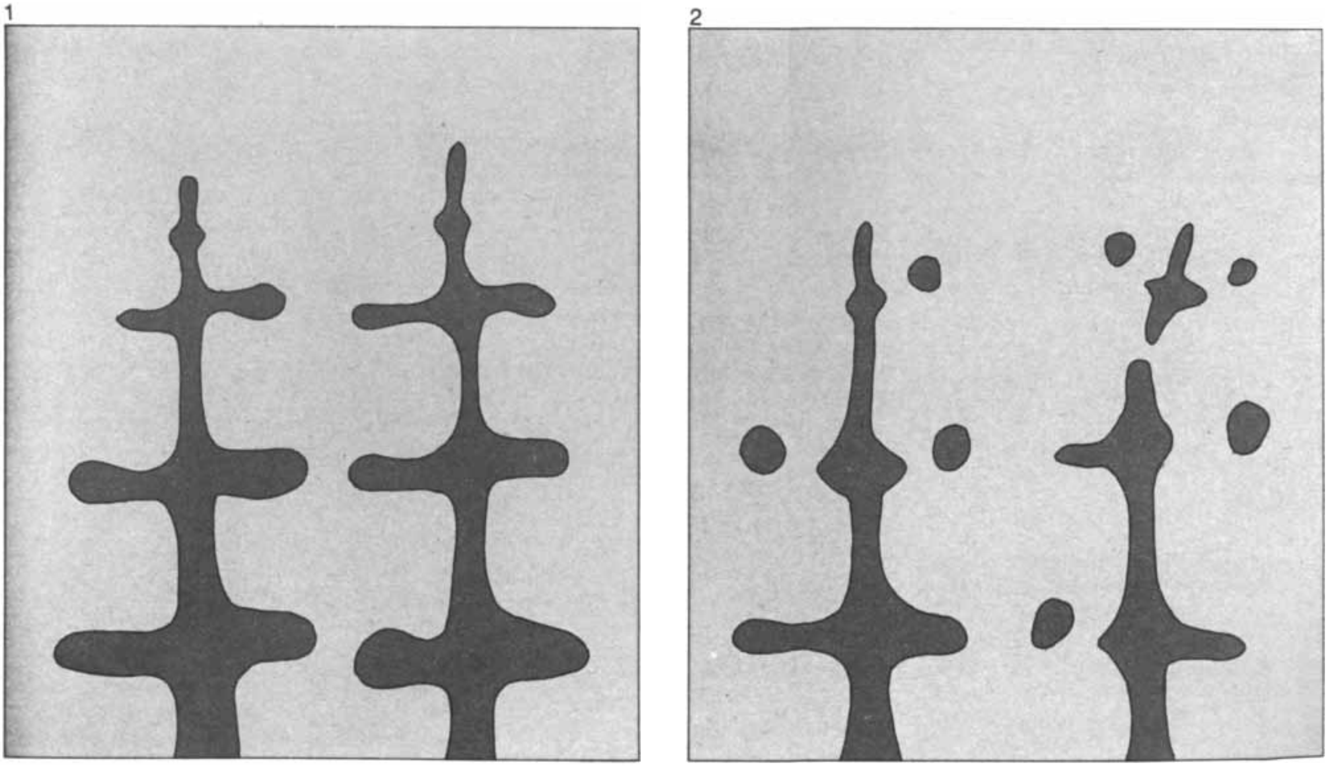


GROWTH OF DENDRITE in a melt normally entails a coarsening process. At first (*1*) the spacing of the arms of a dendrite is quite fine. As solidification proceeds within the mold (*2*), only a few of the arms continue to grow. Others melt, so that the spacing of the arms becomes progressively larger (*3*). A fine spacing of dendrite arms, resulting in a finished product with small grains, is desirable for castings of high quality. Therefore the sooner the foundryman can curb the process of coarsening, the better the casting will be.

The theoretical aspects of the growth of *in situ* composites and of structures obtained at intermediate values of G/R have been intensively studied in recent years. These investigations will no doubt continue. Many interesting and important areas for study remain, including how to get around the G/R restriction, how structure changes at angles and corners and how solidification occurs when one or more phase is faceted.

Turning now to a large-scale feature of solidification, the development of new ingot-making processes over the past decade or so has led to unforeseen forms of segregation in castings. As I have mentioned, segregation is a casting defect resulting from the separation of alloying elements in the course of solidification. The new forms of segregation include "freckles" in consumable-electrode ingots and inverse centerline segregation (a depletion of solutes, rather than the usual enrichment of them, near the center of the casting) in continuous castings [*see illustration on opposite page*].

The textbook explanations of segregation did not suffice for these new forms. As a result extensive fundamental work has been done on the problem, yielding entirely new explanations. These explanations, introduced originally by Robert Mehrabian and others while they



DENDRITE MULTIPLICATION is a recently discovered solidification event that opens a way to improved castings. As a dendrite grows (*1*), some of the arms melt off their roots and are carried elsewhere by convection (*2*). Convection also reduces the temperature in the bulk liquid, so that when a melted-off arm arrives in its new location it can grow into a new grain. Stirring of the melt promotes multiplication. Gentle stirring results in fine-grained equiaxed dendritic castings, vigorous stirring in nondendritic castings.

were students in my laboratory and now confirmed as being largely correct, have led in turn to significant advances in ingot-making processes that reduce or eliminate segregation.

A simple example is the kind of segregation often found under the "hot top" in large ingots. (A hot top is an insulated reservoir on the top of an ingot mold. The reservoir holds molten metal, which flows downward from the hot top to feed the contraction that occurs in the metal in the ingot during solidification.) According to conventional wisdom, the segregation results because this area freezes after the bulk of the ingot, so that the alloying element (such as carbon) is pushed there by the advancing solidification front.

The recent experimental work has shown that this type of segregation appears only when hot-topping practice is inadequate, resulting in a negative temperature gradient during solidification of the hot top. In this circumstance the structure in or just below the hot top has cooler, more solute-rich liquid in the interdendritic spaces above the warmer, solute-poor liquid. Convection currents are thereby set up in the interdendritic regions; they carry the solute-rich liquid downward and result in the solute-rich segregation observed. The problem is eliminated by improving the hot-topping practice.

On a scale finer than dendrite arms, inclusions constitute an important structural feature. They affect the surface finish of the ingot and the machinability and mechanical behavior of the final cast or wrought structure. In the past much attention has been paid to the thermodynamics of the formation of inclusions. Only recently have inclusions been seen as a problem of solidification.

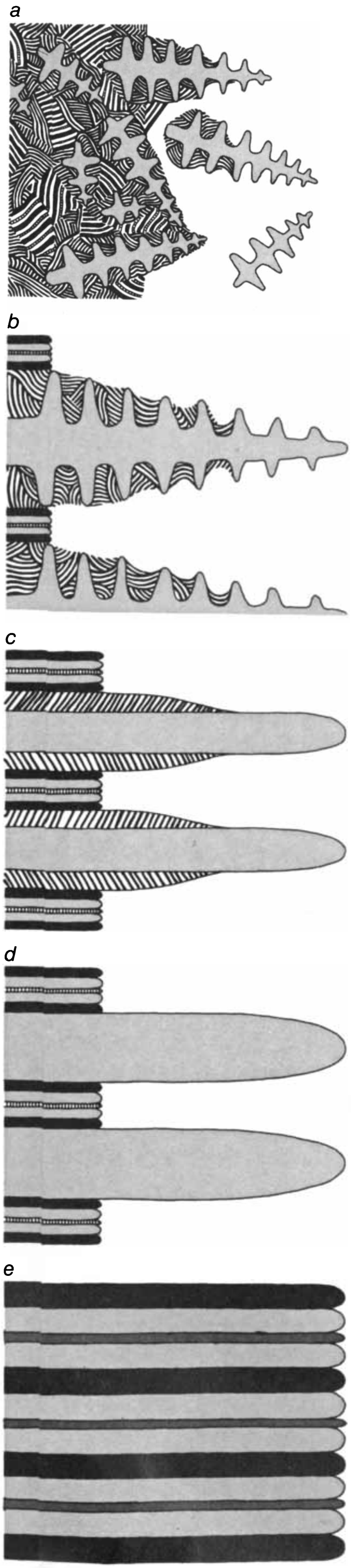
Inclusions can nucleate, multiply and grow dendritically just as the primary metallic phase does. They may grow before or during the growth of the metallic phase. They interact with growing metal phases in several ways. Nucleation of the metal sometimes seems to occur preferentially on inclusions. Sometimes inclusions are entrapped by the growing metal, as with oxysulfides in steel, and sometimes they are pushed into interdendritic spaces, as inclusions of silica and alumina tend to be.

A particularly interesting problem involving inclusions is isothermal solidification, which occurs (to take a single case) during the deoxidation of steel by aluminum. When aluminum is added to a steel melt that contains oxygen, the driving force for the isothermal solidification of aluminum oxide (Al_2O_3) is substantial. Solidification of this phase takes place rapidly, apparently limited only by convection and diffusion of the aluminum into the melt. A number of inclusion structures have been found in melts deoxidized by aluminum, including dendritic, faceted and clustered.

A possible but not yet proved explanation is that these three structures form sequentially. Fine dendrites form first because of the large driving force for the solidification of the aluminum oxide. Then the dendrites break up by the dendrite-multiplication process. Finally, the inclusions collide and partly coalesce. This coalesced structure is the cause of blockage in nozzles when metal is poured into a mold and also of poor surface finish in the casting of ingots, particularly continuous casting. Many interesting and unsolved problems in the solidification of inclusions seem ripe for attack.

Another area of advance is in understanding the dendrite multiplication that I have mentioned. Until recently it was thought that all new grains in castings and ingots arose from a nucleation event, with the nucleus being generated as described by the classical theories of homogeneous and heterogeneous nucleation. Now it is known that this view is incorrect. Indeed, it seems likely that in steel, at least, the great majority of grains (perhaps all but one) arise from a dendrite multiplication [*see bottom illustration on preceding page*].

A dendrite, once it forms, quickly becomes highly unstable because of its large surface area. That area can be reduced by the coarsening process that I discussed above. It results in a decline in the number of dendrite arms as solidification proceeds. The same process, however, can also cause arms to melt or dissolve off completely. The process is enhanced by convection, which brings thermal pulses to the tip of the dendrite, possibly straining it somewhat, and has



RANGE OF STRUCTURES can be obtained in directionally solidified, multicomponent alloys by varying G/R, which is the thermal gradient divided by the growth rate. The structures, which appear here in order of increasing G/R in a three-phase alloy, include one with equiaxed grains, a two-phase eutectic and a three-phase eutectic (*a*), aligned single-phase dendrites with two-phase eutectic cells and an aligned three-phase composite (*b*), aligned cells of each type of phase (*c*), aligned single-phase cells with an aligned three-phase composite (*d*) and aligned three-phase composite (*e*).

the additional effects of carrying the remelted dendrite arm to another place, where it can grow to a new grain, and of reducing the temperature gradient in the bulk liquid so that the melted-off arm is less likely to be dissolved while it is being carried away.

Had this multiplication mechanism and its effectiveness been known earlier, far less time and money would have been spent in trying to promote heterogeneous nucleation (and thus fine grain structure) by such techniques as vibrating the mold and applying ultrasonic waves to the solidifying melt. In fact, as one looks over the meager successes of experiments in grain refinement by these methods, it seems likely that even these refinements were attributable to grain multiplication, and that the same or better results could have been obtained by gently stirring the melt, as is now on occasion done commercially to achieve fine grain size.

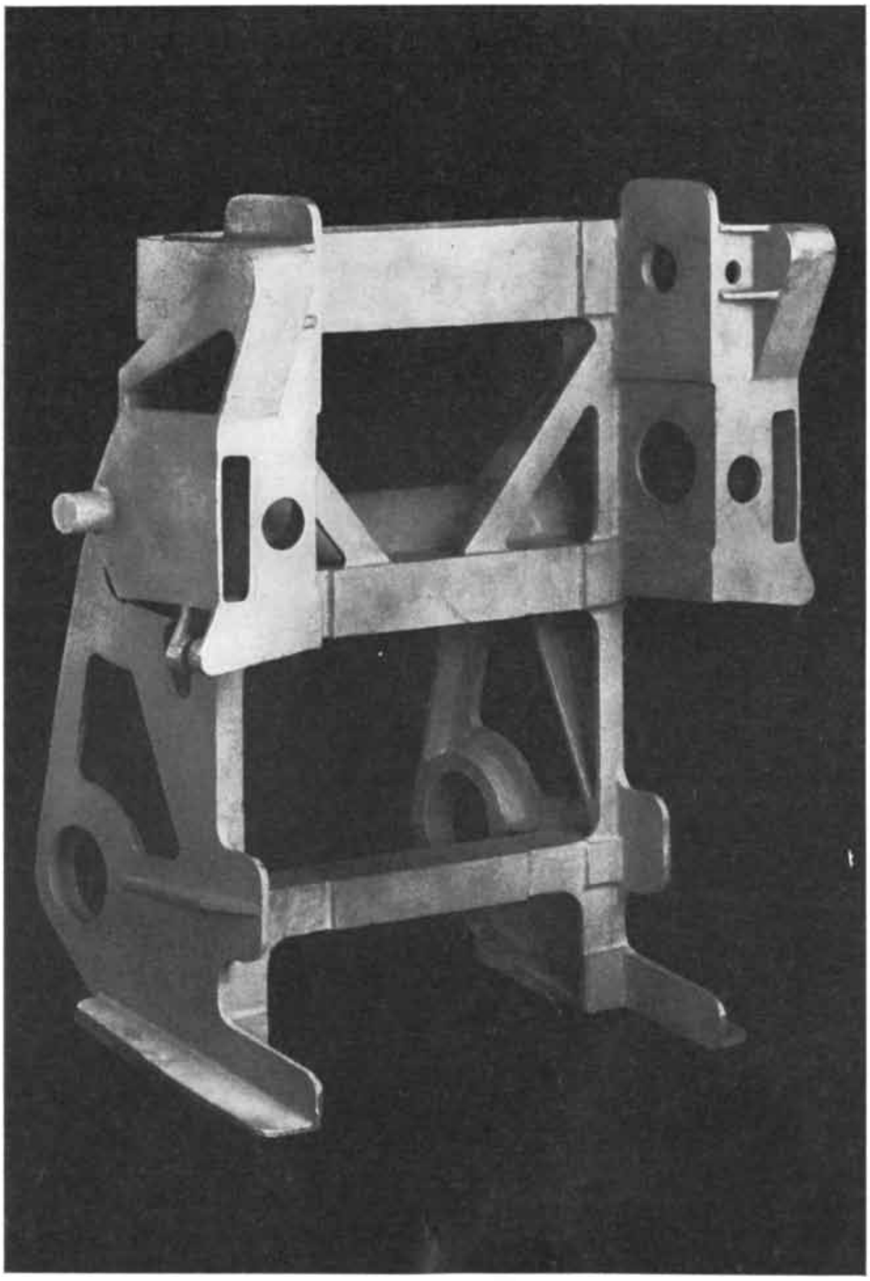
An interesting nondendritic structure is obtained when the melt is stirred with exceptional vigor. It is the structure identified as "equiaxed nondendritic" in the illustration on page 90. This structure and its unique rheological properties for metal were first observed by David Spencer during his graduate work at M.I.T. He partly solidified a tin-lead alloy between two counterrotating cylinders. The resulting structure of solidified small spheroids suspended in a liquid was found to flow as a slurry.

As a result it became evident that vigorously agitated, partly solidified slurries could be formed into shapes by conventional casting processes, such as die casting, or (if the fraction of solid material was higher) by forging or pressing. Extensive studies, both in my laboratory at M.I.T. and in industry, have demonstrated these possibilities of shaping slurries.

Until now all metal forming of commercial significance has been done in either the fully liquid or the fully solid state. The reason is that foundrymen could not run their castings satisfactorily when solid dendrites began to form and that solids could not be worked satisfactorily when liquid was present. The unique structure obtained by vigorous agitation during solidification appears to circumvent both of these difficulties and to lay the foundation for the development of improved or entirely new metal-forming processes, which promise to be more economical than the present ones.

One area of application of this process that is being explored is continuous casting, the aims being to speed up the process and thereby make it more economical and to achieve a finer, more easily workable structure. Another area of application being studied is die-casting metals of relatively high melting points, particularly ferrous metals. They are not die-cast in significant quantity now because the thermal shock to the metal components, including the mold, is so severe that the life of the components is short. The casting of a somewhat viscous, semisolid alloy of the slurry type can result in much reduced thermal shock to the surface of the die.

Another potential application for forming semisolid metals is the casting of composite materials. An example is 30-percent silicon carbide in an aluminum-alloy matrix. A wide variety of particulate and fibrous materials can be mixed readily into an agitated semisolid metal. They are retained as a dispersion because the solid in the mixture apparently prevents them from floating or agglomerating. Parts made in this way could be expected to have improved resistance to abrasion. Low-friction applications are also possible. The method may also provide a practical way of lowering the cost and weight of die castings that do not need to exhibit high strength. One can envision diluting the metal in the melt with a material such as recycled glass.



ALUMINUM CASTING of premium quality was made as an airplane part. A fine spacing of its dendrite arms, achieved by appropriate chilling in the mold, gives it a high strength.

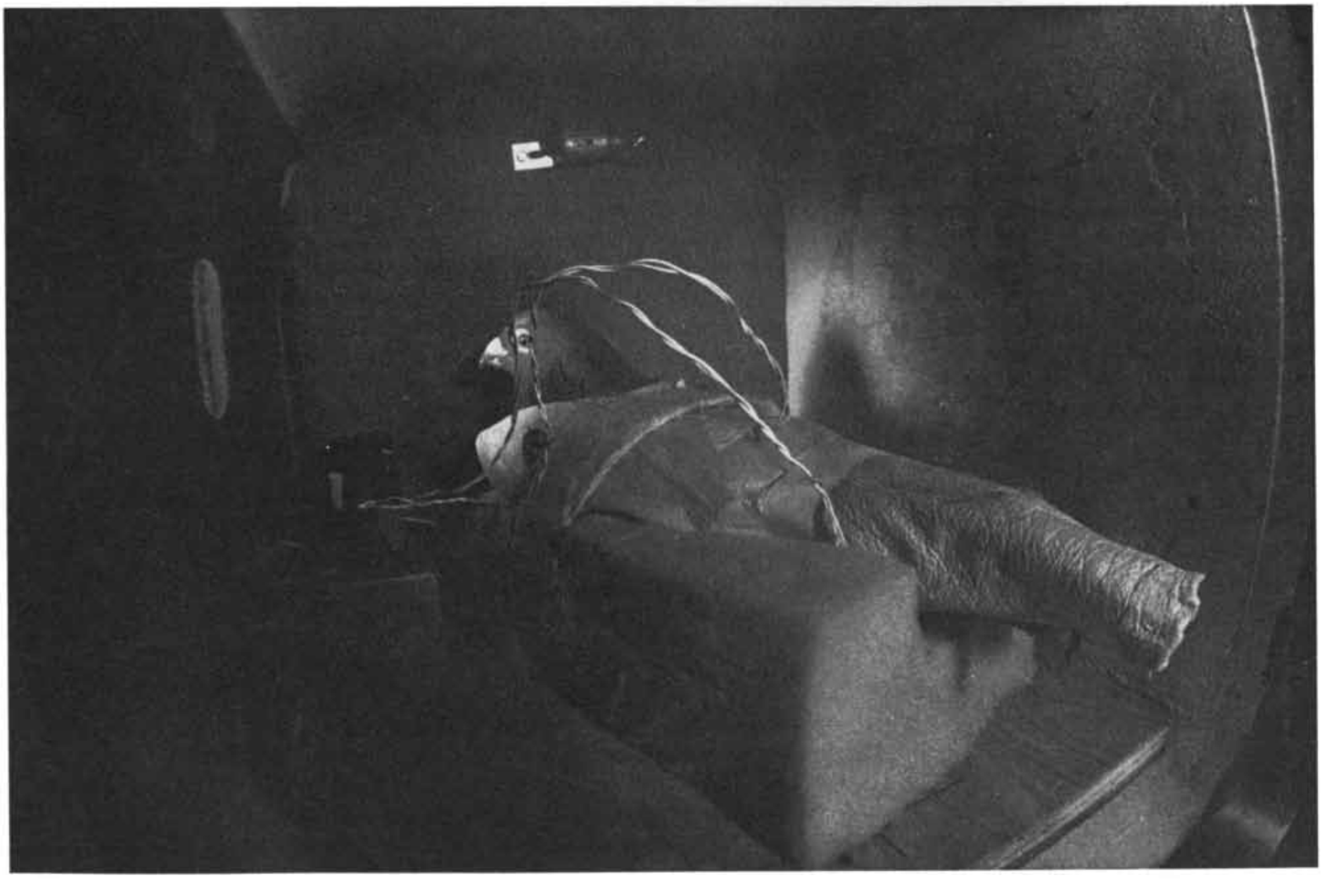
THE MYSTERY OF PIGEON HOMING

Recent findings have upset previous explanations of how pigeons find their way home from distant locations. It appears that they have more than one compass system for determining direction

by William T. Keeton

How does a homing pigeon find its way back to its home loft from hundreds of miles away? The answer does not lie in visible landmarks; pigeons taken in covered cages to areas they have never seen before have little trouble finding their way home. Nor does it lie entirely in the bird's ability to determine compass directions from the sun or the earth's magnetic field. Even when a pigeon can determine compass directions, how can it know which way home is? Although the homing prowess of the pigeon has long engaged the curiosity of man, the full story of how the bird navigates still remains a mystery. Nonetheless much has been learned about the pigeon's navigational abilities in the past two decades, particularly in the past six years.

The modern homing pigeon, a de-



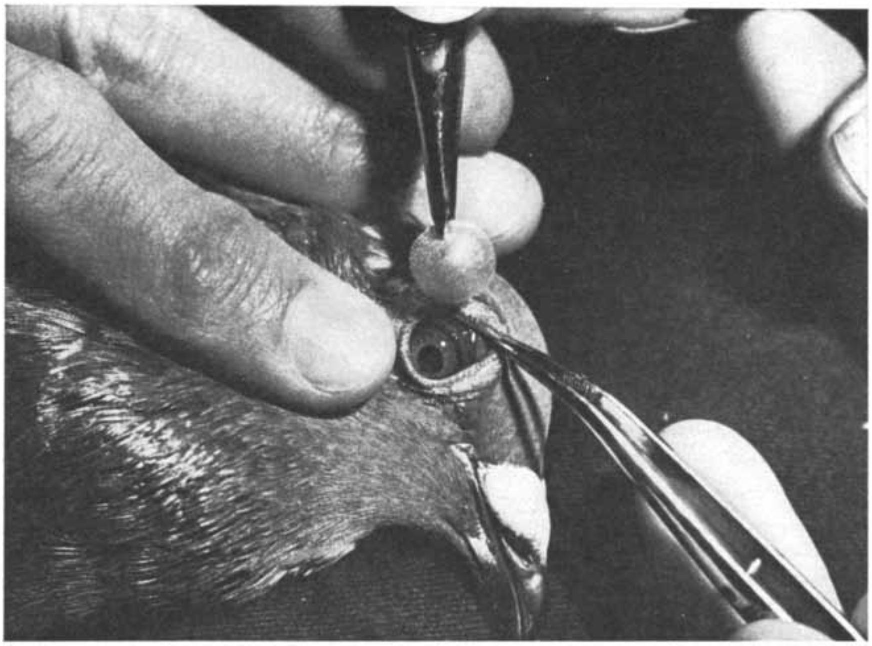
PIGEON IN ISOLATION CHAMBER has been prepared for tests of its unusual sensory capabilities. Two wires go to electrodes on the pigeon that give it a mild electric shock during the test, and two other wires are connected to electrodes that pick up the bird's heartbeat. The pigeon is restrained by a harness to keep it from moving. In a typical experiment the bird receives a shock following a specific stimulus, which might be a change in the strength of an induced magnetic field, a change of air pressure or a change in the plane of polarized light falling on the pigeon's eye. If the bird is able to sense the changes in the stimulus, it begins to anticipate the shock and its heart rate increases at the beginning of the stimulus, which is given at random time intervals. Experiments in the author's laboratory indicate that the pigeon is capable of sensing tiny fluctuations in atmospheric pressure. In addition pigeons, like honeybees, can detect changes in the plane of polarized light. The photograph was made with a camera placed inside the chamber.

scendant of several earlier breeds of pigeons, was developed in Belgium in the middle of the 19th century. Today, in addition to serving as message carriers, homing pigeons are raised for competitive racing. This sport is widespread and very popular in Europe, and has become firmly established in many parts of the U.S. as well. Often several thousand pigeons are entered in a single race. The birds are shipped to a designated point, usually between 100 and 600 miles away, and are then released simultaneously. After the owners of the birds have recorded the arrival times at home, using special devices designed for the purpose, the speed of the individual birds is calculated to determine the winners. Speeds of 50 miles per hour are common; the best pigeons can make it home from 600 miles away in a single day.

The remarkable ability of pigeons to find their way home has been known for at least as long as there has been written history. The armies of the ancient Persians, Assyrians, Egyptians and Phoenicians all sent messages by pigeon from the field. It is known that regular communication via pigeon existed in the days of Julius Caesar. During the siege of Paris in 1870 more than a million messages reached Parisians by means of pigeons that had been smuggled out of the city in balloons. Pigeons did such valuable service in both world wars that monuments in their honor were erected in Brussels and in the French city of Lille. In the U.S. some famous pigeon "heroes" were stuffed and mounted after their death; they are on display at the Army Signal Corps Museum and the National Museum.

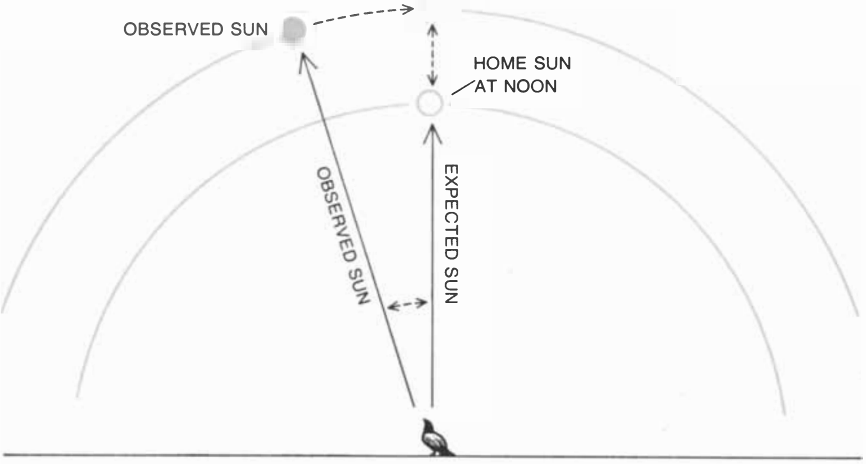
In 1949 Gustav Kramer and his students at the Max Planck Institute for Marine Biology at Wilhelmshaven in Germany demonstrated that a pigeon in a circular cage with identical food cups at regular intervals around its periphery could easily be trained always to go to a food cup located in a particular direction, for example the northwest, even though the cage was rotated and the visual landscape around it was changed. They found that the pigeon's ability to determine a direction depended on the bird's being able to see the sun. Under a heavy overcast the bird's choice of food cups became random. If the sun's apparent position was altered by mirrors, the pigeon's choice of food cups was correspondingly altered.

It is obvious that if birds can use the sun as a compass to determine directions, they must be able to compensate for the change in the sun's apparent position

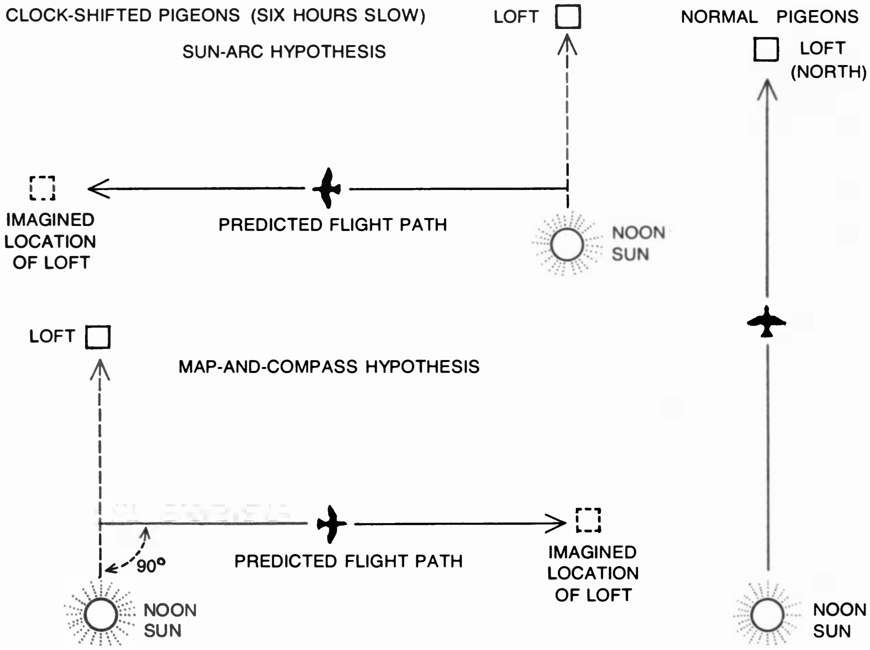


FROSTED CONTACT LENS is placed on a pigeon's eye before a test release. When both of the pigeon's eyes are covered by the lenses, the pigeon is unable to see objects that are more than a few yards away. Control pigeons have clear lenses put on their eyes and are released at the same time from the site. Experiments by Klaus Schmidt-Koenig and H. J. Schlichte of the University of Göttingen, who developed the technique, have demonstrated that pigeons wearing the frosted lenses are able to orient their flight in a homeward direction when they are released at a distant site, and that some pigeons are able to fly back to their home loft. Lenses currently in use are made of a gelatin that dissolves in a few hours.

NIGHT NAVIGATION of pigeons is being studied by Cornell workers. Pigeons with radio transmitters on their back are released and tracked by a radio receiver in the truck. In this time exposure three successive light flashes were used to illuminate the flying pigeon.



SUN-ARC HYPOTHESIS was proposed by G. V. T. Matthews of the University of Cambridge in the 1950's to explain how pigeons could obtain from the sun alone all the information required to determine their north-south and east-west displacement from home. For example, if a pigeon were released at noon at an unfamiliar site that is southwest of its home loft, the bird would observe the sun's motion and quickly extrapolate along the sun's arc across the sky to the noon position. It would then compare the sun's noon altitude with the remembered altitude at noon at home. Since the bird is south of home, the sun would be higher at the release site, and the pigeon would know that it has to fly north to make the sun appear lower. In order to determine its east-west displacement the pigeon would compare the position of the sun at the release site with the position the sun should have according to the bird's internal clock. In this instance the bird's clock would inform it that the time at home is noon. The sun at the release site, however, is at an altitude lower than that at noon, so that the bird knows it must fly east. By combining the two displacements, the bird would know that it should start flying to the northeast to get to its home loft.



TESTS OF TWO SUN NAVIGATION HYPOTHESES were made with pigeons whose internal clocks had been shifted six hours slow by altering their day-and-night time periods in the laboratory. According to the sun-arc hypothesis, when the clock-shifted pigeons are taken south of their home loft and released at noon, their internal clock tells them that it is 6:00 A.M. at home. They observe that the sun at the release site is too far along its arc for 6:00 A.M. and that they therefore should fly west (*top left*). The alternative map-and-compass hypothesis suggests that the pigeons know where they are relative to home from some kind of map, and that they use the sun only to get compass direction. Their internal clock says it is 6:00 A.M., and they therefore assume that the sun is in the east. Since the sun is really in the south, they should begin flying east, thinking that this direction is north (*bottom left*). When the experiment was carried out, the clock-shifted pigeons flew east, thereby supporting the map-and-compass hypothesis and contradicting the sun-arc hypothesis. Normal pigeons released at the same site departed in the correct homeward direction (*right*).

during the day. In the Northern Hemisphere the sun rises in the east, moves through south at noon and sets in the west. If a pigeon is to determine a particular direction, it cannot simply select a constant angle relative to the sun. It must change the relative angle by about 15 degrees per hour, which is the average rate of change of the sun's position throughout the day. In short, the bird must have an accurate sense of time, an internal clock, and that clock must somehow be coupled with the position of the sun in the sky if an accurate determination of direction from the sun is to be possible.

In a simple but elegant fashion Kramer demonstrated that birds do indeed compensate for time when they are using the sun as a compass. He trained some birds, in this case starlings, to use the sun to go in a particular direction to get to a food cup. He then substituted a stationary light for the sun. The starlings responded to the light as though it were moving at 15 degrees per hour. Since the light was in fact stationary, the bearing taken by the birds shifted approximately 15 degrees per hour.

Klaus Hoffmann, one of Kramer's students, went an important step further in demonstrating the role of the internal clock in sun-compass orientation. He kept starlings for several days in closed rooms where the artificial lights were turned on six hours after sunrise and turned off six hours after sunset. It is known that the internal clocks of most organisms can be shifted to a new rhythm in this manner; the process is very similar to what is experienced by a human being who flies from the U.S. to Europe in a few hours and then takes several days to adjust to European time. When the starlings whose internal clocks had been shifted six hours slow were tested in a circular cage under the real sun, they selected a bearing 90 degrees to the right of the original training direction. Since their internal clocks were a quarter of a day out of phase with sun time, they made a quarter-circle error in their selection of food cups.

Although Kramer and his colleagues had clearly demonstrated that some birds, including pigeons, can use the sun as a compass, their discovery by itself cannot explain how pigeons home. As I have indicated, homing requires more than a compass. If you were taken hundreds of miles away into unfamiliar territory, given only a magnetic compass and told to start walking toward home, you would not be able to get there. Even though you could determine where north


Admiral
now a part of
Rockwell International
Admiral
SUPER-SOLARCOLOR
3

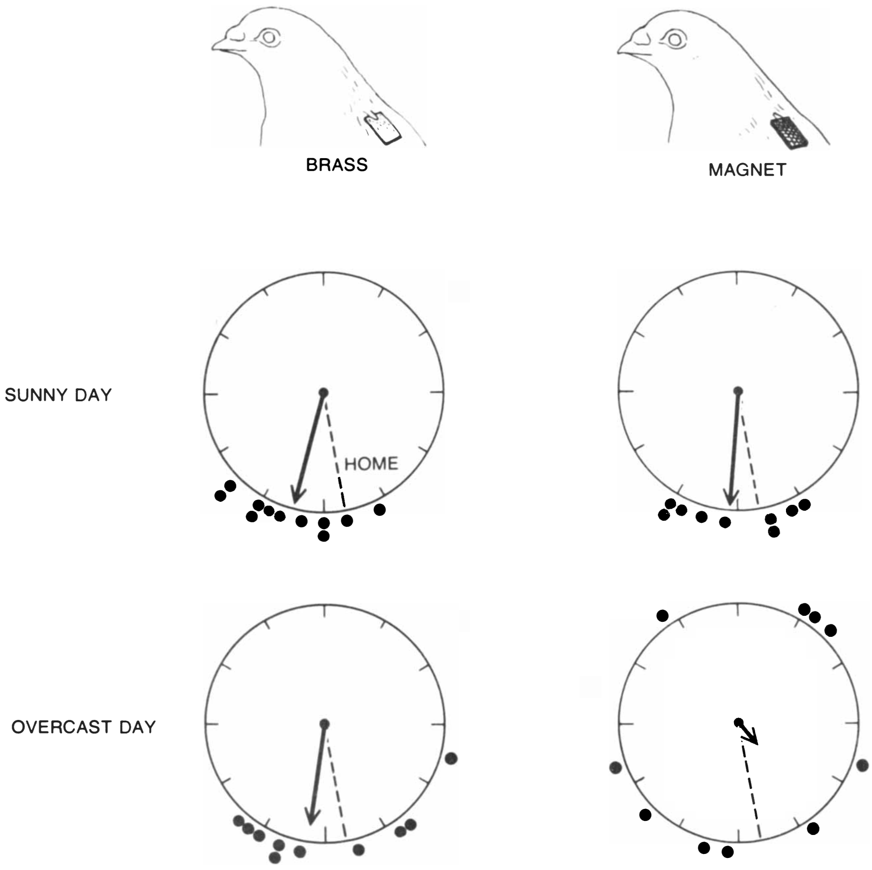

SMIRNOFF SILVER®
THE SILVER MARTINI.
FOR PEOPLE WHO WANT A SILVER LINING WITHOUT THE CLOUD.
SMIRNOFF SILVER. NINETY POINT FOUR PROOF. SMIRNOFF LEAVES YOU BREATHLESS®
SMIRNOFF SILVER® VODKA, 90.4 PROOF. DISTILLED FROM GRAIN. STE. PIERRE SMIRNOFF FLS. (DIV. OF HEUBLEIN, INC.) HARTFORD, CONN.

was, you would not know where you were with respect to home, hence such compass information would be nearly useless.

In 1953 G. V. T. Matthews, who was then working at the University of Cambridge, suggested that pigeons get far more information than compass bearings from the sun. He hypothesized that the sun gives them all the information they need to carry out true bicoordinate navigation. Stated briefly, on release at a distant site a pigeon would determine its north-south displacement from home by observing the sun's motion along the arc of its path across the sky, extrapolating to the sun's noon position on that arc, measuring the sun's noon altitude and comparing it with the sun's noon altitude at home (as the bird remembered it). If the sun's noon altitude at the release site was lower than it was at home, the bird would know that it was north of home; if the sun was higher than it was at home, the bird would know that it was south of home. To calculate the east-west displacement the bird would determine local sun time by observation of the sun's position on its arc at the release site and compare the local time with home time as indicated by its internal clock. A local time ahead of home time would indicated the bird was east of home; a local time behind home time would indicated the bird was west of home. Thus, according to Matthews, the bird would determine its north-south displacement from the sun's altitude and its east-west displacement by the time difference; combining these data would indicate the homeward direction [*see top illustration on page 98.*]

Matthews' sun-arc hypothesis was a major stimulus to further research on pigeon homing, and it formed the basis for many of the experiments conducted in the following decade. Unfortunately, however, nearly all the results of these experiments contradicted the hypothesis, and investigators actively engaged in research on pigeon homing today no longer regard it as being probable. The evidence against the hypothesis is so extensive that most of it cannot be discussed here. For the moment I shall mention only one kind of experiment to help the reader understand some of the more recent research.

Klaus Schmidt-Koenig, another of Kramer's students, showed in 1958 that when pigeons whose clocks have been artificially shifted are released at a distant site, their initial choice of direction is shifted. Their vanishing bearings (the bearings at which they vanish from the view of an observer using high-power binoculars) deviate from those of normal pigeons by 15 degrees for each hour the birds have been clock-shifted.

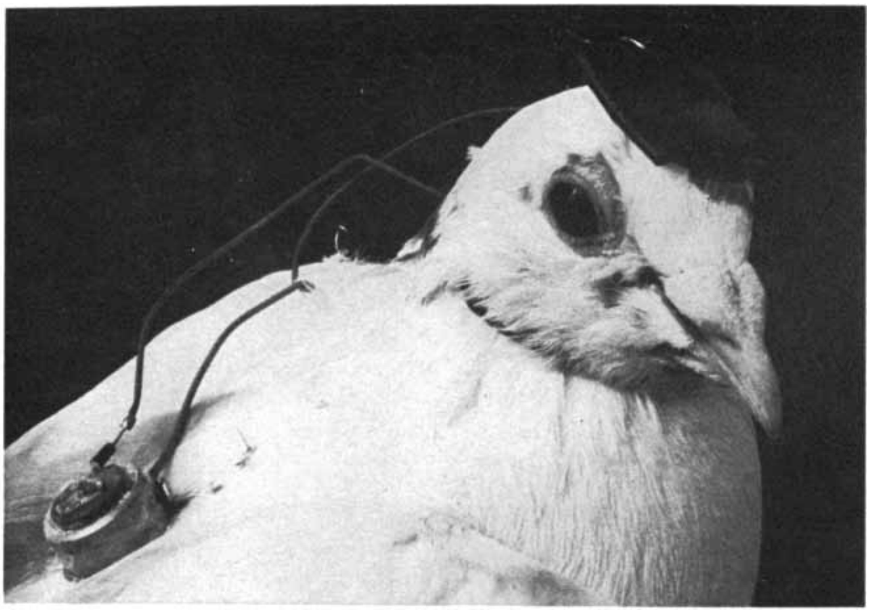


MAGNETIC-FIELD HYPOTHESIS, proposed more than a century ago, had been rejected until recently because earlier experiments failed to show that putting a magnet on a pigeon disorients its homing. Recent tests show, however, that pigeons with bar magnets attached to them are disoriented when they are released at an unfamiliar site under a total overcast but are not disoriented when the sun is visible. Control pigeons with brass bars attached to them show little difference in their mean vanishing bearing under the sun or an overcast. The vanishing bearings of individual pigeons, as determined by an observer with binoculars, are shown by the solid circles. The broken line indicates the true home bearing. The mean vector, or directional tendency, of all the birds in a test group is shown by the arrow. The length of the mean vector is a statistical representation of the degree of agreement among the birds in selecting a direction. Perfect agreement would give a vector length equal to the circle's radius; the more scattered the departing directions, the shorter the vector.

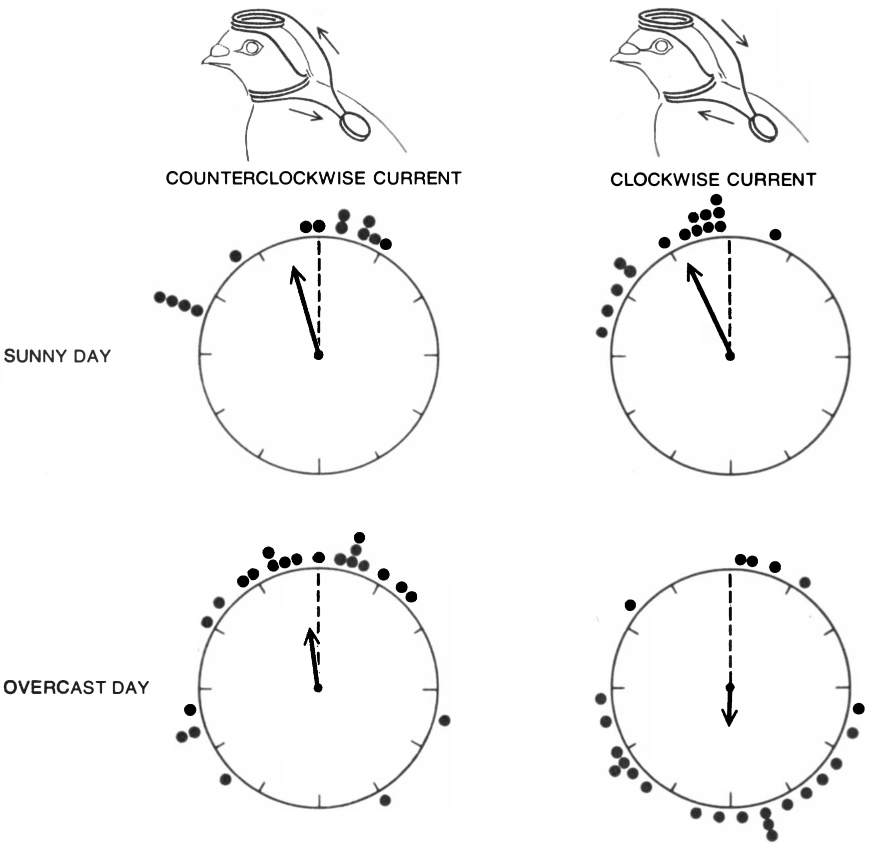
Let us examine a test involving clock-shifted pigeons to see whether or not the results agree with what would be predicted by the sun-arc hypothesis. Suppose we shift the birds' internal clock so that it is six hours slow and then release them at noon 100 miles south of their home loft. According to the sun-arc hypothesis, the birds would observe that it is noon at the release site, but their internal clock would tell them it is only 6:00 A.M. at home. They should therefore react as though they were thousands of miles east of home, and they should start flying almost due west. When such an experiment is actually performed, however, the birds vanish nearly due east, exactly opposite what the sun-arc hypothesis predicts [*see bottom illustration on page 98*].

Is there any way we can make sense of these results? The answer is yes, but to do so we must turn from Matthews' sun-arc hypothesis to an alternative model proposed by Kramer. Kramer emphasized that all the evidence supports the conclusion that pigeons get only compass information from the sun and nothing else. They appear to behave in a manner analogous to a man who uses both a map and a compass, as though they first determine from some kind of map where they are relative to home and in which direction they must fly to get home and then use the sun compass to locate that direction.

Since Kramer could never explain



HELMHOLTZ COILS above the pigeon's head and around its neck induce a relatively uniform magnetic field through its head. The coils are powered by a small mercury battery on the bird's back. Direction of the induced field can be reversed simply by reversing the connections of the battery. The strength of the magnetic field can be varied by controlling the amount of current passing through the coils. Battery is exhausted in two or three hours.



PIGEONS WITH HELMHOLTZ COILS in which the current flows counterclockwise (south-seeking pole of a compass in the induced magnetic field points up) fly almost directly homeward on both sunny and overcast days. When the current in the coils is made to flow clockwise (north-seeking pole of a compass in the induced magnetic field points up), the pigeons still fly homeward on sunny days, but on overcast days they fly almost 180 degrees away from home. These results were obtained in several experiments conducted by Charles Walcott and Robert Green of the State University of New York at Stony Brook.

what the source of the map information might be, let us for the sake of our example pretend that before we release each pigeon we whisper in its ear, "Home is due north." Now the bird must use its sun compass to locate north. Its internal clock says it is 6:00 A.M., when the sun should be in the east; hence north should be approximately 90 degrees counterclockwise from the sun. Remember, however, that the bird's clock is six hours slow; it is actually noon, when the sun is in the south. Hence the bird's choice of a bearing 90 degrees counterclockwise from the sun sends it east, not north. We can summarize by saying that no matter what combination of directions and clock-shift we use in actual experiments, the results come out consistent with the predictions of Kramer's map-and-compass model and not with Matthews' sun-arc hypothesis.

Because the pigeon's use of the sun compass in orientation was the one thing that was firmly established, there was a tendency in the 1960's for many investigators to assume that the sun is essential for homeward orientation at an unfamiliar release site. Several discrepancies, however, led me and my colleagues at Cornell University to doubt it. First, I knew of numerous instances of fast pigeon races under heavy overcast. Second, our pigeons seemed to perform well under overcast if they had first been made to fly in the vicinity of the home loft in rainy weather. Third, the published evidence that pigeons were disoriented under heavy overcast was not entirely consistent, and fourth, we and others had been able to get pigeons to home at night.

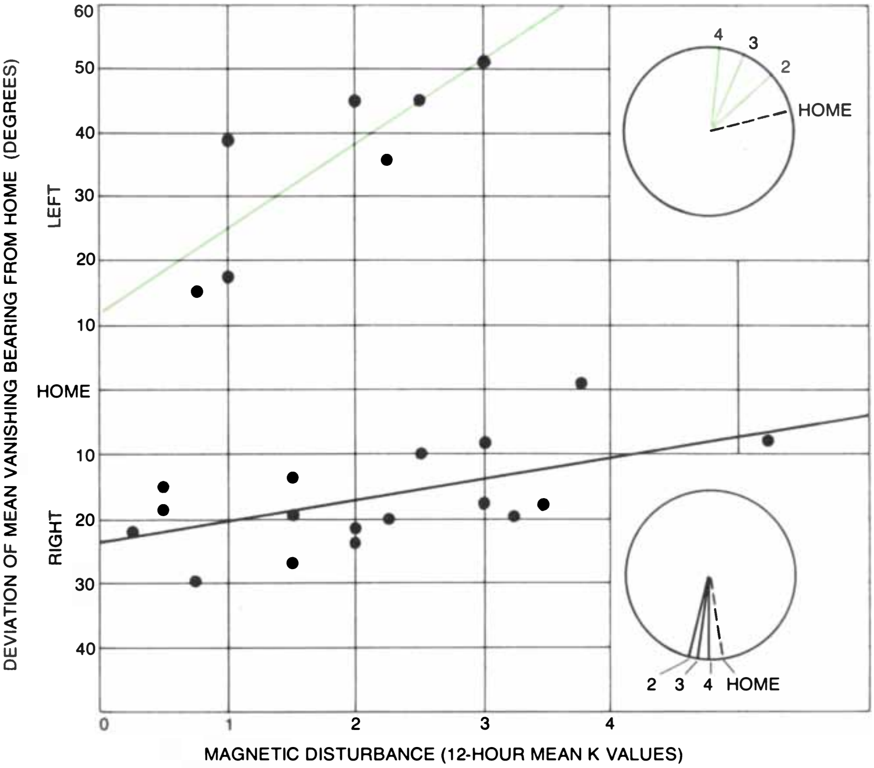
We set out to reexamine the importance of the sun in pigeon navigation. In our most important experiments we too used clock-shifted pigeons. As we expected, when pigeons whose internal clocks had been shifted six hours fast or slow were released under sunny conditions, their vanishing bearings were roughly 90 degrees to the right or left of the vanishing bearings of control pigeons whose internal clocks had not been altered. When the pigeons were released in total overcast, however, the results were quite different: both the clock-shifted birds and the control birds vanished toward home and there was no significant difference in their bearings. This was true even when the release site was completely unfamiliar to the pigeons.

These results led us to several conclusions: (1) Pigeons accustomed to flying in inclement weather are able to orient homeward under total overcast.

Since there is no difference between the bearings of control and clock-shifted birds under such conditions, it is clear they are not able to see the sun through the clouds and hence are no longer using the sun compass. (2) There must be redundancy in the pigeons' navigation system. They use the sun compass when it is available, but they can substitute information from other sources when it is not. (3) The alternative information used in lieu of the sun compass does not require time compensation. (4) The alternative system cannot be pilotage by familiar landmarks, because pigeons can correctly orient themselves homeward under overcast even in distant, unfamiliar territory.

Recognition of the fact that pigeons are able to use alternative cues, depending on the circumstances involved, meant that the results of many older experiments could no longer be accepted at face value. For example, if an experimenter altered cue *A* while keeping other conditions optimum, and if the pigeons continued to orient well, they may simply have used cue *B* as an alternative to *A*. Similarly, if *B* was altered while everything else was kept at an optimum and the birds oriented well, they may have used *A* as an alternative to *B*. In short, we would have been wrong if we had concluded from these experiments that cues *A* and *B* are not elements in the pigeons' navigation system. In fact, such experiments show only that neither *A* nor *B* alone is essential for proper orientation under the particular test conditions.

This kind of reasoning led us to conduct experiments in which we varied several possible orientational cues simultaneously, on the assumption that if we could interfere with enough of them at the same time we could hope to learn which cues are more important and how they interact with one another.



DISTURBANCES OF THE EARTH'S MAGNETIC FIELD caused by solar activity appear to affect a pigeon's initial choice of bearing when it is released at a distant site under sunny conditions. The *K*-index scale is used to indicate the degree of magnetic activity, ranging from quiet (less than 2) to a major magnetic storm (6 or more). In 1972 a series of releases of Cornell University pigeons from a site 45 miles north of the home lofts revealed that, as the degree of magnetic disturbance increases, the vanishing bearing of the birds steadily shifts to the left as seen by an observer facing homeward (*black curve*). At this release site the shift to the left brought the birds' vanishing bearings closer to the true home bearing, but success in homing was not improved. In another series of tests, pigeons from a different loft were released from a site west of their home. A similar leftward shift of vanishing bearings with increasing magnetic disturbance was found (*colored curve*). In this instance the shift to the left caused the vanishing bearings to recede away from the true home bearing.

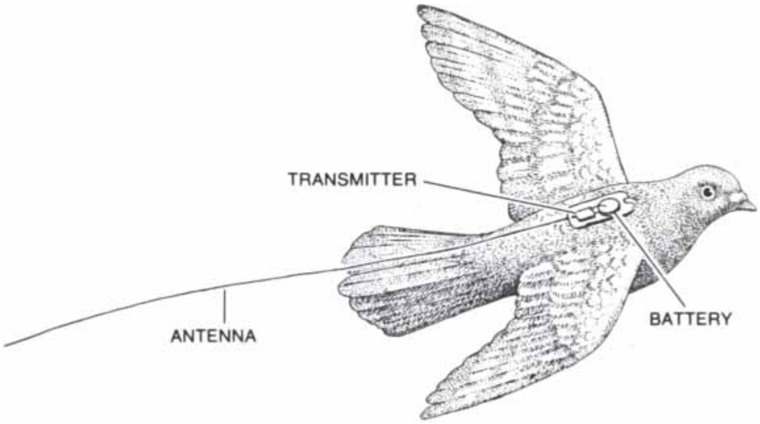
We chose first to look again at the old idea that birds might obtain directional information from the earth's magnetic field. Although this hypothesis had been known for more than a century, there was no evidence for it and much experimental evidence against it. Nonetheless, it seemed worth reexamining. And it was! When we repeated older experiments of putting bar magnets on pigeons to distort the magnetic field around them, we found, as had others before us, that the birds had no difficulty orienting on sunny days. When the test releases were conducted on totally overcast days, however, the birds carrying magnets usually vanished randomly whereas control birds carrying brass bars of the same size and weight vanished toward home. Several other workers have since repeated these experiments, with the same results.

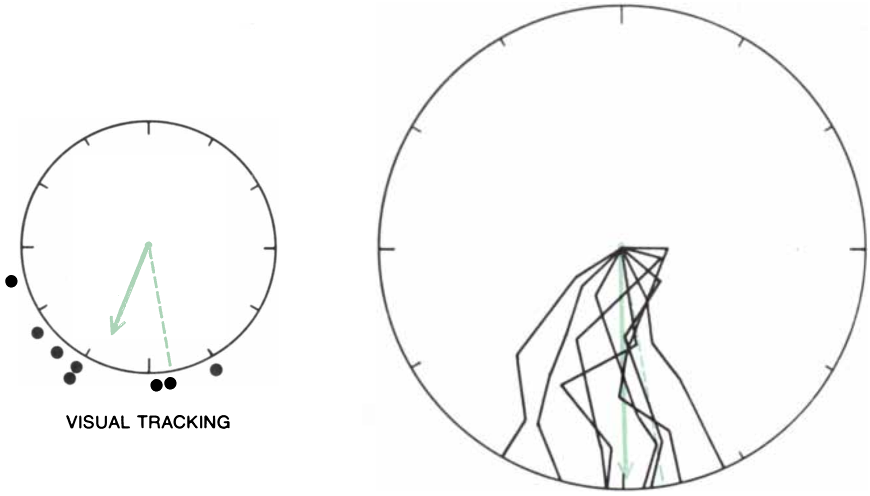
More recently Charles Walcott of the State University of New York at Stony Brook and his student Robert Green have gone one step further. Instead of working with bar magnets, they put a small Helmholtz coil on the pigeon's head like a cap and another coil around its neck like a collar. Power is supplied from a battery on the bird's back. This device makes it possible to induce a relatively uniform magnetic field through the bird's head. The direction of the induced magnetic field can be made to point up through the bird's head or to point down simply by hooking up the battery to make the current in the coils flow clockwise or counterclockwise. Under sunny conditions Walcott and Green found that the direction of the induced magnetic field did not affect the pigeon's ability to orient homeward. Under total overcast, however, the direction of the induced magnetic field had a dramatic effect: when the north-seeking pole of a compass in the induced field pointed up, the pigeons flew almost directly away from home, whereas when the south-seeking pole of a compass in the induced field pointed up, the pigeons oriented toward home.

Our results, together with Walcott's, suggest that magnetic information may play a role in the pigeon navigation system. This is consistent with the recent discovery by Friedrich Merkel and Wolfgang Wiltschko of the University of Frankfurt that European robins in circular cages can use magnetic cues to orient themselves in a particular direction. William Southern of Northern Illinois University also has reported that the orientation of ring-billed gulls is influenced by magnetic activity.

Recently Martin Lindauer and Her-man Martin of the University of Frankfurt have demonstrated that honeybees give orientational responses to magnetic cues several thousand times weaker than the earth's field. Only a few years ago biologists were debating whether or not any organism could detect a magnetic field as weak as the earth's (approximately half a gauss). The responses of honeybees to magnetic cues now makes us wonder if one gamma (10^{-5} gauss) will not prove to be the lower limit. Indeed, a study that my colleagues and I have recently conducted suggests that the magnetic-detection sensitivity of pigeons may rival that of honeybees. In four long series of tests over a period of three years we have found that fluctuations of less than 100 gamma (and probably less than 40 gamma) in the earth's magnetic field, caused by solar flares and sunspots, appear to have a small but significant effect on the pigeons' choice of an initial bearing at the release site.

The question of how organisms detect magnetic stimuli is unanswered. We have very little idea what a magnetic sense organ should look like, or even where in the body we should expect to find it. Since magnetic flux can pass freely through living tissue, magnetic detectors might be anywhere inside the body. The search for these detectors has already begun in our laboratory and in others throughout the world. It promises to be a challenging undertaking.





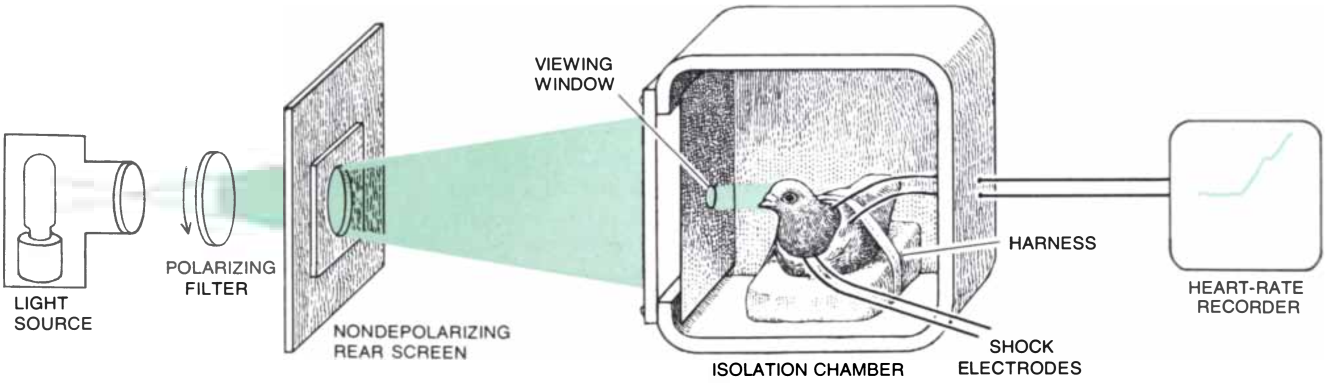
RADIO TRACKING OF PIGEONS is carried out from a receiver on the ground at the release site or sometimes by a receiver on an aircraft. An FM transmitter and a battery are glued on the pigeon's back (*top*). The 19-inch antenna trails behind the bird as it flies. Data from radio tracking reveal that pigeons do not continue to fly in a straight line after they leave a release site but frequently alter their course. The vanishing bearings for eight pigeons, as determined by observers with binoculars, and the bearings determined by simultaneous radio tracking from the site are compared (*bottom*). The scale of the two circles is arbitrary. Visual tracking extends to one or two miles, depending on the flying height of the bird. Radio tracking extends to eight miles or more. Broken line indicates home direction.

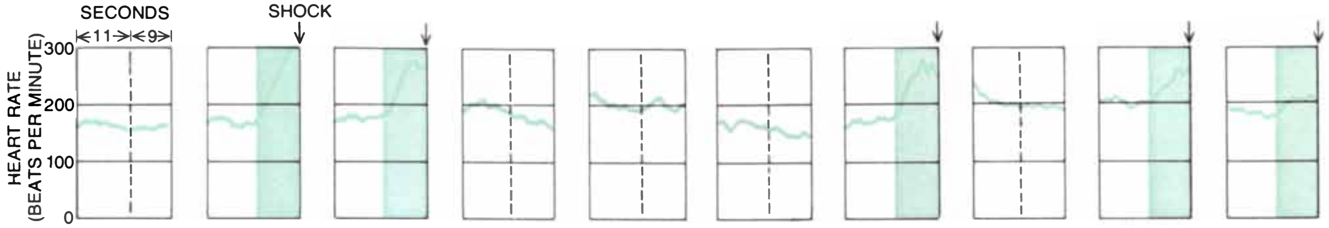
Exciting as the discovery that magnetism plays a part in avian navigation systems may be, we are in a sense back where we started. The weight of the evidence at present suggests that magnetism simply provides a second compass, not the long-sought map. Hence we must continue our search. What other sources of information might the birds have?

One possibility that comes readily to mind in this age of long-range rocketry is that the birds might be capable of inertial guidance, that they might somehow detect and record all the angular accelerations of the outward journey to the release site, then double-integrate them to determine the direction home. Intriguing as this possibility is, all the evidence is against it. Pigeons have been carried to release sites while riding on turntables or in rotating drums, yet this input of additional inertial "noise" has no effect; the birds orient homeward as accurately as control birds not so treated. Other pigeons that were carried to the release site while they were under deep anesthesia were able to determine the direction home with no difficulty. Pigeons with a variety of surgical lesions of the semicircular canals—the principal detectors of acceleration in vertebrates—orient themselves accurately, whether they are tested under sunny conditions or overcast ones.

The hypothesis that pigeons may be able to use olfactory information in navigating has been advocated by Floriano Papi and his colleagues at the University of Pisa. The probability that this is the case seems low in view of the relatively poor development of the pigeon's olfactory system. Nonetheless Papi has some interesting experimental results, and it is too early to make a judgment on his proposal. We are currently conducting experiments to test his ideas.

By now the reader may well be wondering why so little has been said about what might seem the most obvious possible cue for homing pigeons: familiar landmarks. The reason is that there is abundant evidence that landmarks play a very small role in the homing process. In the course of tracking pigeons by airplane Walcott and his colleague Martin Michener have repeatedly noted that when pigeons flying on an incorrect course encounter an area over which they have recently flown, they seldom give any indication of recognizing the familiar territory. Several other investigators, including members of our group, have found that pigeons clock-shifted six hours and released less than a mile from





PIGEON PERCEPTION OF POLARIZED LIGHT is being tested in the author's laboratory by Melvin L. Kreithen. After electrodes for administering electric shock and for detecting the heart rate are attached, the pigeon is put in a harness and then is placed in a sealed soundproof chamber. Light is projected through a polarizing filter on a rotating mount and then through a nondepolarizing rear-projection screen. Part of the light enters the isolation chamber through a small window and falls on the pigeon's eye. The light comes on at random intervals. In some trials, which are also determined at random, the polarizing filter starts to rotate after 11 seconds; in others it does not. When the filter rotates, the pigeon receives a shock at the end of the light signal. When the filter does not rotate, no shock is given. After a number of trials the pigeon's heart rate begins to rise rapidly at the beginning of the rotation of the polarizing filter, indicating that the bird is able to sense the change in the plane of polarized light and is anticipating the shock that is to come. Recordings from a series of tests of a pigeon are shown (*bottom*). The colored block indicates the interval during which the polarizing filter rotates. In control runs during the corresponding interval (*to right of broken line*) no rotation occurs.

home, in territory over which they have flown daily during their exercise period, often vanish 90 degrees away from the home direction. Only a direct view of the loft building itself takes precedence over what their navigation system is telling them; nearby buildings or trees apparently do not serve as reference points under these conditions. In fact, even a view of the loft is not always effective, particularly at distances of a mile or more.

Perhaps most convincing of all are experiments conducted by Schmidt-Koenig and H. J. Schlichte of the University of Göttingen. They put frosted contact lenses over the eyes of pigeons, thus making it impossible for the birds to see any object that is more than a few meters away. Not only do these pigeons orient homeward when they are released as far as 80 miles away but also a surprising number of them actually get home. Schmidt-Koenig conducted some of his experiments at our Cornell lofts, and thus I had the opportunity of observing them at first hand. It was a remarkable experience. The birds arrived very high overhead and fluttered down to a landing in the fields around the loft. Being unable to see the loft, they waited for us to pick them up and carry them the last few feet. These results suggest that the pigeon navigation system is often accurate enough to pinpoint the home location almost exactly without reliance on familiar landmarks; vision is necessary only for the final approach, frequently at a distance of less than 200 yards.

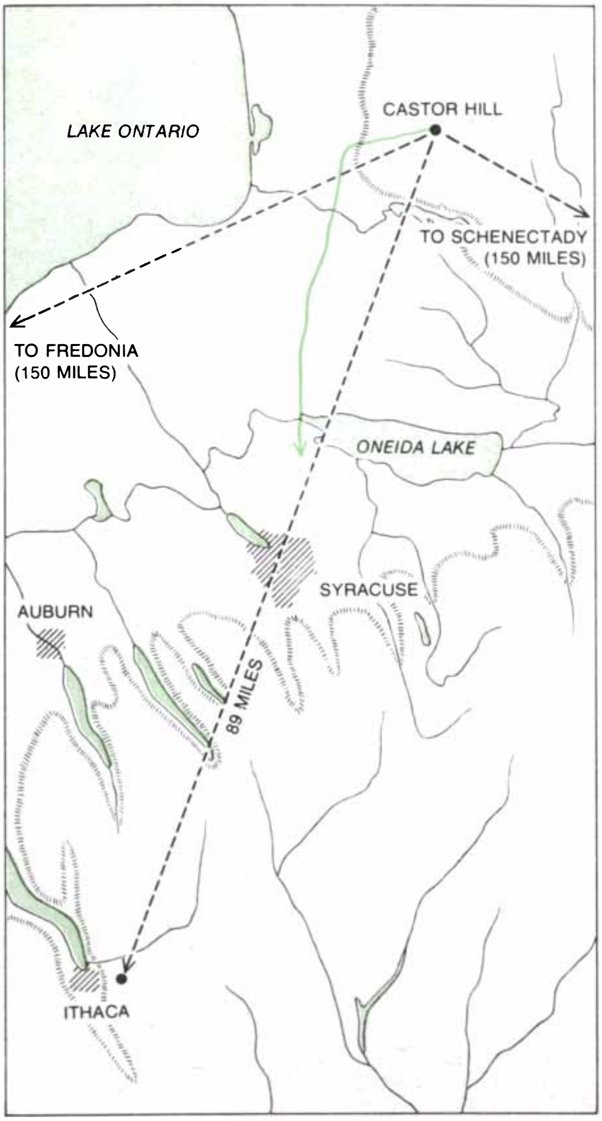
It will be apparent from all I have said that the task of uncovering the pigeon's navigation system is going to be a difficult one. The old idea that birds use a single method to determine the home direction has given way to the realization that there are probably multiple components in the system and that these components may be combined in a variety of ways, depending on such factors as weather conditions, the age of the bird and the bird's experience.

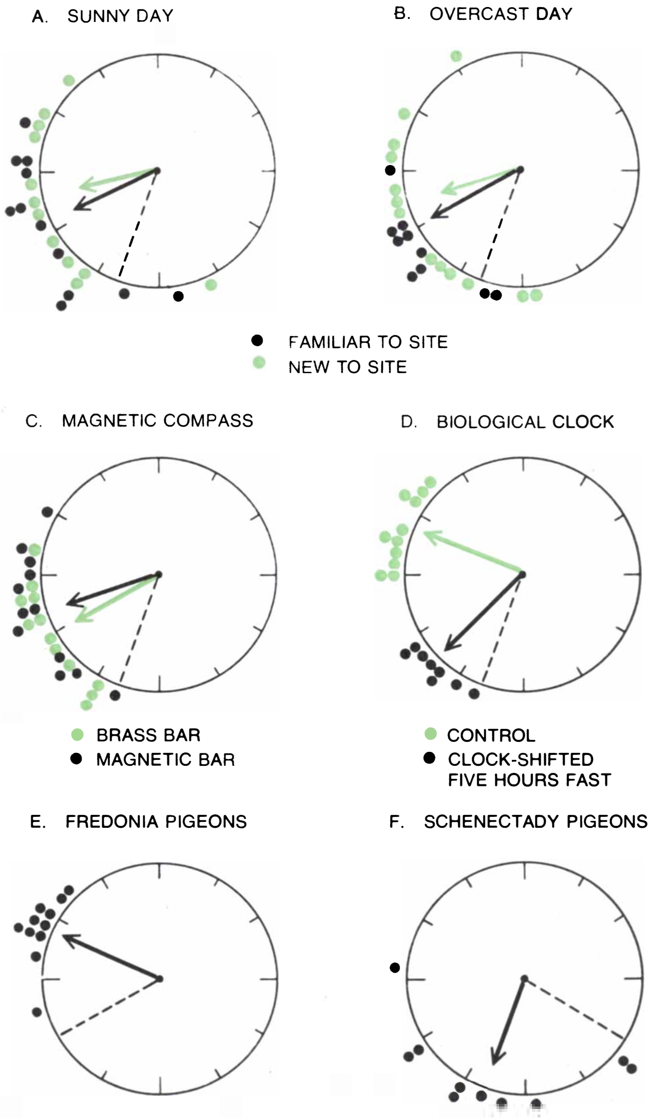
One approach that holds much promise for helping us tease apart the many elements in the system is the study of the ontogeny of navigational behavior. For example, we have found that bar magnets disrupt the initial orientation of very young pigeons released away from home for the first time in their life, even when the sun is visible. Moreover, normal first-flight youngsters cannot orient under total overcast, even if they have previously been released for exercise in inclement weather. It seems, then, that inexperienced homing pigeons need both sun information and magnetic information. Various other manipulations that have little effect on experienced birds also disorient first-flight pigeons.

Perhaps with experience a pigeon learns to orient accurately with less information. Or perhaps experience is necessary to enable the pigeons to settle on a weighting scheme that allows them to decide what to do when they get conflicting information from different sources. Early results from some current experiments indicate that by training very young pigeons under conditions in which we severely restrict, or eliminate altogether, certain normally important environmental cues such as the sun, we may be able to induce the birds to settle on weighting schemes for evaluating directional cues that are quite different from normal. The availability of such birds would greatly facilitate the carrying out of experiments designed to clarify the roles of cues that are normally difficult to alter.

Another approach we are actively pursuing is an attempt to learn more about the sensory capabilities of pigeons. The more we learn, the more we become convinced that birds live in a sensory world very different from our own. For example, my student Melvin L. Kreithen has recently demonstrated that pigeons are remarkably sensitive to tiny changes in barometric pressure. Such a sensitivity might enable pigeons to get navigationally useful information from pressure patterns in the atmosphere. Kreithen and I have also recently obtained experimental evidence that pigeons can detect the plane of polarized light, which might mean that pigeons, like honeybees, can continue using the sun as a compass on partially overcast days, when the sun's disk is hidden from sight but some blue sky remains visible.

It has long been known that the bearings chosen by pigeons at distant release sites, although roughly in the homeward direction, are almost never oriented directly toward home. Moreover, the mean bearings of repeated releases at any given site usually show a consistent deviation from home; there is, in effect, a relatively stable "release-site bias" that is characteristic of each location. At some sites the bias is apparent only in the vanishing bearings obtained by visual tracking. The bias often becomes less marked when the final-contact bearings obtained by radio tracking are used, but at some sites the bias is still manifest when the birds move out of radio range (between six and 10 miles from the release site). In the hope that these biases might prove to be a key to local geophysical factors that could provide at least part of the map information for pigeons, we chose for intensive study several release sites where the biases were unusually large.





CASTOR HILL is approximately 89 miles northeast of the Cornell University pigeon lofts in Ithaca, N.Y. At Castor Hill the Cornell pigeons regularly choose an initial bearing that deviates clockwise from the true home direction. This occurs both on sunny days and on overcast days (*A* and *B*); thus the characteristic bias must not depend on the sun compass. The bias probably is not due primarily to the magnetic compass either, since under sunny conditions birds with bar magnets have about the same vanishing bearing as pigeons with brass bars (*C*). Pigeons whose internal clocks have been shifted five hours fast choose a more homeward direction than normal pigeons, but the clock-shifted pigeons are less successful at getting home (*D*). Pigeons from another loft at Fredonia, N.Y., 150 miles west of Ithaca (*E*), and pigeons from Schenectady, N.Y., 150 miles east of Ithaca (*F*), choose bearings that also deviate clockwise from the true home bearing. It appears that the bias is a function of the site and not of the pigeons. The actual flight path of normal Cornell pigeons released at Castor Hill and tracked by an airplane is shown by the colored line on the map.

One such site is the Castor Hill Fire Tower, located 89 miles north northeast of our Cornell lofts. Here our pigeons regularly depart with a mean bearing that deviates roughly 60 degrees clockwise from the direction of home [*see illustration on opposite page*]. In a long series of experiments we have found that this clockwise bias is evident not only with experienced pigeons new to the site but also with pigeons that have been released at the site before. The same bias is found in very young pigeons on their first homing flight. It is found on both sunny and overcast days, so that it apparently has nothing to do with the sun compass, and it is found in pigeons wearing magnets, so that it probably has nothing to do with the magnetic compass either. It is even found when the pigeons are wearing frosted contact lenses; hence it must not depend on anything the pigeons see.

Wild bank swallows captured near Cornell and released at Castor Hill showed the same clockwise bias, indicating that the biasing factor, whatever it may be, affects other bird species in the same way. Pigeons borrowed from lofts 150 miles east and west of Cornell and released at Castor Hill show a similar clockwise departure bias relative to their home. Finally, pigeons clock-shifted five hours fast depart nearly straight toward home from Castor Hill but nonetheless have poorer homing success than control pigeons that depart with the usual 60-degree bias. In a joint experiment with Walcott, normal pigeons with radio transmitters attached to them were tracked by an airplane after their release from Castor Hill. We found that the birds turn onto a more homeward course when they are approximately 14 to 18 miles west of Castor Hill. It may be that the clock-shifted birds that have poor homing success make a corresponding turn when they are a similar distance from the release site and thus become directed away from home. We hope soon to learn if this is so.

We conclude, then, that the bias in the birds' initial bearings is not a biological error, that it is due not to some peculiarity of the birds but to a peculiarity of the location. The birds are probably reading the map cues correctly but the map itself is twisted clockwise at Castor Hill. Perhaps if we can learn what geophysical factors are responsible for this distortion of the map we will finally be on the way to understanding the ancient mystery of how pigeons home.

The Detection of Neutral Weak Currents

The recent experimental discovery of a hitherto unobserved class of elementary-particle interactions supports the view that a deep link exists between the weak force and the electromagnetic force

by David B. Cline, Alfred K. Mann and Carlo Rubbia

Progress in the understanding of nature has often come through the recognition that seemingly diverse phenomena have a common origin. The classic example of such a unification was provided in the 19th century when electricity, magnetism, light and radio waves were all found to be linked by the equations of electromagnetism formulated by James Clerk Maxwell. We may now be on the verge of a comparable unification in the domain of elementary-particle physics. At present the interactions of matter at the subatomic level are thought to involve four distinct kinds of force: gravitational, electromagnetic, strong (or nuclear) and weak. Recent experiments, however, tend to support the theoretical view that at least two of these forces—the electromagnetic force and the weak force—are aspects of the same force [see "Unified Theories of Elementary-Particle Interaction," by Steven Weinberg; SCIENTIFIC AMERICAN, July].

Here we shall describe a series of experiments in which a beam of high-energy neutrinos is used to produce a hitherto unobserved class of weak interactions characterized as neutral-weak-current interactions. The results of these experiments, obtained over the past year or so by investigators at several high-energy accelerator centers, including our own group at the Fermi National Accelerator Laboratory ("Fermilab"), supply important evidence suggesting that the weak force is connected at an underlying level not only with the electromagnetic force but also possibly with the strong force.

The modern conception of the electromagnetic force operating between two moving electrically charged particles visualizes the interaction as taking place in two stages: first a photon, or quantum of electromagnetic radiation, is emitted by one charged particle, and then the photon is absorbed by the other particle [*see upper illustration on page 112*]. The exchanged photon is viewed as being the carrier of the electromagnetic force.

In another description of the same process the motion of one charged particle gives rise to an electric current, which attracts or repels a similar current associated with the other particle. Here the force is regarded as operating between the two currents. Such a current-current interaction is analogous to the interaction between two wires carrying electric current.

The two descriptions lead to equivalent results. A fundamental characteristic of the electromagnetic interaction is that the net electric charge of the particles before and after the interaction is strictly conserved. Accordingly the process is described as an interaction between two neutral currents, meaning that the interaction does not change the electric charge of the participating particles.

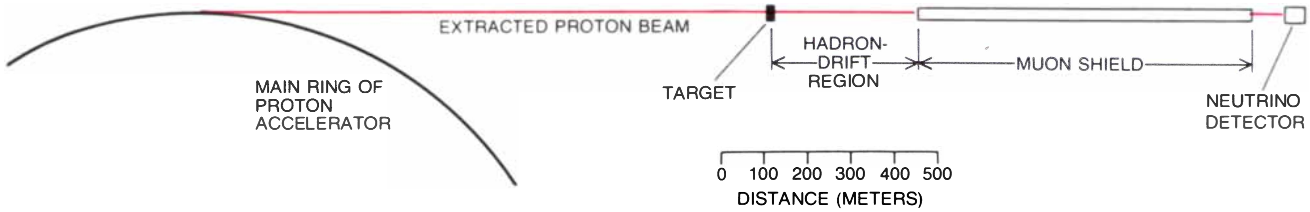
The weak force was unwittingly discovered in 1896 by Henri Becquerel when he first observed the radioactive decay of an atomic nucleus. As we now understand it, the nuclear transformation that manifests itself as radioactivity is mediated by the weak force and in general results in the creation of two particles: an electron and a neutrino. The electron could be readily observed in the experimental equipment of Becquerel's day, but the neutrino, having no mass and being electrically neutral, escaped detection. It was not until some 20 years ago that the neutrino was finally observed indirectly through its interaction with matter.

Beginning with the experiments of Becquerel the weak force has been studied intensively at low energies, with the experimental events being supplied by radioactive decay. In a sense this work is analogous to the study of static electricity, a line of inquiry that preceded the study of electricity as we know it by centuries. In this analogy the investigation of the weak force at high energies with the aid of neutrino beams corresponds to the study of electricity as we know it.

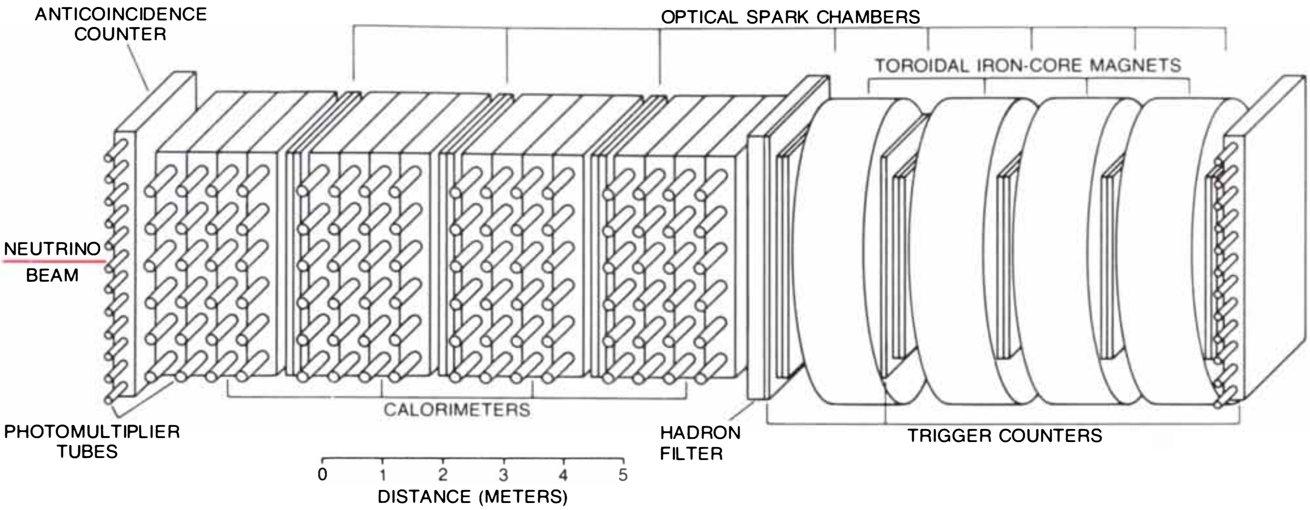
Since weak interactions are not yet completely understood, a detailed theoretical account of the mechanism that generates the weak force cannot now be given. Nonetheless the general rule that the forces between two particles result from an exchange of other particles is basic to much of our present understanding of elementary-particle interactions, and it is natural to apply this rule to the weak interactions as well. The hypothetical particle that has been assigned to carry the weak force is called the intermediate vector boson or the *W* particle. Over the past 20 years there have been repeated attempts to discover the *W* particle, but none have met with success. Because of a fundamental principle of quantum theory which states that the range of any force is inversely proportional to the mass of the exchanged particle, it is evident that if the *W* particle exists, it must be very massive.

Of course, it could be that the description of the weak force as being carried from one point to another by a particle is incorrect. Just as it was impossible to infer the behavior of moving electric charges from static electricity, it might also be impossible to infer the nature of high-energy weak interactions from radioactive decay. That is why there is now such great interest in the high-energy neutrino experiments, since it appears that for the first time the nonstatic behavior of the weak force is being observed.

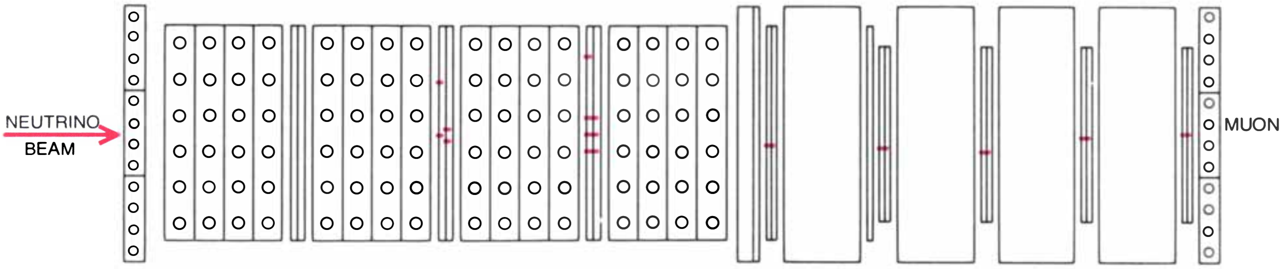
Just as some people suspected long be-



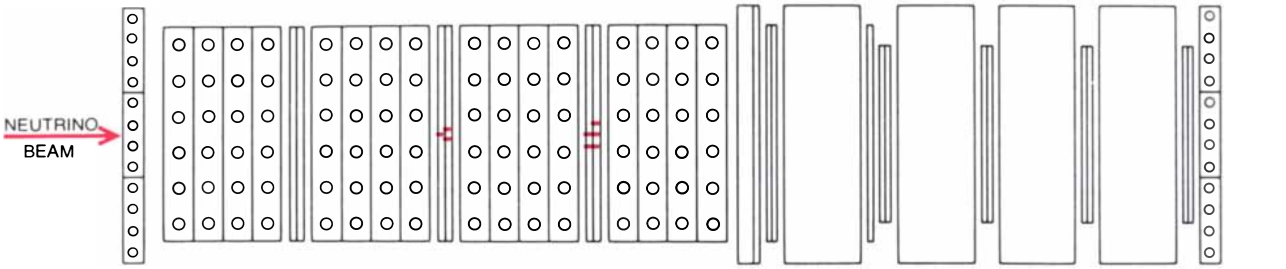
ONE OF THE EXPERIMENTS in which neutral weak currents were detected was conducted at the Fermi National Accelerator Laboratory in Batavia, Ill. ("Fermilab"). Protons from the 30-billion-electron-volt accelerator struck a target, generating a shower of other particles. The hadrons (particles that participate in strong interactions) decayed in the drift region; the muons (members of the class of leptons, particles that do not participate in strong interactions) were filtered out in a shield consisting of a long mound of earth. The result was a pure beam of neutrinos and antineutrinos, which were used for the neutral-weak-current experiment.



NEUTRINO INTERACTIONS WERE OBSERVED at Fermilab in an apparatus consisting of a sequence of calorimeters, spark chambers and magnets. The calorimeters were filled with a liquid that scintillates when it is struck by a charged particle; the scintillations were detected by photomultiplier tubes, 12 to each calorimeter. Tracks of charged particles were traced by the spark chambers.



CHARGED-CURRENT EVENT left a characteristic track (*color*) in the spark chambers of the apparatus. The multiple tracks in the left side of the apparatus were made by the hadrons that were produced by the event; the single track at far right was made by a muon.



NEUTRAL-CURRENT EVENT is characterized by a track in which the muon is absent. The distance between the first spark chamber and the last in the apparatus is some 50 feet, but in photographs the tracks are brought close together by a series of mirrors.

Energy from

How Exxon is putting America's newest fuel to work for you today.

Energy from the atom is one of our country's most promising sources of power. Today, nuclear power supplies 7% of America's electricity.

Ten years from now, nuclear energy should be generating about one-third of our electrical power. By the year 2000, over half of it will probably be supplied by nuclear energy.

Where uranium is found.

Uranium ore is the raw material for nuclear fuel. So far, the known commercial deposits of uranium that have been found in America are mostly in the western states.

One of the largest deposits discovered in recent years is the site of Exxon's Highland Mine near Casper, Wyoming. Here, we are mining 2,000 tons of uranium ore a day.

Once the ore is mined, we extract the uranium from the ore and turn it into a fine powder known as "yellow cake."

It is then converted into a gas and sent to plants run by the Atomic Energy Commission. Here, the fissionable, or energy-producing, component is concentrated. This process is called enrichment.

The fuel for a reactor.

We make the enriched uranium into fuel for nuclear reactors at Exxon Nuclear Company's plant in Richland, Washington.

First, we convert it into pellets about twice the size of an aspirin tablet. After numerous checks for precise weight, size, shape and uniformity, the pellets are sealed in zirconium or stainless steel tubes. The tubes are then grouped together in a precise geometric pattern to form a fuel bundle or assembly.

When these assemblies have passed several thousand checks and inspections, they are shipped to nuclear electric-power generating plants. It takes about 600 assemblies to make up the core or heat source of a nuclear reactor. And such a reactor can generate enough electricity to run a city the size of Boston.

Making nuclear power plentiful.

Exxon is currently working on several projects to expand the availa-

A single tiny pellet of enriched uranium can generate as much energy as 150 gallons of fuel oil. Thousands of these pellets go into a nuclear fuel bundle. A single bundle can generate enough electricity to run 4,000 average-sized homes.

This is Exxon's uranium mine in Wyoming. When the surface mining is completed, most of the area will be filled in, contoured and planted to match the rest of the countryside. Grass, wildflowers and shrubs that Exxon planted are already growing on earth that has been excavated.

Uranium.

An engineer at Exxon Nuclear Company shows a young visitor how a fuel bundle is assembled. Dozens of zirconium or stainless steel tubes are filled with uranium pellets, then grouped together to form a nuclear fuel bundle. As many as 600 bundles make up the core or heat source of a reactor.

bility of nuclear energy.

One involves the use of plutonium to conserve uranium ore. Plutonium is a by-product of today's commercial nuclear reactor. By substituting plutonium for uranium, the same amount of nuclear fuel can be produced while saving uranium.

In another project, Exxon Nuclear Company is looking into the technology and economics of uranium enrichment.

At the moment all enrichment in the U.S. is done by the Atomic Energy Commission. As the demand for enriched uranium grows, private industry may take on part of this job.

What's delaying nuclear power?

As we pointed out, today only 7% of America's electrical power comes from nuclear plants. The figure is small because the construction of new plants has been delayed. Environmental concerns, labor shortages, technical problems, late deliveries of specialized equipment and changes in regulatory procedures have all taken their toll.

As these matters are resolved, the nation will begin to depend more and more on nuclear energy to produce the electricity to run its factories, heat its homes and cook its meals.

Not an easy job.

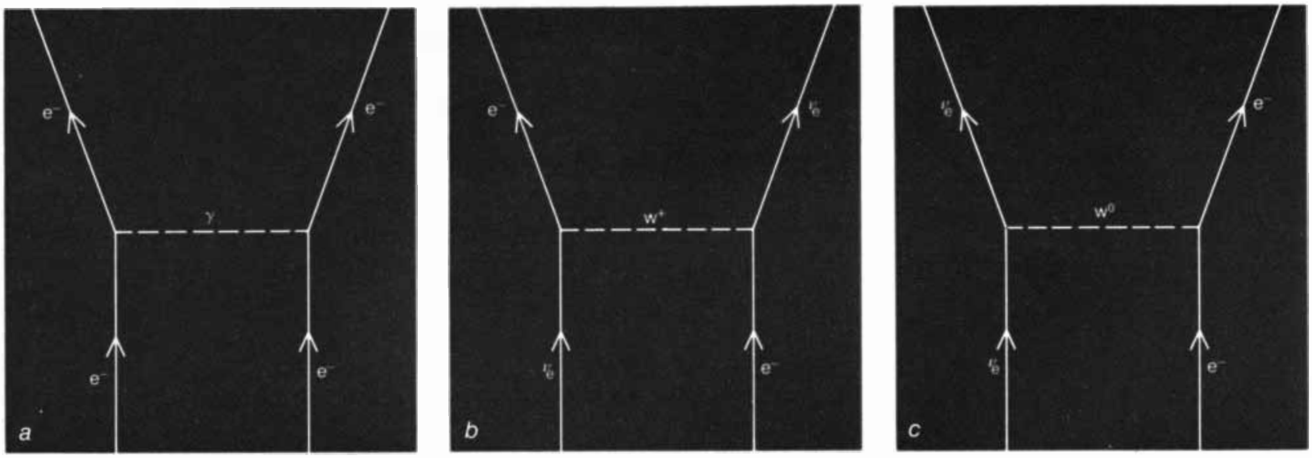
Making America more self-sufficient as an energy-producing nation through the use of nuclear power will not be an easy job. But Exxon believes we need to reach this goal as soon as possible.

EXXON

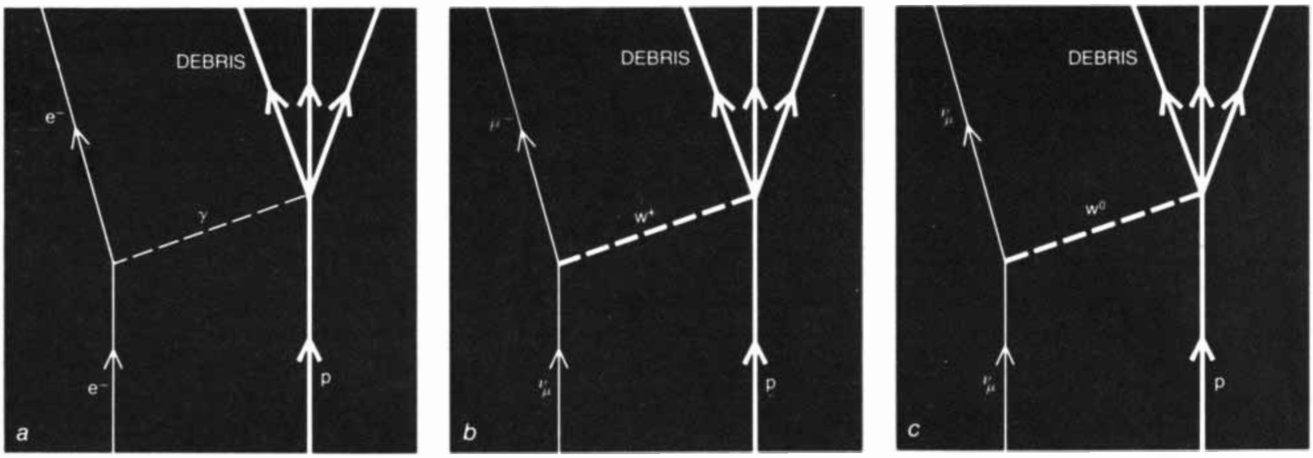
fore Maxwell's time that there must be a connection between electricity and magnetism, there has been speculation for many years concerning the existence of some underlying connection between the weak force and the electromagnetic force. The discovery that there was a connection between these two forces would serve to unify the description of the two kinds of interaction and would undoubtedly represent a major advance in our understanding of the basic forces of nature. How, then, are the two forces alike, and how do they differ?

As it happens, the dissimilarities between the weak force and the electromagnetic force are much more obvious than the similarities. In all studies of the weak interaction until about a year ago the interaction was observed to proceed through a charged weak current. For example, when a neutrino collides with an electron, their interaction (known to particle physicists as a scattering event) appears to result in the incoming neutrino's turning into an outgoing electron and the incoming electron's turning into an outgoing neutrino. Since the neutrino is neutral and the electron is charged, charge is exchanged in the collision. In an electromagnetic interaction, on the other hand, the charge of the incoming electron does not change in the course of the collision. Thus the first difference between the weak force and the electromagnetic force is that when the two kinds of force are viewed as current-current interactions, the currents in the weak interaction have always appeared charged, whereas the currents in the electromagnetic interaction have always appeared neutral.

The concept of currents in elementary-particle interactions can be extended to the interactions of electrons and neutrinos with hadrons: the class of comparatively heavy particles, such as the proton, that interact by means of the strong force. A typical high-energy collision between a neutrino and a proton gives rise to a shower of other hadrons. In such a charged-weak-current interaction the incident neutrino always turns into a muon. A muon is not a hadron but one of the leptons: the class of comparatively light or massless particles that do not respond to the strong force.



FORCES BETWEEN PARTICLES are analogous to forces between two electric currents carried in wires. These Feynman diagrams show examples of the electromagnetic interaction (*a*) and two kinds of weak interaction (*b* and *c*). In the electromagnetic interaction between two electrons (e^-) a photon (γ) is exchanged. In the weak interaction between a neutrino of the electron type (ν_e) and an electron, a hypothetical particle, the intermediate vector boson, or *W* particle, is exchanged. If a positive *W* particle is exchanged (*b*), the interaction is a charged-current process; if a negative *W* particle is exchanged (*c*), it is a neutral-current process.



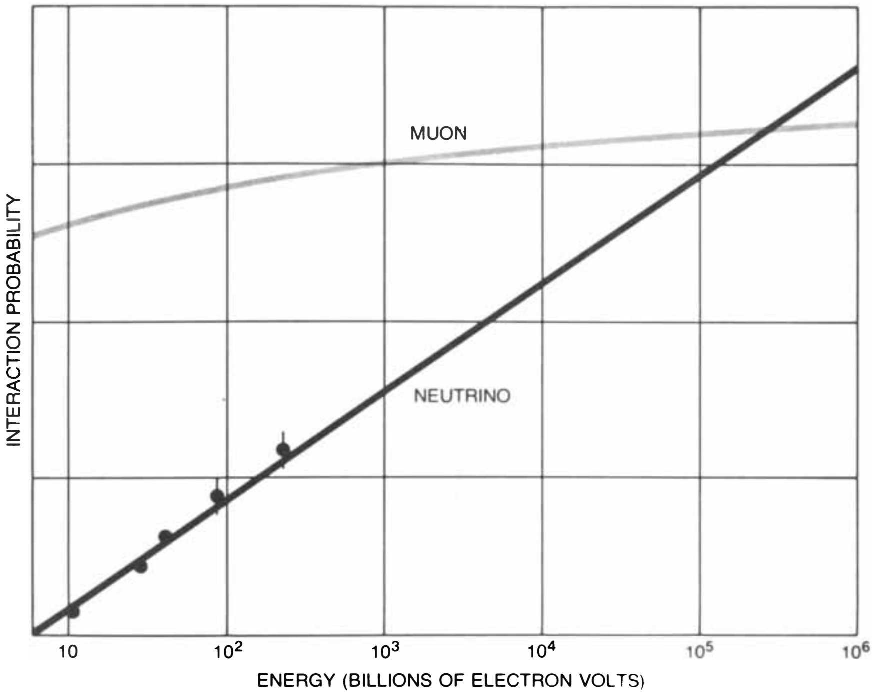
CURRENT BEHAVIOR is also exhibited when leptons interact with hadrons. When a proton (*p*) is struck by a high-energy electron (*a*), the proton is disrupted into a debris of other particles but the electron is present in the final state. This is a neutral-current electromagnetic interaction. When proton is struck by a neutrino of the muon type (ν_μ), either a negative muon (μ^-) or a muon neutrino is present in the final state (*b* and *c*). First interaction is a charged-current one; second is a neutral-current one.

Another difference between the weak force and the electromagnetic force is the vast disparity in strength between the two. The weak force manifested in such processes as the radioactive decay of an atomic nucleus is roughly 10 billion times weaker than the electromagnetic forces present in the same nucleus. The time it takes for a nucleus to decay by the intervention of a force is related to the force's strength. Thus whereas it takes seconds, minutes, days or even years for a nucleus to decay radioactively (depending on the energy available for the transition), a comparable electromagnetic process would take place some 10 billion times faster.

(It comes as a surprise, therefore, that in recent experiments at Fermilab the strength of the weak interaction has been observed to grow as the energy of the transition is increased. In collisions involving neutrinos, for example, the probability that a weak interaction will occur grows with increasing energy [*see illustration at right*]. The strength of the electromagnetic interaction also increases slightly with increasing energy but at a much lower rate. If this trend continues at still higher energies, the strength of the two forces will become the same, and one of the fundamental differences observed between the two forces at low energies will have disappeared. At extremely high energies the weak force may even become stronger than the electromagnetic force!)

Still another striking difference between the weak force and the electromagnetic force is that weak interactions violate parity, whereas electromagnetic interactions do not. The conservation of parity, once thought to be an inviolable principle of physics, states in essence that nature does not distinguish between a given reaction and its mirror image. Parity violation is now recognized as a standard feature of all weak interactions that are mediated by charged currents.

A collision involving a high-energy neutrino provides a graphic example of a weak interaction that violates parity. In nature neutrinos always spin in a characteristic direction: let us say left-handed for neutrinos and right-handed for antineutrinos. In particle collisions that conserve parity the interaction probability is expected to be independent of the handedness of the spin of the participating particles. In other words, at the subatomic level the direction of a particle's spin should have no effect on its overall scattering properties. For electromagnetic interactions the following principle is observed: The interaction probability for



PROBABILITY OF INTERACTION between a muon or a neutrino and a hadron increases with energy. The interaction between the muon and the hadron is electromagnetic; the interaction between the neutrino and the hadron is weak. Where the two curves intersect the neutrino interaction may become more probable than the muon interaction. This is another way of saying that weak interaction may become stronger than electromagnetic.

left-handed and right-handed electrons or muons is the same. Hence electromagnetic interactions involving incident beams of electrons or muons are said to conserve parity.

For scattering events involving beams of neutrinos or antineutrinos, in contrast, the interaction probability for charged currents is observed to depend strongly on the handedness of the incident particle. In fact, the ratio of the interaction probabilities for antineutrinos and neutrinos approaches 1 : 3, which is the value expected for maximum parity violation. Such charged-weak-current interactions have been found to violate parity at energies ranging from a low of a .0001 electron volt to a high of 100 billion electron volts. Over approximately the same range of energies the electromagnetic interaction has never been found to violate parity.

One final difference between the weak force and the electromagnetic force is the effective distance over which each force acts. The strength of the electromagnetic force, like that of the gravitational force, falls off inversely as the distance squared. Hence even at very large distances this force continues to act. The long-range nature of the electromagnetic force arises from the fact that the force is transmitted by a photon, a particle with zero mass. The weak force, on the other hand, appears to act only over very short distances; sometimes it is even imagined as acting at a point. Thus the collision of two weakly interacting particles resembles the collision of two billiard balls. As we have mentioned, assuming that the weak force is transmitted by the exchange of a particle, the extremely short range of the force leads to the expectation that the exchanged particle is very massive.

Now, what about the similarities between the weak force and the electromagnetic force? One of the most striking is that the interaction probabilities for both forces follow the same "current times current" rule. In fact, when Enrico Fermi formulated the first model of the weak interaction in 1933, he did so by analogy with the electromagnetic interaction, and his model has proved remarkably successful in describing low-energy weak interactions to this day. Another similarity lies in the universality of the two interactions in comparison with the nuclear (or strong) interaction. All particles, including hadrons, participate in the weak and electromagnetic interactions but only hadrons are affected by the nuclear force.

Thus on the one hand the form and

universality of the two interactions are similar but on the other the strength, the charge of the currents, the range of the force and the relation to parity are completely different. A particularly puzzling aspect of the weak interaction has always seemed to be the heavy dependence of the weak currents on the electric charge of the particles. Unless there were a deep connection between the two forces, it is difficult to see why the weak interaction should depend so critically on electric charge.

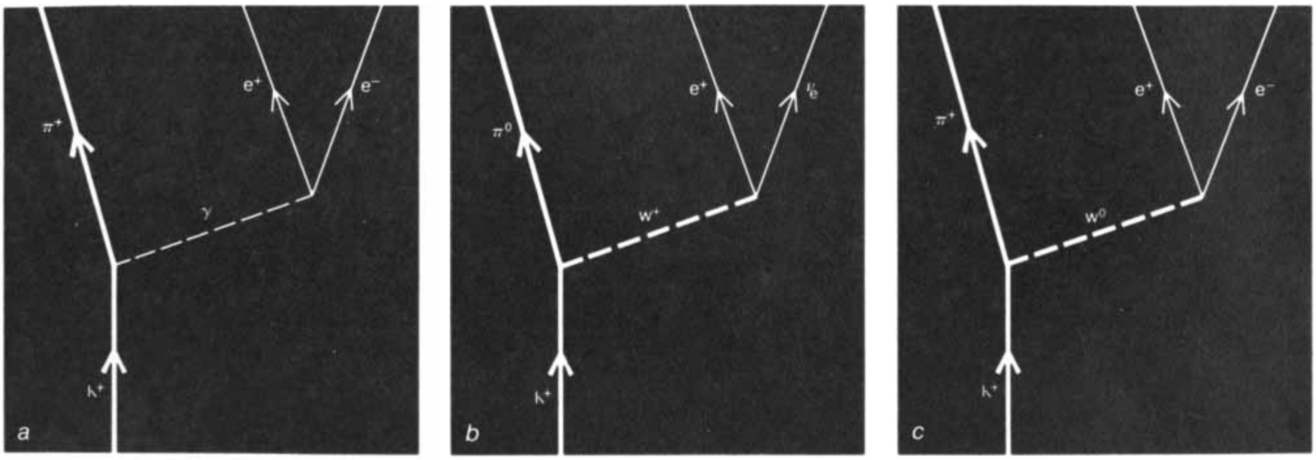
At first the striking differences between the weak interaction and the electromagnetic interaction would seem to outweigh the similarities, precluding any common origin for the two forces. Nonetheless, several theoretical physicists have suggested models that account for both interactions in terms of a single underlying interaction. The first concrete proposal along these lines was made in 1957 by Julian Schwinger of Harvard University. Later Steven Weinberg, who was then at the Massachusetts Institute of Technology, and Abdus Salam of the International Center for Theoretical Physics in Trieste independently put forward a new version of the unified field theory, based on the principle of "gauge invariance," in which an additional unobserved symmetry is assumed to exist between the weak force and the electromagnetic force. Almost all of the proposed models require that the weak interaction operate through a neutral current as well as through a charged one. According to the models this hypothetical neutral weak current would be analogous to the observed neutral current in the electromagnetic case and hence would provide a direct link between the two kinds of interaction.

At the time these models were formulated an experimental search for neutral weak currents appeared to be out of the question, since the probability that a nucleus would become involved in a weak interaction through the intervention of a neutral current was even in the most favorable case many orders of magnitude smaller than the probability of a corresponding electromagnetic interaction. Thus the question of the existence of neutral weak currents was not immediately put to experimental test. In order to separate the two kinds of neutral current one must select a nuclear process where the electromagnetic force does not interfere. Although such processes are known to nuclear physics, they had never been much studied.

The earliest meaningful test for the existence of a neutral weak current focused on the decay of certain particles distinguished by a property known to high-energy physics as strangeness. It had been established previously that strangeness is conserved in interactions that proceed by the intervention of the strong force or the electromagnetic force. In other words, the participating particles have the same degree of strangeness before the interaction and after it. In the course of a charged-weak-current interaction, in contrast, it had been observed that the strangeness of the participating particles could either change or stay the same.

The detailed study of charged-weak-current interactions had revealed a universality that connects the decay of strange particles, nonstrange particles and even of leptons such as the muon. Indeed, it was this universal nature of the weak interaction that originally made it possible to classify such a diverse group of elementary-particle collisions and nuclear transitions under the single heading of the weak interaction. This universality is clearly an important property of charged-weak-current interactions, and it is also, as we have noted, an important property of the electromagnetic interactions.

The search for neutral currents in strange-particle decays was first carried out with positively charged *K* mesons. If neutral weak currents exist, and if they obey the same universality rule that charged weak currents are known to obey, then it was expected that the neutral currents would be readily observed in the decay of the *K* mesons. An experimental search was undertaken in 1962 by one of us (Cline) and his colleagues at the University of Wisconsin and the Lawrence Berkeley Laboratory of the University of California. The search established that the probability for a neutral-weak-current interaction was less than a hundred thousandth of the probability of the corresponding charged-weak-current interaction. It was clear that weak neutral currents were either absent or extremely improbable in positive-*K*-meson decays. Subsequent detailed searches based on other strange-particle decays were equally fruitless. These results appeared to undermine the hoped-for connection between weak and electromagnetic interactions.

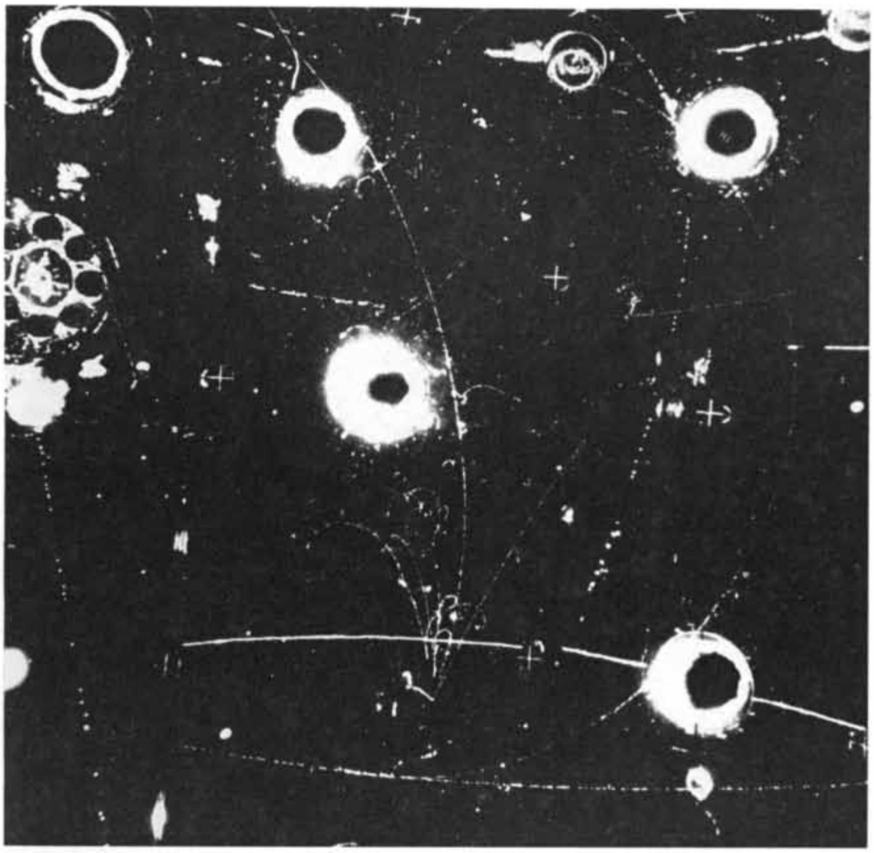


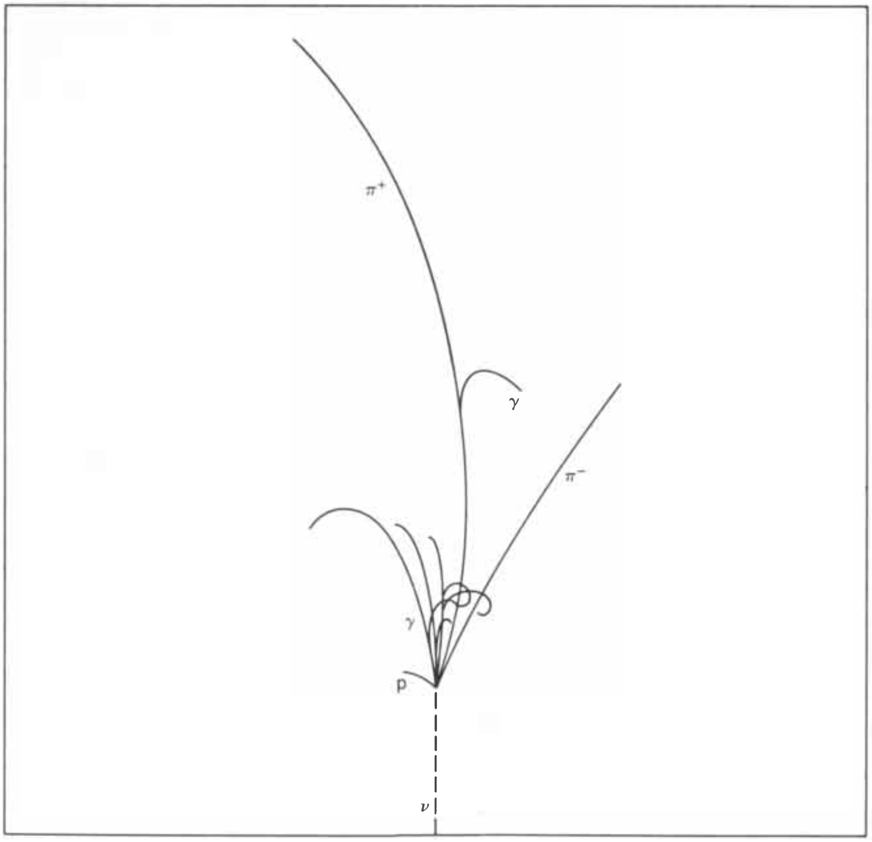
POSITIVE K MESON (K^+) has the quantum property known as strangeness. In electromagnetic interactions strangeness is conserved, that is, it cannot change, and so the decay of the particle into a positive pion (π^+) and an electron-positron pair (*a*) is forbidden because none of those particles have the property of strangeness. Strangeness can change, however, in weak interactions (*b* and *c*). The decay of the positive *K* meson into a neutral pion (π^0), a positron (e^+) and an electron neutrino (ν_e) is a charged-weak-current event. The decay of the particle into a positive pion (π^+) and an electron-positron pair is a neutral-weak-current event. For unknown reasons the neutral-weak-current event is 100,000 times less likely to take place than the charged-weak-current event is.

At approximately the same time that detailed searches were under way for neutral weak currents in strange-particle decays the first generation of high-energy neutrino-beam experiments had been undertaken at the Brookhaven National Laboratory and at the European Organization for Nuclear Research (CERN). Intense beams of neutrinos and antineutrinos are produced as by-products at large proton accelerators. Both the neutrino and its antiparticle interact with matter only by means of the weak force. Since the time of the earliest neutrino experiments at Brookhaven it has been known that there are actually two kinds of neutrino: the muon neutrino, which interacts with matter to produce a muon, and the electron neutrino, which interacts with matter to produce an electron. At proton accelerators muon neutrinos are produced about 100 times more copiously than electron neutrinos.

Neutrino interactions provide another way of searching for weak neutral currents without the overriding influence of the electromagnetic interactions, and also without the complications of strangeness. Basically these experiments are quite simple. A charged-current interaction of a muon neutrino always results in a charged muon in the final debris of the collision. A neutral-current interaction always produces a neutrino instead. All that is required of the experimenter is to separate one reaction from another. Of course it is not that easy, and in fact the first round of neutrino experiments at Brookhaven and CERN failed to detect any neutral-current events. In some cases it was determined that neutral currents are very improbable. Combined with the null results obtained from strange-particle decays, these results appeared to confirm the universal absence of neutral weak currents. As a result, by the end of the 1960's it was taken for granted that neutral weak currents were either absent in nature or extremely rare.

Within the past year, however, two experiments utilizing rather different techniques have provided evidence that neutrinos can interact with matter to form new neutrinos, and at a rate that is a substantial fraction of the rate for neutrino scattering to form charged leptons. One of the experiments was done at CERN with a liquid freon (CF_3Br) bubble chamber, which served as both the target and the detector of the neutrino interactions. The impact of high-energy protons from the CERN proton synchrotron on a metal target yielded, among other particles, charged mesons that could be focused magnetically and then allowed to decay in flight. These decays produced either a





CANDIDATE FOR NEUTRAL-CURRENT EVENT appears in photograph made with Gargamelle, the large bubble chamber at the European Organization for Nuclear Research (CERN). As the diagram shows, a neutrino (which leaves no track in the chamber) first strikes a proton. The main products are a positive pion, a negative pion and a neutral pion (which also leaves no track). All these particles are hadrons. The absence of a charged lepton such as a negative muon suggests that the interaction is a neutral-weak-current one.

beam of neutrinos or a beam of antineutrinos, depending on the charge of the incident particles.

The analysis was based on some 80,000 bubble-chamber photographs made with the neutrino beam and more than 200,000 photographs made with the antineutrino beam. A total of 102 neutrino events were found without muons in the final state compared to 428 neutrino events with muons; the corresponding numbers for antineutrino events were 64 and 148. An example of a muonless event is shown in the bubble-chamber photograph on the preceding page.

The other experiment was conducted at Fermilab by a group of physicists from Harvard University, the University of Pennsylvania, the University of Wisconsin and Fermilab. In this experiment the impact of high-energy protons on a target generated the secondary hadrons that decayed to provide the incident neutrino and antineutrino beams.

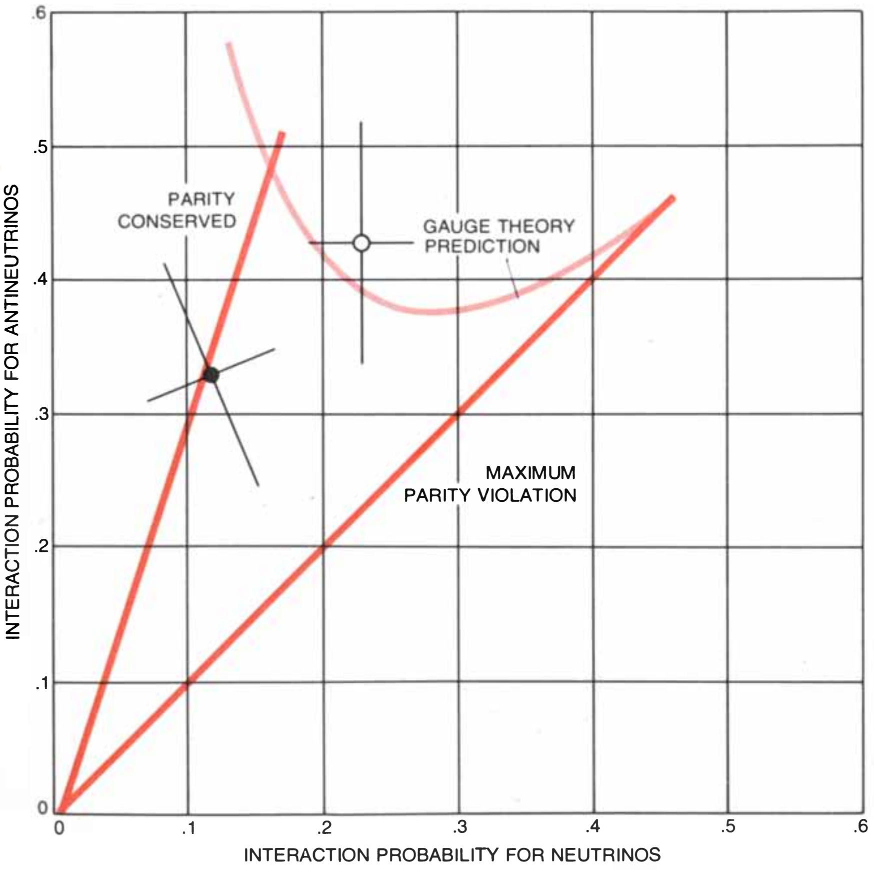
The target-detector of the neutrino interactions was a large ionization calorimeter consisting of 16 segments, each of which was filled with a liquid that scintillates when a high-energy particle strikes it. Each segment was monitored by 12 photomultiplier tubes, the summed output of which measured the ionization energy deposited in the liquid scintillator by the products of a neutrino or antineutrino interaction. Most of the energy in the cascade of particles from such an interaction was contained within the total volume of the ionization calorimeter. Interspersed between every four segments of the calorimeter were optical spark chambers, which displayed the cascade of secondary particles visually. Immediately downstream from the ionization calorimeter was a muon detector consisting of four large scintillation counters and toroidal magnets with accompanying optical spark chambers, to identify muons and to measure the momentum of a subsample of the muons.

The experiment was triggered by the arrival of particles with an energy greater than a certain minimum value. Following a trigger signal the electrical outputs of all segments of the ionization calorimeter were recorded individually, as were the outputs of the muon identifier counters, and all the spark chambers were fired. Typical neutrino events observed in the detector have a muon that penetrates the muon detector and is deflected in the magnet. The search for neutral currents was carried out by looking for events lacking a muon in the final state. Any such event is very probably an example of a neutral-current interaction.

A total of 4,181 events were triggered in a number of separate runs with various neutrino-antineutrino beam mixtures, yielding 991 useful events, of which 220 showed no final-state muon and 771 exhibited a muon. The principal correction to the raw data of this experiment was for wide-angle muons that miss the muon identifiers.

The primary conclusion from these experiments is that, unlike earlier neutrino-scattering experiments, they showed a significant number of neutrino-induced (and antineutrino-induced) events lacking charged muons in the final state. It is important to note that this positive signal was obtained in both experiments even though the experimental methods were quite different. Moreover, the average energy of the events observed in the two experiments was quite different: in the CERN experiment it was about three billion electron volts and in the Fermilab experiments it was roughly 40 billion electron volts. Recently three additional experiments undertaken by different groups working at the Argonne National Laboratory, Brookhaven and Fermilab were analyzed. The results from these groups confirm the effect.

The existence of neutral weak currents is probably the manifestation of an intimate relation between weak and electromagnetic interactions, but at present the form of this relation is not clear. It is possible that the neutral weak currents share other properties with the electromagnetic interaction, unlike the charged weak currents. For example, the neutral weak current may conserve parity, as the electromagnetic interaction does, providing a more direct link between the two. In that case there would be a distinct difference between the intrinsic properties of the charged and neutral

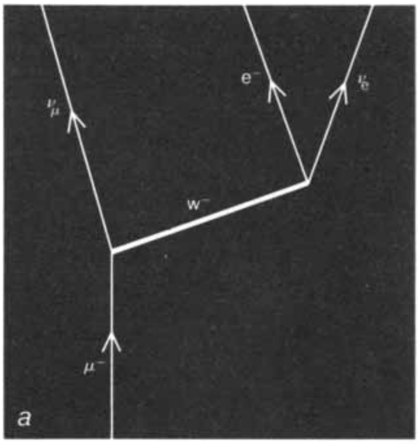


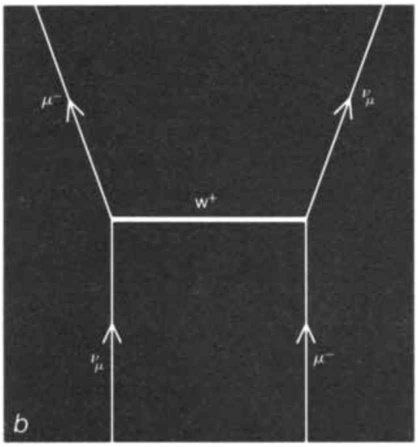
THEORETICAL PREDICTIONS are indicated by curves on a graph that plots the probability that neutrinos will interact with hadrons through neutral-weak-current interactions against the probability that antineutrinos will so interact. The result from the Fermi National Accelerator Laboratory experiment is indicated by the black dot; the result from the CERN experiment, by the white dot. The prediction of gauge theory is that the experimental result will lie above the bent curve toward the top of the graph. Since the two results straddle the curve, they are inadequate to test the correctness of the gauge theory. Since they are closer to the straight curve at left than the straight curve at right, however, they do suggest that parity is not maximally violated in such neutral-weak-current interactions.

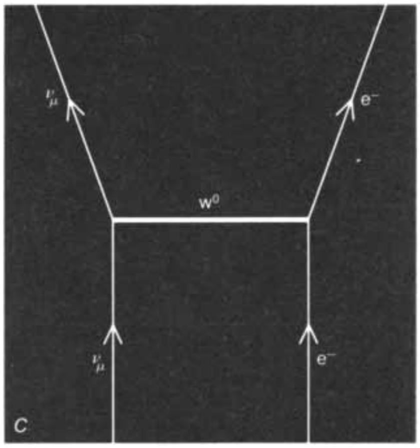
weak currents. A direct test of the parity-violating or parity-conserving nature of the neutral-current interaction can be made by comparing the interaction probabilities for neutrinos and antineutrinos. The results of the CERN and Fermilab experiments suggest that the neutral-current interaction is probably not maximally parity-violating, in contrast to the charged-current interaction, which is. Better experimental measurements that are currently under way will soon be able to test for this possibility.

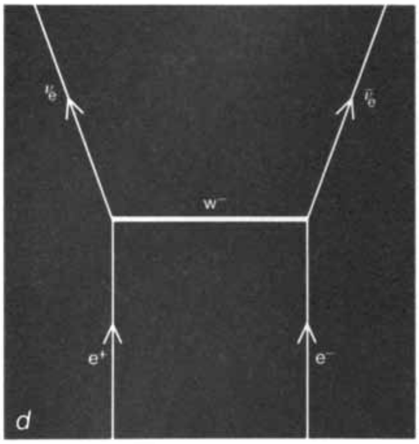
Another possibility is that the neutral weak current is a natural component of the gauge theories that have been proposed to unify weak and electromagnetic interactions. Continued experimental study of neutral-current reactions over the next few years will probably make the details of the unification clearer. Many profound questions remain to be answered, such as why weak interactions violate parity and why the range of force in the two cases is so different. Nonetheless, the discovery of neutral currents has generated great interest precisely because of the likelihood that such unification will eventually emerge, whatever the exact version turns out to be. In any event neutral currents will also undoubtedly serve as a tool for the further exploration of many other properties of elementary particles.

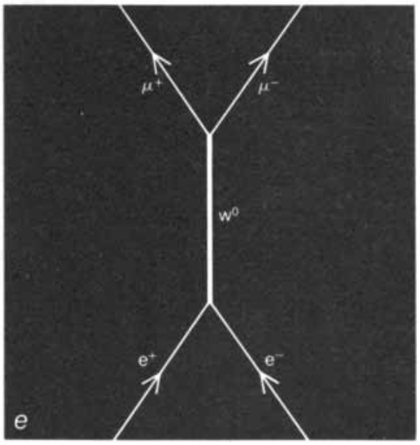
Along with the possible unification of weak and electromagnetic interactions the study of weak-interaction processes has played an important role in acquiring an understanding of the nature of strong interactions. This has come about because the study of weak interactions has been largely confined to the study of the weak decays of strongly interacting particles. An unavoidable complication in these studies is the nature of

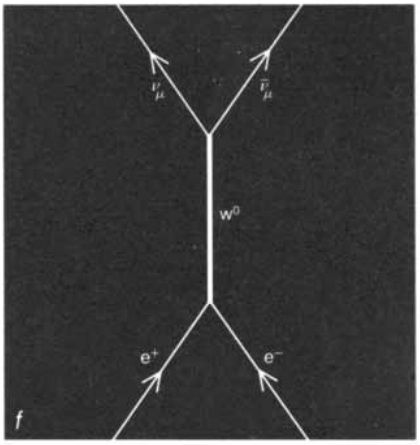


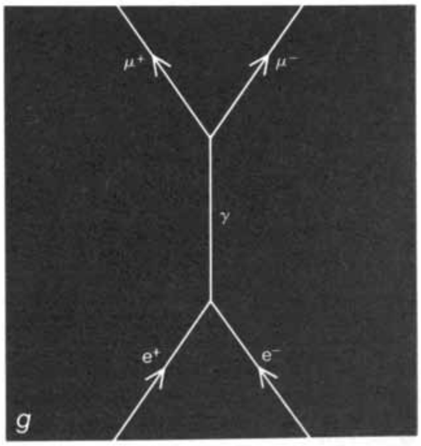












EXAMPLES OF INTERACTIONS BETWEEN LEPTONS are outlined in Feynman diagrams. One example (*a*) is the decay of a muon (μ^-) into a muon neutrino (ν_μ), an electron (e^-) and an electron neutrino (ν_e). Another example (*b* and *c*) is the scattering of neutrinos by muons and electrons. A third example (*d*, *e* and *f*) is electron-positron scattering through the agency of the weak interaction. A fourth example (*g*) is electron-positron scattering through the agency of the electromagnetic interaction. Before the detection of neutral weak currents some of these processes (*c*, *e* and *f*) were believed to be absent. Their existence now makes it possible to investigate the weak interactions between leptons.

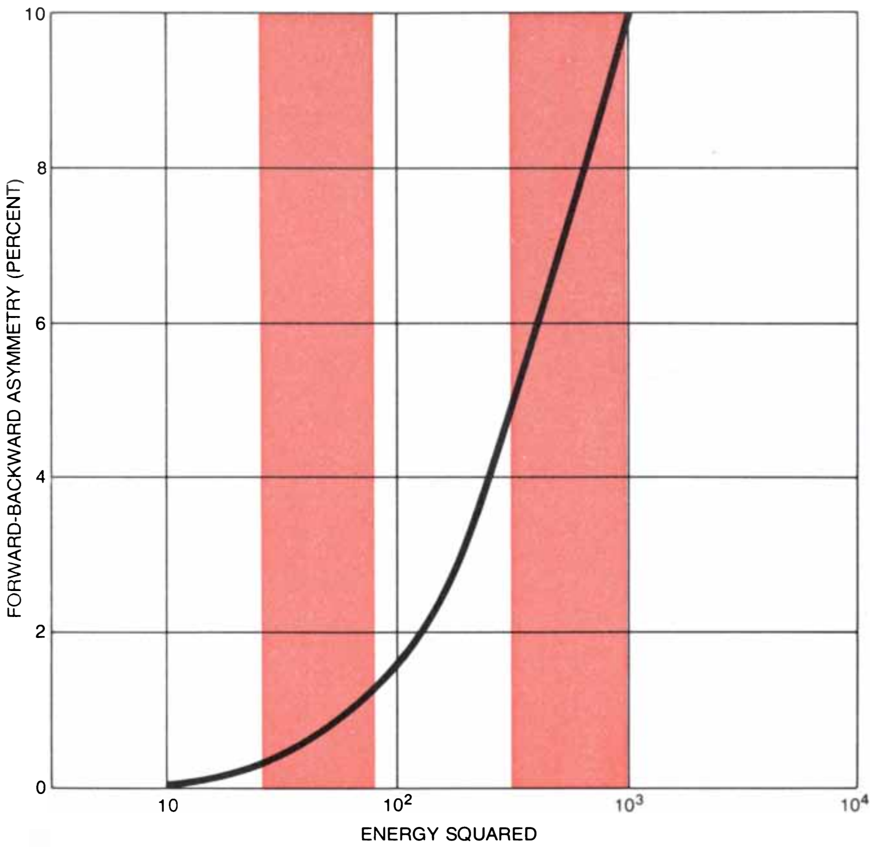
the strong interaction, which plays a central role in the binding and composition of such particles. It is ironic that through the mediation of the weak interaction new knowledge about the properties of the strong interaction might be gained. The neutral-weak-current interaction will undoubtedly extend this knowledge. The most direct manifestation of the strong interaction in weak decays comes in the form of certain selection rules that have been discovered in the study of charged-current processes. One such rule states (in simplified terms) that when the strangeness changes by one unit the charge must also change by one unit. Another rule that has been observed requires that the total change in strangeness in the reaction be no greater than one. Both of these rules have been observed with charged-current reactions; they can now be tested with neutral currents. Although the reasons for such selection rules are unknown, there can be little doubt that they are related to the fundamental composition of hadrons and that they will play an important role in the ultimate explanation of these particles.

Neutral weak currents will also play an important role in the study of purely leptonic weak interactions. Although nature abounds in leptons, the study of weak interactions in which only leptons participate has barely begun. As leptons are currently understood, they are less complicated than hadrons, being almost pointlike particles. The classic example of a purely leptonic reaction is when a muon decays to form an electron and two neutrinos, one of the muon type and the other of the electron type. With the advent of high-energy, high-intensity neutrino beams a new leptonic reaction is available for study. This process occurs through the charged-current interaction in which a muon neutrino is scattered by the static electric field of an atomic nucleus to become two muons and a muon neutrino.

The existence of neutral currents provides other purely leptonic reactions for examination, considerably extending the range of study of leptonic processes. For example, the neutral-current interaction of a muon neutrino and an electron may possibly exist. There is already some preliminary experimental evidence for such a process from the CERN experiments. Another exciting prospect is the detection of weak interactions in high-energy electron-positron collisions. The weak interaction of an electron and a positron to form a neutrino and an antineutrino, both of the electron type, surely exists, but it cannot be detected by any known means because there are no final products other than neutrinos. On the other hand, the reaction of an electron and a positron to form two muons, one positive and one negative, offers an excellent opportunity to study neutral-weak-current interactions between leptons. This reaction can proceed either by a weak interaction or by the electromagnetic interaction.

Here the weak interaction is expected to interfere with the electromagnetic interaction in a way that is related to the parity-violating nature of the weak interaction. The result is a forward-backward asymmetry in the production of positive and negative muons. Thus an excess of positive muons will be made that will go forward in the direction of the incoming positron. Observation of this asymmetry would be additional evidence for neutral weak currents. The asymmetry is expected to grow with the strength of the weak interaction [*see illustration at left*]. A group of physicists from the University of Pennsylvania and the University of Wisconsin are searching for this asymmetry at the electron-positron colliding-beam machine ("SPEAR") at the Stanford Linear Accelerator Center. Since weak interactions grow in strength with energy, the asymmetry is expected to grow to a relatively large value at the higher energies that will become available in the next generation of machines in which electrons and positrons are made to collide in counterrotating beams. One such machine, designated PEP (for positron-electron, Phase 1), has been proposed by Stanford University and the University of California at Berkeley. There is little doubt that the study of weak interactions between leptons colliding in high-energy beams will be a giant step forward in the study of weak interactions.

As we have indicated, the existence



RATIO OF PARTICLES SCATTERED FORWARD AND BACKWARD in certain experiments should indicate an interference between neutral electromagnetic currents and neutral weak currents. In such experiments electrons and positrons in counterrotating accelerator beams collide and give rise to positive and negative muons. The curve in this graph shows the predicted degree of asymmetry, if the interference exists, between the number of muons scattered in the forward direction and the number scattered in the backward direction. The colored band at left shows the range of energies in an experiment now being set up in a colliding-beam facility ("SPEAR") at the Stanford Linear Accelerator Center. Colored band at right shows the range of energies that would be attained in a new facility ("PEP") proposed by physicists at Stanford and the University of California at Berkeley.

HYPOTHETICAL PROCESSES		CHANGE OF PROPERTIES IN THE INTERACTION		DECAY PROBABILITY
$\delta^0 \rightarrow \pi^+\pi^-\pi^0\pi^0$	STRONG INTERACTION	NO CHANGE OF STRANGENESS	CONSERVES PARITY	VERY PROBABLE
$\delta^0 \rightarrow \gamma\ \gamma$	ELECTROMAGNETIC INTERACTION	NO CHANGE OF STRANGENESS	CONSERVES PARITY	ONE IN 10^4 TIMES
$\delta^0 \rightarrow \pi^+ e^- \bar{\nu}_e$	WEAK INTERACTION (CHARGED CURRENT)	NO CHANGE OF STRANGENESS	VIOLATES PARITY	ONE IN 10^{10} TIMES
$\delta^0 \rightarrow \pi^0 \nu_e \bar{\nu}_e$	WEAK INTERACTION (NEUTRAL CURRENT)	NO CHANGE OF STRANGENESS	UNKNOWN	ONE IN 10^{11} TIMES
$\delta^0 \rightarrow \kappa^+ e^- \bar{\nu}_e$	WEAK INTERACTION (CHARGED CURRENT)	CHANGE IN STRANGENESS	VIOLATES PARITY	A FEW IN 10^{12} TIMES
$\delta^0 \rightarrow \kappa^0 \nu_e \bar{\nu}_e$	WEAK INTERACTION (NEUTRAL CURRENT)	CHANGE IN STRANGENESS	UNKNOWN	LESS THAN ONE IN 10^{19} TIMES
$\delta^0 \rightarrow \pi^+\pi^-$	WEAK INTERACTION	NO CHANGE OF STRANGENESS	VIOLATES PARITY AND TIME REVERSAL INVARIANCE	LESS THAN ONE IN 10^{20} TIMES

DECAY OF NEUTRAL DELTA MESON (δ^0), a hypothetical particle, suggests that the existence of neutral weak currents as a subclass of the weak interactions shows a deep connection between the weak force and the electromagnetic force; this and other evidence may also hint at the existence of some new force in addition to the four already known. The hypothetical particle is assumed to have a mass large enough to give rise to groups of particles of lower mass. The symbols of the decay particles are given in the preceding illustrations. The significance of the various interactions, changes of property and probabilities of decay are discussed in the text.

of neutral weak currents is suggestive of a deep connection between the weak interaction and the electromagnetic interaction; it is also possible that a new puzzle has been raised about the nature of all interactions between elementary particles. The customary classification of such interactions into the categories gravitational, electromagnetic, strong and weak is based in part on the relative strengths of those forces. Some idea of those strengths can be gathered from a consideration of the decay of a hypothetical particle called the neutral delta particle. Although this particular particle is not known to exist, there are other known particles (for example the eta meson) with the same properties except for a different rest mass. The hypothetical neutral delta particle is assumed to have a rest mass large enough to make it possible for the particle to be transformed into groups of particles of lower mass: mesons, electrons, neutrinos and so on. The neutral delta particle is also assumed to carry zero units of strangeness. The table above lists the various transformations that the neutral delta particle can make, along with the probability of each interaction. Thus, for example, the probability that the particle will participate in a strong interaction is about 10,000 times greater than the probability that it will be involved in an electromagnetic interaction. The probability that it will participate in a weak interaction, however, is shown to depend heavily on the properties (quantum numbers and symmetry parameters) that change in the transformation and also on whether the transformation involves a charged current or a neutral current.

It is frequently conjectured that the strength of a particular interaction may actually be derived from the number of properties that are changed in the operation of forces. For example, it is sometimes said that the strong interaction may be strong because no properties are changed and that the weak interaction may be weak because many properties are changed. We see from the example of the decay of the neutral delta particle that within the subclass of the weak interactions the different strengths (that is, the different transformation probabilities) depend on the number of properties that change; this may be an example of the conjectured fundamental relation between the strength of interaction and the number of properties that change. To follow this train of thought further, it may be that there are really more than four basic interactions in nature. For example, there is the possibility that a new "superweak" force is needed to provide the transformation of the neutral delta particle that violates the invariance of time reversal. So far too little is known about which properties of elementary particles are actually changed by the neutral weak current to reach a definite conclusion.

The existence of neutral weak currents also promises to help solve some perplexing problems in cosmology. The universe is literally filled with neutrinos. An important source of these neutrinos is electron-positron annihilation, a process thought to have been important in the early evolutionary stages of the universe. These neutrinos were originally suspected of carrying the electron's "tag." By means of the neutral-weak-current interaction, however, electrons and positrons might have annihilated each other to give rise to neutrinos that carry the muon's tag, thus possibly changing the composition of the universe's supply of neutrinos. If, in particular, the muon neutrino carried even a small mass, vast amounts of additional energy could be "hidden" in the universe. The present experimental limits on the mass of the muon neutrino are quite rough and do not rule out this possibility.

Neutrino production plays an important role in the cooling of stars, because the neutrino, once it is manufactured in the interior of a star, escapes, carrying energy with it. The suggestion has recently been made that the escaping neutrinos could, through the neutral-current interaction, exert a large force on the surface of the star and thus blow the surface off. This phenomenon could be what we observe as the explosion of a supernova.

These examples touch only a few of the ramifications that the discovery of neutral currents are likely to have in the areas in nature where the weak interaction plays an important role. Optimism is currently running high that this discovery and other pieces of evidence are leading to a common understanding of the weak interaction and the electromagnetic interaction. There are many remaining obstacles, but it appears that the crucial step has been taken.


IMP X
ACT
SC
SC

COUNTERFEITING IN ROMAN BRITAIN

A runaway inflation in the third century debased the value of coins and drove out of circulation the coins whose value was not debased. The need for exchange then gave rise to widespread counterfeiting

by George C. Boon

Counterfeiting is not generally regarded as being a public service. There have been times in the past, however, when it virtually was. Consider a typical list of coins from a Romano-British site of the third century. The record of the excavation of the fort of Segontium, near Caernarvon in northwestern Wales, includes this notation: " 'Radiates': orthodox, 74; counterfeits, 73." For those who are not familiar with numismatics this cryptic entry will raise at least two questions. First, what kind of coin is a radiate? Second, under what circumstances were counterfeits of this coin so common that they almost equaled the number of orthodox (that is, legitimately minted) radiates from the fort?

An informative answer to the first question requires some familiarity both with Imperial Roman history and with Roman money. To deal with the history first, the third century was a time of trouble in the Roman world. The last decade of the preceding century had seen the worthless son of the great Marcus Aurelius meet a violent end. The imperial purple passed to Septimius Severus, a decent soldier, who ruled effectively until A.D. 211. Thereafter matters deteriorated rapidly through civil and foreign war. One of the successors to Severus, the emperor Decius, was killed by the Goths in 251; another, Valerian, was captured by the Persians in 260. Over the next decade Rome lost all its districts across the Rhine, while the Persians maintained a warlike posture on the eastern frontier. In the west Roman Spain and Britain and all Gaul, from the Rhine to the Atlantic and from the English Channel to the Mediterranean, became independent of the Empire and were ruled by the Gallic emperor Postumus and an ephemeral succession of local pretenders proclaimed by the troops stationed in the region. The "Gallic Empire" was recovered during the time of Aurelian (270-275), but not until the reign of Diocletian (284-305) was the long decline of third-century Rome halted.

Roman money too had a long and complex history. At the risk of oversimplification, a good place to begin is at the beginning, that is, with the basic unit established in the early days of the Roman Republic to serve the dual purpose of a weight and a measure. This unit was the *as;* the Latin word literally means "unity," and its root is the same as that of "ace" in English. An *as* was subdivided into 12 *unciae;* these were "ounces" as units of weight and "inches" as units of measure. Initially the *as* as a unit of weight consisted of a metal casting weighing one Italic pound, or *libra* (the word from which the English abbreviation "lb." is derived). The Italic pound weighed 273 grams; the casting metal was an alloy of copper, tin and lead.

Inflation goes hand in hand with war, and the Punic Wars of the third and second centuries B.C. were no exception. Under the pressure of devaluation the weight of the *as* fell first from a pound, or 12 *unciae,* to two *unciae* and then to one. The decline continued, and by the time of Julius Caesar (100-44 B.C.) an *as* weighed only a third of an ounce. Now, a national coinage that consists only of pennies, which is how one might characterize the lightweight *asses* of the later Republic, is simply not serviceable. Indeed, during the third century B.C. the Romans had added a silver piece to their coinage; it was called the *denarius,* that is, a "piece of 10," and it was marked with the Roman numeral "X" (the root *deni-* comes from *decem,* or 10). The silver *denarius* was originally equal in value to 10 *asses* but was later equal to 16.

One of the major reforms instituted by Caesar Augustus after he took power in 27 B.C. was stabilization of the currency. His monetary decrees in 15 B.C. led to the minting of *denarii* from fine silver at the rate of 84 per new Roman pound (about 325 grams). The silver coins bore a fixed relation to gold: 25 *denarii* were equal in value to one gold *aureus.* These gold pieces were minted at the rate of 42 to the pound. Augustus also minted lesser coins; his *sestertius,* struck from yellow brass, was equal in value to four red copper *asses.*

The slow decline of the Imperial currency began under the emperor Nero (A.D. 54-68) who added between 5 and 10 percent base metal to the silver of his *denarii.* By the time of Septimius Severus the proportion of base metal in the *denarius* had increased to 40 percent. Severus' successor, popularly known as Caracalla because of the Gallic cloak (*caracallus*) he habitually wore, was actually named Marcus Aurelius Antoninus Bassianus. He held power from 211 to 217, and the *denarii* struck at his mints were 50 percent base metal.

Caracalla's half-silver *denarii* obvi-

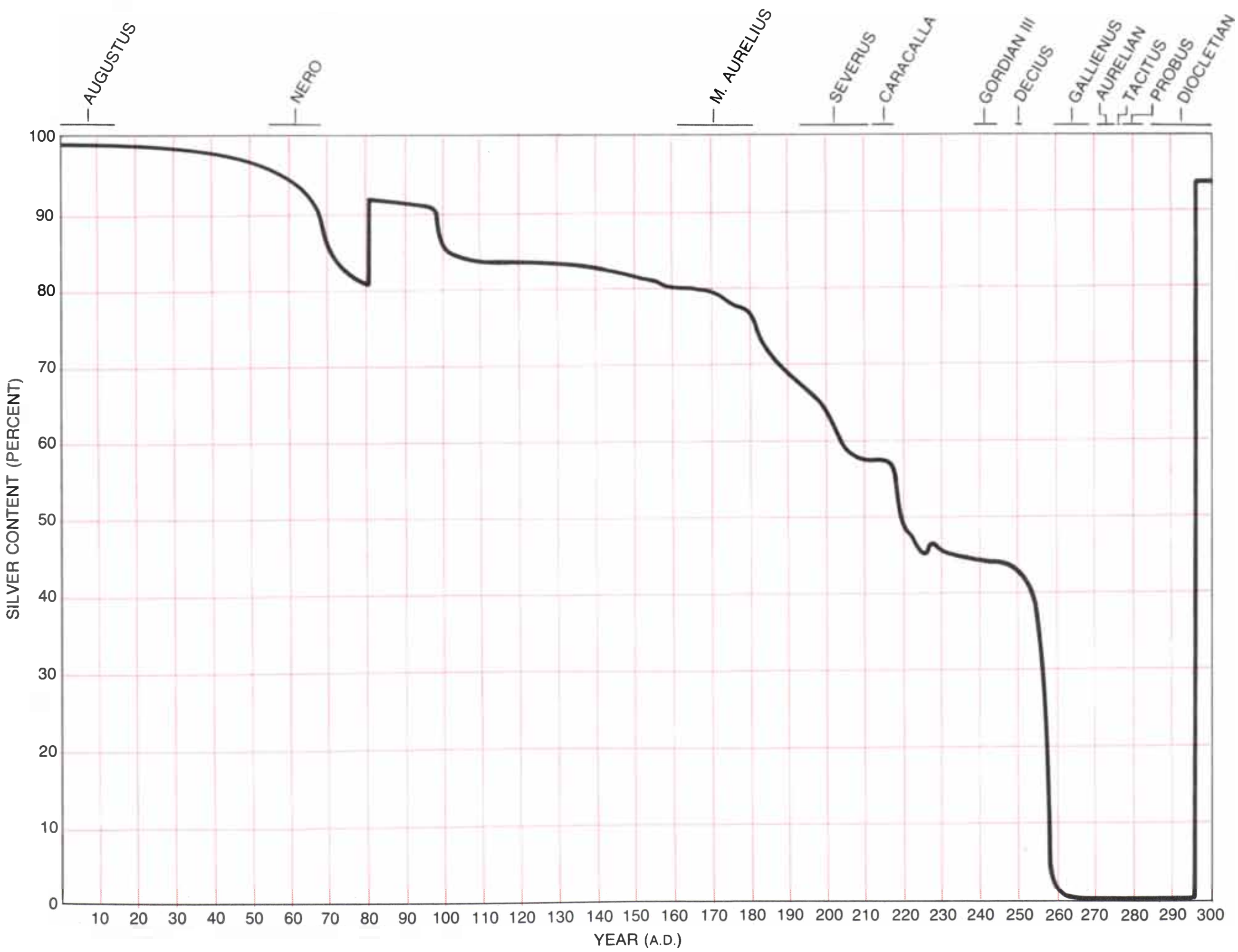
IMPERIAL ROMAN COINAGE, issued by order of Caesar Augustus, is exemplified by the four coins on the opposite page. They are (*from top*) a gold *aureus,* weighing 7.5 grams, a silver *denarius,* weighing 3.9 grams, a yellow brass *sestertius,* weighing 24 grams, and a copper *as,* weighing 10 grams. Four of the copper coins equaled one *sestertius* in value, four of the brass coins equaled one *denarius* and 25 of the silver coins equaled one *aureus.* The coins are enlarged in these photographs; silhouette beside each indicates its true size.

ously could not be exchanged for gold at the same rate that the *denarii* of Augustus had been exchanged some two centuries earlier. It was perhaps with the hope of reestablishing the Augustan relationship of a silver coin to a gold one that Caracalla decreed the minting of a new silver piece. Like Caracalla's *denarius* the coin was still only half silver, but it was minted at the rate of 64 to the pound and so it weighed half again as much as a *denarius:* some five grams compared with 3.3 grams. Called an *antoninianus* after its issuer, it displayed on the obverse a portrait bust of Caracalla, shown wearing a rayed crown [*see illustration on opposite page*]. When the emperors who followed Caracalla minted their own *antoniniani,* they too were shown wearing rayed crowns. It is for this reason that numismatists call all these *antoniniani* radiates.

Caracalla's successors progressively debased the new coins, replacing more and more of the silver with bronze until at last the only silver in them was a surface "wash" that briefly gave the radiates a shiny appearance. By the time of Gordian III (238-244) *denarii* were scarcely ever issued, and radiates had achieved a dominant position in the currency. It is possible to trace the decline of the radiates both in weight and in fineness. For example, by A.D. 250 old *denarii* were being turned into radiates by being overstruck at the mint. This is an admission that the *denarius* had lost a third of its original value.

Thereafter a precipitous decline began; by the mid-260's the coins were only 5 percent silver. By the end of the same decade the Imperial mintings reached an all-time low of barely 1 percent silver. The last rulers of the Gallic Empire issued equally debased coins within a year or two thereafter [*see top illustration on page 125*]. Thus, bearing in mind both the reduced weight of the *antoninianus,* now down to some 2.75 grams, and the similarly reduced weight of the *aureus,* which continued to be very pure gold, one can see that the old Augustan ratio of 25 silver *denarii* to one *aureus* had fallen to a ratio of 1,500 debased radiates to one gold coin.

In practice, however, gold coins were seldom seen. Because their weight was variable and their intrinsic value was high, they were probably regarded more as bullion than as currency. This state of affairs was destined to become enshrined in Diocletian's celebrated Edict on Maximum Prices, issued in 301. The edict set the top value of a pound of gold, either in bar or in coin, at 50,000 *denarii.* The radiate was by then reckoned to be equal to two of the seldom-seen *denarii.* In other words, the three centuries that



PROGRESSIVE DEBASEMENT of Rome's silver coins during the first three centuries of the Christian Era is shown in this graph. A bracket (*top*) marks the period when an emperor whose name is shown held authority. The decline in silver content, slow at first, accelerated late in the second century and hit bottom in the third century. The recovery in A.D. 294, in a reform instituted by the emperor Diocletian, was short-lived, and fine-silver Roman coins did not appear again until the latter half of the fourth century.

separate the rules of Augustus and Diocletian had seen the intrinsic value of these once fine-silver coins fall to a fiftieth of the original level.

Now, metal by itself, including such far from precious alloys as bronze, always has some intrinsic value, and in preindustrial days metal was not cheap. If the time came when, because of debasement, the intrinsic value of a coin exceeded its nominal value, the coin promptly disappeared. It was either hoarded or melted down. To offer a modern example, the U.S. Mint a few years ago began to issue, in place of various fine-silver pieces, cupro-nickel coins of only nominal value. The older coins promptly vanished from circulation. Those who did not hoard the silver sold the coins for their intrinsic value, which substantially exceeded their nominal value. Some saw this as an increase in the value of silver; it can equally well be viewed as a decline in the value of the U.S. dollar.

One can find a rough parallel in Roman practices of the second and third century. The intrinsic ratio of silver to copper was then about 1 : 100. In 184, a time when the *denarius* was struck from an alloy of 25 percent base metal and 75 percent silver, the coin weighed nearly three grams. This meant that its exchange value, at the silver-copper ratio cited, was equal to about 220 grams of bronze. As numismatists say, a silver-alloy *denarius* would "support" some 220 grams of bronze. Now, the recognized exchange of bronze coins for silver was still, as it was in the days of Augustus, four *sestertii* or 16 *asses* per *denarius*. A *sestertius*, however, weighed some 24 grams at this time and an *as* weighed more than 10 grams. It is small wonder, with more profit to be made by minting the *sestertius* than the *as*, that the larger bronze coin begins to predominate in the archaeological finds from this period. Indeed, after the time of Commodus (180-192), few new bronze coins from Rome, the only western mint, even reached Gaul, let alone Britain. In those provinces, which were on the brink of separation, aging second-century bronze coins were retained in use, willy-nilly, down to the time when the ever declining silver content of the radiates could no longer support the equivalent in bronze. That crossover point was reached in about 260, under Postumus, and thereafter bronze coins became subject to serious hoarding.

One example is a hoard of 500 bronze coins found at Leysdown in Kent. Only 14 of the 500 were orthodox third-century coins. Another hoard, found at Bordeaux, consisted of 86 bronze coins. One of them, struck during the rule of Severus (193-211), was worn to such an extent that it must have circulated for many decades before being buried. Even bronze coins that have been unearthed singly at various sites are often seen to be worn nearly flat. An *as* of Hadrian (117-138), found in Coventina's Well at the fort of Carrawburgh on Hadrian's Wall, had weighed 11 grams or so when it was struck but had been reduced by wear to less than four grams before it was thrown into the well for good luck.

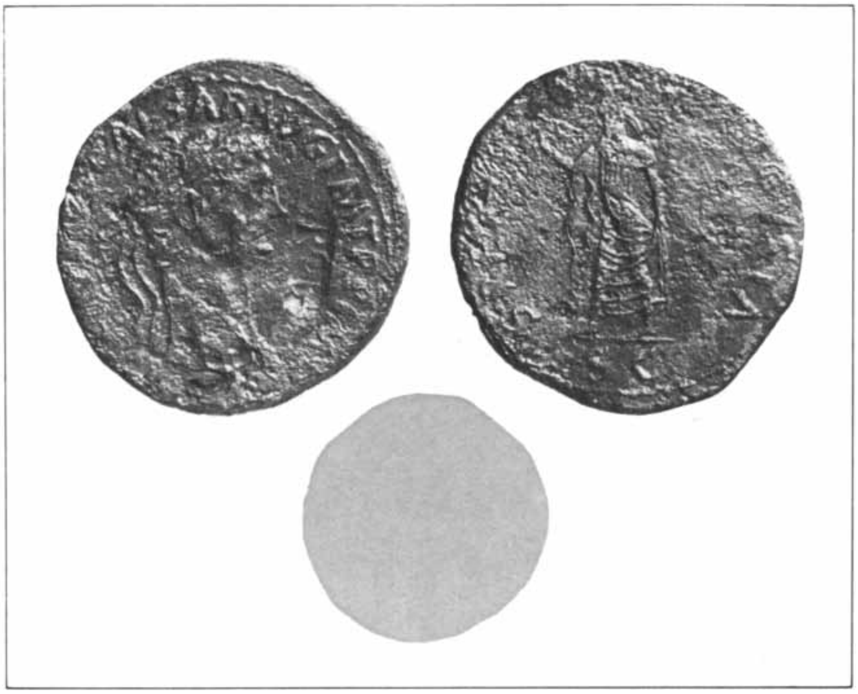
FIRST "RADIATE," the *antoninianus* minted by order of the emperor Caracalla, featured the emperor's portrait bust on its obverse (*left*). The figure on the reverse is Sarapis, a god originally worshiped in Egypt. The ray-decorated crown worn by Caracalla and by later emperors who also minted *antoniniani* gave the name radiate to these coins and to the counterfeits modeled after them. This coin is in the collection of the National Museum of Wales. Like the other coins illustrated it is enlarged; the silhouette shows its true size.

By 270, under Claudius II, the new base radiates reached their low point: only 1 percent silver. Now, as the Swedish scholar Sture Bolin has pointed out, a coin of only 1 percent fine-silver will support about twice its weight in copper, that is, its own weight and that of a silver-free coin of the same weight. The radiates struck by Claudius II, however, weighed rather less than three grams, whereas the lightest *sestertii* of the period averaged 18 grams. Thus it soon came about that virtually the only coins in circulation were base radiates; everything else, including the orthodox bronze coinage, had been abstracted, that is, melted down or hoarded.

We live today at a time of great international anxiety over the general inflation of currencies. It is thus exceptionally interesting to puzzle out the probable effect on ordinary citizens of this other inflation of so many centuries ago. The salient fact to be kept in mind is that the society of the time was far different from the society of today. Industry, if the term is to be given its usual modern connotation, scarcely existed. The fossil fuels on which our life is built were to all intents and purposes unexploited: coal was a wonder and petroleum a medicine. With few exceptions towns were small; outside the two or three big cities of the Mediterranean there was no urban proletariat.

In Britain one Roman town in particular—Calleva Atrebatum (Silchester in Hampshire)—is the most intensively explored of its kind in the Roman empire. It may be taken as an example of a typical local tribal capital or, more precisely, an overgrown village with some urban amenities. How did the inflation affect Silchester?

On my latest estimate the town can hardly have had more than 1,200 inhabitants, and it may have had fewer. Whatever their number, they either were intimately connected with the peasant agriculture on which the economy of the town depended or were handcraftsmen working for a local market, chiefly in local materials. During the inflation the price levels in the marketplace and the shops no doubt soared in nominal terms. On balance, however, so did the prices of the primary products of the land, of labor and of craftsmanship. Land leases were renewed every five years, so that money rents stood to be adjusted at short intervals. According to the late A. H. M. Jones such taxes as were collected in specie appear not to have been increased in proportion to the devaluations of the silver coinage; thus they were a heavy burden only when they were periodically exacted in gold on a community basis. Taxes collected in kind were also a burden, but perhaps only when they are considered in money terms. Viewed in this way the inflation of the third century was certainly a nuisance but by no means a poison applied to the roots of a many-branched society such as ours today.



EARLY COUNTERFEIT, a copy of a *sestertius* minted under the emperor Claudius, not only found its way to Roman Britain but was validated for official circulation. The letter "p" is missing from the counterstamp [P]ROB, or "proved," beside the emperor's nose. Found at Silchester (the Calleva of Roman Britain), the coin is in the Reading Museum. Counterfeiting bronze coins did not violate the laws of Rome until the fourth century.

DEBASED DENARIUS minted under the emperor Caracalla (*left*), with a silver content of only 50 percent, is seen beside a British copy produced by casting. By Caracalla's day the counterfeiting of silver coins had been illegal for some 300 years. Both the orthodox and the counterfeit coins are in the National Museum of Wales; the latter is from Caerleon.

With these notions in mind, let us examine how the people of the period managed to circumvent the damage that inflation wrought on the monetary means of exchange, observing at the same time to what extent counterfeiting served a socially useful purpose. To begin, it is important to realize that the position of the bronze coinage was a peculiar one. To omit further discussion of gold, goods were priced either in silver or in bronze. If one did not have the appropriate coins for the desired purchases, one had to buy them from a money changer, in much the same way that a tourist buys foreign currency at a bank today. For example, at Pergamum in Asia Minor in Hadrian's time, if one wanted to change bronze into silver, the money-changers' price for a *denarius* was 18 *asses;* the cost of changing silver into bronze was 17 *asses* to the *denarius.* From Britain itself comes a reflection of this system. A bronze statuette of a Celtic Mars, found in Lincolnshire and now in the British Museum, bears an inscription on its base. It refers to the cost of some of the casting metal in terms of *denarii* and to the cost of the finished statuette in terms of *sestertii.*

The bronze coinage was peculiar in another respect. From late in Republican times the counterfeiting of Roman silver and gold specie had been prohibited. (The statute was Sulla's *Lex cornelia de falsis,* promulgated in 81 B.C.) As far as we know, however, it was not until A.D. 371 that the bronze coinage also received the protection of the law. By then, however, the character of the coinage had changed completely. It might fairly be said that during the more than 450 intervening years the manufacture and circulation of counterfeit bronze coins gave their makers no pangs of conscience, only an honest profit.

Huge numbers of counterfeit Roman bronze coins arrived in Britain as early as the entry of Claudius' expeditionary force of A.D. 43, not only in the purses of the legionaries but also in the regimental pay chests. At this time the mint in Rome had shut its doors, and no other mint functioned in the western part of the empire. Inevitably a shortage of necessary small change arose, and it could be alleviated only by unofficial means.

Since the counterfeit coins were not illegal, they circulated freely until the Imperial mints stirred themselves sufficiently some 20 years later and recommenced the striking of good supplies of bronze to the proper standard of weight. Governmental acceptance of the counterfeits in the meantime is attested by the fact that, when circulating coins were occasionally subject to inspection, counterfeits of reasonable appearance were often adopted into official use and even were counterstamped PROB, or "proved."

This kind of validation was almost certainly the work of the army officials whose concern it was to pay in the best available coin the wages of the legionaries, who were notoriously unruly individuals. The pay of a legionary private was 225 *denarii* a year; the money was handed over in three annual installments. Late in the first century a fourth payday was added, raising the total wage to 300 *denarii,* and under Caracalla the legions received a 50 percent raise. Caracalla's apparent bounty may well have taken the form of dispensing free rations and equipment, which up to that time had been deducted from pay. Before then the deductions and a system of compulsory savings probably ate up 60 of the 75 *denarii* due to the foot soldier every payday; the remaining 15 were probably paid in bronze. One may conjecture that the practice was halted, probably toward the end of the second century, and the legionaries were paid in silver instead. Otherwise we should find numerous bronze coins when we excavate Roman sites of the third century. (The need of meeting military costs in silver was also an important contribution to the progressive debasement of the *denarius* and the radiate.)

As long as the silver coinage retained an appreciable value, however, the need for bronze small change continued. Under the existing conditions of scarcity the time was ripe for further counterfeiting. Now, as we have seen, counterfeiting can serve a useful social purpose. It is never, however, a charitable enterprise. If the counterfeiter is to extract a profit from his labors, copies of an orthodox coin in the same metal or a similar one must be lighter than the original. Over a long period of shortage the false coins tend to become lighter and lighter. This is also true because the counterfeiter frequently used earlier counterfeits, already reduced in weight, as the models for new forgeries.

In the early decades of the third century the counterfeiters' models were usually orthodox coins of the preceding century, and the false coins were produced both by casting and by die-striking. Evidently too few orthodox coins of contemporary mintage were available to serve as models. An example from Gaul

DEBASED RADIATES, minted by the independent rulers of Gaul in imitation of the orthodox Imperial coinage, show (*top*) the emperor Tetricus I and (*bottom*) Salonina, wife of the preceding emperor Gallienus. The top coin, heavily silver-plated, is in the collection of P. J. Casey; the bottom coin, in the National Museum of Wales, is from Caernarvon.

LIGHTWEIGHT COUNTERFEIT (*bottom*) copies a bronze *as* minted under the emperor Commodus (*top*). The difference in weight between the orthodox coin (11.5 grams) and its copy (3.08 grams) was the counterfeiter's margin of profit. The orthodox *as* is in the British Museum; the imitation, found at Caerleon, is in the National Museum of Wales.

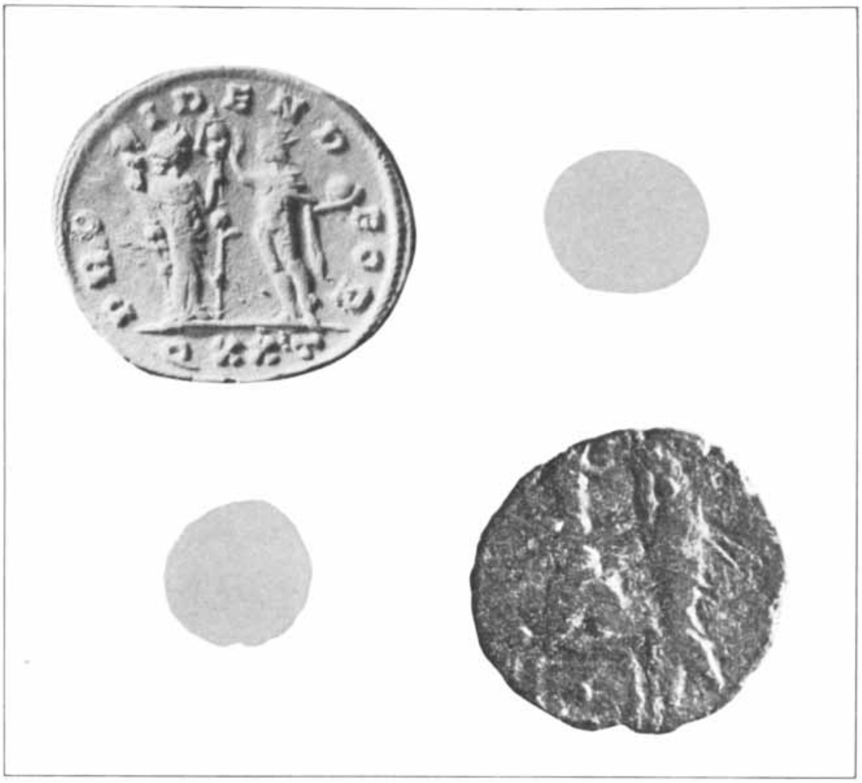
demonstrates that the counterfeiters of bronze coins even practiced conserving metal by a technique used for centuries by the forgers of silver coins: plating a base blank. The sample coin, from a private French collection, is a counterfeit *as* bearing the portrait bust of Lucilla, wife of the emperor Verus (161-169). The core of the coin is iron; its bronze skin was applied either as a wrapping of foil or by dipping the iron core in molten bronze. The composite blank was then die-struck.

Not long afterward there appeared in Germany, Gaul and Britain a distinctive type of very light cast-bronze counterfeit. Its presence has only rather recently been recognized in these areas, although such coins have long been known from sites along the Danube and its hinterland. They are at most only half the orthodox weight and often only a quarter of it. In the words of the Austrian scholar Wilhelm Kubitschek, who

COUNTERFEITERS' MOLDS, found in France and England, were used to make imitations of the radiates issued under three Gallic emperors. The interlocking pair of molds (*top*) was found at Lyons and is now in the Rouen Museum; they imitate a coin of the emperor Postumus. The mold with a standing figure (*middle, impression of mold at right*) was found at Whitchurch. Now in the Bristol City Museum, it imitates a coin of the emperor Victorinus. The third mold (*bottom, impression at right*) is also from Whitchurch and now at Bristol; it imitates a coin of Tetricus I. Casting, an inefficient way to make small coins, was not continued for long.

studied a series of these forgeries from the site of Carnuntum, near Vienna, they are of "unspeakable wretchedness." The soundness of Kubitschek's remark may be judged by comparing an orthodox *as* of Commodus (180-192) with lightweight cast counterfeits of the same issue found at Caerleon in Wales [*see bottom illustration on page 125*]. The original weighs a full 11.55 grams; the counterfeits weigh only a little more than three grams. In addition to being very thin they are poor reproductions of the original and are also cast in bad metal.

The most recent orthodox coins used as models for these cast counterfeits in Britain are scarcely later than the time of the emperor Decius (249-251). Nonetheless a deposit of counterfeiter's molds found at Whitchurch, near Bristol, includes among many intended for the forging of radiates one mold for a counterfeit lightweight *as* of the time of Hadrian (117-138). The mold was evidently in active use in about 273. This is proof that the cast-bronze lightweights still had some part to play even after the Imperial "silver" coinage had reached its low point.



SUN GOD INTO FISH, a transformation that suggests at least a failure in communication and probably illiteracy on the part of the counterfeiter, is documented by this radiate of the emperor Aurelian (*left*) and its imitation, found in Wales. The two figures on the orthodox coin are (*left to right*) Providence, holding military standards, being saluted by Sol, who holds an orb in his left hand. The British imitation transforms Sol into a great gaffed fish and Providence and the standards into a tiny fisherman on a knobby platform. Orthodox coin is in the British Museum; the imitation is in the National Museum of Wales.

Whereas the need for small change made the forgery of bronze coins a profitable "social service," counterfeiters were not reluctant to forge silver coins as well. As a French antiquary, the Abbé Mahudel, noted in 1723, commenting on the discovery at Lyons of a cache of molds for counterfeit *denarii*, "The decadence of [die] engraving, already considerable by the time of Septimius Severus, and the alteration which he had introduced in the fineness of the coinage, favored . . . counterfeiters . . . more and more by making their deceptions easier." Molds for *denarii*, and also molds for radiates once that denomination had achieved dominance, have been found at scores of sites all over Europe. There are two dozen such sites in Britain alone. The "silver" counterfeits were made of base metal that was tinned or silver-plated in order to pass. An example from Caerleon is a forged *denarius* of the time of Caracalla [*see bottom illustration on page 124*]. Some such coins were made entirely of tin; in a deposit of thousands of molds unearthed at Edington in Somerset was found not only a lump of tin but also a tin coin still in place between the two halves of a mold.

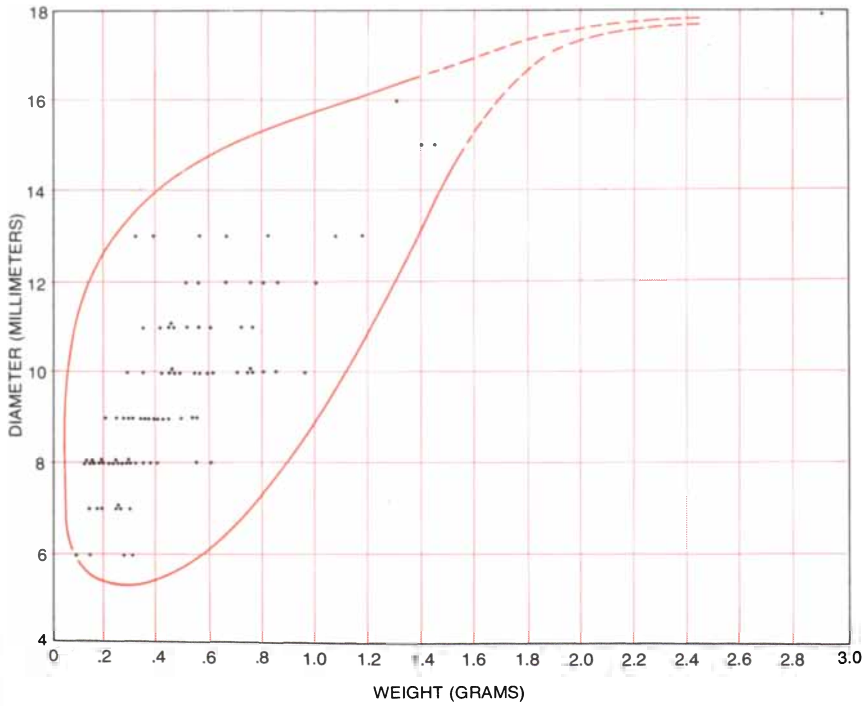
It seems reasonable to date the wave of silver-coin counterfeiting represented by these finds to the 40 years or so before control of Gaul reverted to Rome in 274. Certainly the orthodox "silver" coinage issued by the Gallic emperors was not exempt from imitation by the counterfeiters. Baked clay molds for the production of counterfeit radiates of the time of Postumus (259-268) have been found at Lyons, and molds for the radiates of his successors, Victorinus (269-270) and Tetricus I (270-274) are among those discovered at Whitchurch [*see illustration on opposite page*].

Much is made in archaeological circles of the evidence that literacy in Romano-British society had fallen to a remarkably low level during these troubled decades. The discovery of writing implements at various sites and the appearance of graffiti on contemporary tiles and pottery certainly denote a basic ability to write and cipher, but one gains the impression that effective literacy was limited to such practical matters as marks of ownership, tally-keeping and very little else. The work of many counterfeiters throughout the period provides strong evidence in support of this contention.

To take one striking example, the emperor Aurelian (270-275) attempted to reform the Roman coinage by introducing a radiate that was larger and better-looking than those of his predecessors. It even contained a little more silver: as much as 3 to 4 percent. The relationship between the new coins and the old remains problematical, but the Roman numerals "XX I" that appear on some suggest that they were valued at 20 *sestertii*, that is, five *denarii*. The new coins did not reach Britain in adequate quantities, and one of them, bearing a design inscribed "Providence of the Gods," attracted the attention of a British counterfeiter. Aurelian's design shows the sun god, Sol, on the right, with his arm raised in a salute to Providence, who stands at the left, carrying two military standards.

Consider now a coin less than half the diameter of the Aurelianic original; it was unearthed in the ruins of a Roman farm at Dinoelen near Abergele in Wales. Its design shows a great fish at the right, being gaffed by a diminutive fisherman who stands on a platform at the left. When the British counterfeit is compared with the Aurelianic original, the transformation is easy to reconstruct. Sol, standing with his feet apart, has become a fork-tailed fish. The god's outstretched left arm, hand grasping an orb,

SHRUNKEN SIZE of the later radiates counterfeited in Britain is exemplified by this imitation (*bottom*) of a coin of the emperor Probus (*top*). The orthodox radiate, minted during the decades when the very poor alloy was only "washed" with silver, was unearthed in Silchester and is in the Reading Museum. Its counterfeit, found at Caernarvon and now at the National Museum of Wales, was also probably made shiny with a silver or tin wash.



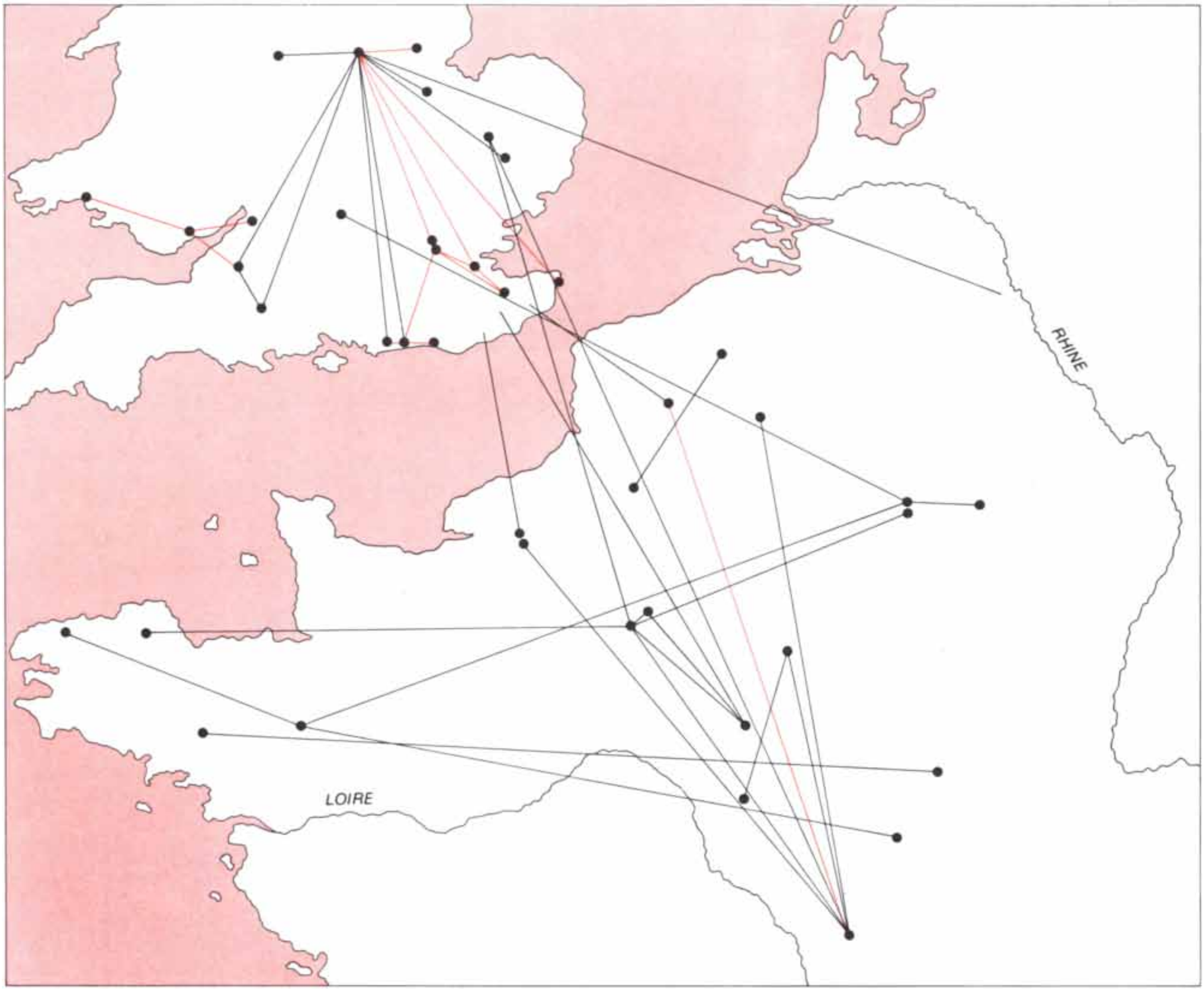
HOARD OF COUNTERFEITS, discovered at Caerleon in 1955, included 95 miniature radiates. The scattergram shows the position of each radiate in terms of weight and diameter. As the enclosing envelope indicates, except for one large and heavy specimen (*right*), most of the coins were smaller than .5 inch in diameter and less than .2 ounce in weight.

has been transformed into the fish's broad dorsal fin, and all that is left of Providence's standards is a knobby platform. It is clear that whoever engraved the dies for the British counterfeit had found the inscription and design of the original totally unintelligible.

The requirement, imposed by inexorable economics, that counterfeit coins be lighter than the orthodox coins they imitate probably contributed to the abandonment of casting in favor of die-striking in the last quarter of the third century. The time consumed in making and assembling composite molds and in detaching the casts meant in any case that casting could not meet the demand for the product. Moreover, casting is insufficiently flexible as a means of reducing the size of the counterfeits; die-striking is not only faster but also can produce coins only five millimeters in diameter that weigh as little as .3 gram. An example of the flexibility of die-striking is provided by one deposit of counterfeits found in Wales. Unearthed near a Roman farm at Coygan, all the coins had been struck by the same combination of dies. They varied only a little in diameter, which ranged from 10 to 12 millimeters. In weight, however, the variation was as much as threefold: from .27 to .85 gram.

The last Imperial Roman radiates to be counterfeited in Britain are those of the emperor Probus (276-282) and such as were found at Coygan. It seems that, even though Carausius and Allectus, pretenders to the purple, ruled in Britain between 286 and 293, the long arm of Rome managed somehow to bring the British counterfeiting epidemic under control beforehand. At least no counterfeits of Imperial coins later than those of Probus are known. The Coygan counterfeits bore every sign of hasty concealment, as if the counterfeiters had been warned of approaching punitive forces. The pit where they were concealed had been carefully masked by stones, but vestiges of the coiners' activities were apparent in the floor crevices of an adjacent hut. In any event Diocletian's sweeping reform of 294, together with subsequent developments, certainly killed any survivors of the small-radiate epidemic. Only a few counterfeits of larger coins are found in fourth-century sites; together with diminishing numbers of orthodox Gallic coins they evidently served as small change for many decades. The problem of counterfeiting in the fourth century is not considered here.

One large question is left in the mind



MINIATURE RADIATES circulated very widely in Britain and Gaul, disproving the contention that they were used only as local token money. Lines in black connect points where radiates struck from the same dies have been found; where a terminal dot is missing the precise site of the find is uncertain. Lines in color connect points where radiates of similar style have been found. Data are from studies by Harold B. Mattingly of the University of Leeds, by Jean-Baptiste Giard of the Bibliothèque Nationale in Paris and by Marcel Thirion and Mlle J. Lallemand of the Bibliothèque Royale in Brussels, provided through the courtesy of Pierre Bastien.

of the modern scholar. How could such tiny counterfeit coins, only a fraction of the orthodox weight, have served in company with orthodox coins as a medium of exchange? It is true that very small coins have existed in many different countries and at many different periods, but in most instances they have been "value coins" of silver or even gold. When we look at these miniature radiates today, however, we see little more than irregular disks of green decay.

To answer the question we must look at the tiny counterfeits with hindsight. Although they were base, they were, we must remember, imitations of coins that were equally base and yet counted as silver denominations. Moreover, even the worst of the orthodox radiates wore a silver wash when they were first issued, and there is every reason to believe that the worst and smallest of the counterfeits were similarly shiny when they were new. The plating may have been silver (or tin, which is easier to apply). Furthermore, even on the best of the orthodox coins the plating was fugitive. If we can imagine the little counterfeits as being bright and silvery when they were first produced, we are well on the way to accepting the fact that they could be, and were, regarded by the folk of third-century Britain as possessing value.

What was that value? As the result of recent discoveries in a small cave in the Mendip Hills south of Bristol the question can be answered with almost mechanical precision. Here a collection of the raw material used for making small radiates was found a few years ago. As was so often the case, the raw material consisted of contemporary coins: orthodox radiates and large copies of them. These coins were cut into quarters [*see illustration on next page*]. Each of the four segments was then heated and die-struck.

Now, it is obvious that each such new counterfeit must have been worth more to the counterfeiter than a quarter of the value of an orthodox coin; otherwise there would have been no point in the procedure. To reiterate, counterfeiting may occasionally be a public service, but it is never conducted as such. Indeed, it seems fair to say that during this period such counterfeits, whatever their size, probably passed for the same value as full-size orthodox coins. Ninety-five counterfeit radiates in a hoard unearthed

QUARTERED RADIATES, found in a counterfeiter's den hidden in the Mendip Hills, included one (*top*) minted in Gaul under the emperor Postumus. It had been filed for quartering and one quadrant had been partially detached. The counterfeiters used each quarter as raw material for minting a miniature radiate. The coins of Rome were not held to be worth more than those of Gaul. One quarter of such an orthodox coin is seen at bottom; the incomplete inscription [TA]CITVS shows that it was minted under the emperor Tacitus (275-276), the last Roman ruler but one whose coins were so used by counterfeiters in Britain.

in 1955 at Caerleon showed wide divergences in size and weight. In all likelihood, however, all of them were identical in value. That value, as we have seen, was small and grew steadily smaller; inflation was certainly not halted by Aurelian's efforts at reform. The same Mendip cave includes a neat instance of the depreciation of the reformed Imperial currency. One of the many quartered pieces bears part of the name of Aurelian's successor, the emperor Tacitus (275-276). The radiate, although supposedly of established Imperial worth, had been treated as raw material by the British counterfeiters in the same way that they had treated any Gallic radiate or counterfeit thereof, and had therefore presumably fallen to the same low value.

It was long customary among numismatists to explain the miniature radiates as being essentially a local coinage, a token money provided to a dependent peasantry by the owners of great estates. Nothing could be further from the truth. The counterfeiters' illicit mints were indeed widely scattered. They were often located in remote places, such as the Coygan farmstead or the cave in the Mendip Hills, but this was not always so. There were counterfeiters' dens in London itself, and many of the coiners' wares traveled far outside the region where they were struck. A great deal of painstaking work by Harold B. Mattingly in Britain, by Marcel Thirion and Mlle J. Lallemand in Belgium and by Jean-Baptiste Giard in France has succeeded in tracing coins struck by the same dies or in the same style over wide areas [*see illustration on preceding page*]. Such discoveries further attest the value that people of this period ascribed to the counterfeit radiates, for all their small size and base metal.

The history of the inflation in the third century is not without interest to us today. It is difficult to see how the Roman authorities could have avoided the debasement of silver that was responsible for the inflation, a process that began in the first century under Nero. It is equally difficult today to see (or at least economists, of whom I am not one, cannot agree) how the printing of worthless money can be restrained now without dire effects on the fabric of society. The riddle existed, though in different form, in antiquity. It is still with us.

MATHEMATICAL GAMES

The arts as combinatorial mathematics, or how to compose like Mozart with dice

by Martin Gardner

"Mathematics and music! The most glaring possible opposites of human thought! Yet connected, mutually sustained!"

—Hermann von Helmholtz

There is a trivial sense in which any work of art is a combination of a finite number of discrete elements. Not only that, the precise combination of the elements can be expressed by a sequence of digits or, if you will, by one enormous number.

Consider a poem. Assign distinct numbers to each letter of the alphabet, to each punctuation symbol and so on. A certain digit, say 0, can be used to separate the numbers. It is obvious that one long string of digits can express the poem. If the books of a vast library contain every possible combination of words and punctuation marks, as they do in Jorge Luis Borges' famous story, "The Library of Babel," then somewhere in the collection is every poem ever written or that can be written. Imagine those poems coded as digital sequences and indexed. If one had enough time, billions on billions of years, one could locate any specified great poem. Are there algorithms by which one could find a great poem not yet written?

Consider a painting. Rule the canvas into a matrix of minute cells. The precise color of each cell is easily coded by a number. Scanning the cells yields a chain of numbers that expresses the painting. Since numbers do not decay, a painting can be re-created as long as the number sequence is preserved. Future computers will be able to reproduce a painting more like the original than the original itself, since after a few decades the original will have physically deteriorated to some extent. If a vast art museum contains every combination of colored cells for matrixes not exceeding a certain size, somewhere in that monstrous museum will hang every picture ever painted or that can be painted. Are there algorithms by which a computer could search a list of the museum's code numbers and identify a sequence for a great painting not yet painted?

Consider a symphony. It is a fantastically complex blend of discreteness and continuity; a violin or a slide trombone can move up and down the scale continuously but a piano cannot produce quarter-tones. We know, however, from Fourier analysis that the entire sound of a symphony, from beginning to end, can be represented by a single curve on an oscilloscope. "This curve," wrote Sir James Jeans in *Science and Music*, "*is* the symphony—neither more nor less, and the symphony will sound noble or tawdry, musical or harsh, refined or vulgar, according to the quality of this curve." On a long-playing record a symphony is actually represented by one long space curve.

Because curves can be coded to any desired precision by numbers, a symphony, like a painting or a poem, can be quantized and expressed by a number chain. A vast library of computer tapes, recording all combinations of symphonic sounds, would contain every symphony ever written or that could be written. Are there algorithms by which a computer could scan the number sequences of such a library and pick out a great symphony not yet written?

Such procedures would, of course, be so stupendously complex that man may never come close to formulating them, but that is not the point. Do they exist in principle? Is it worthwhile to look for bits and pieces of them? Consider one of the humblest of such aesthetic tasks, the search for rules that govern the invention of a simple melody. Is there a procedure by which a person or a computer can compose a pleasing tune, using no more than a set of combinatorial rules?

If we restrict the tune to a finite length and a finite number of pure tones and rhythms, the number of possible melodies is finite. John Stuart Mill, in his autobiography, recalls that as a young man he was once "seriously tormented by the thought of the exhaustibility of musical compositions." Suppose our tune is made up of just 10 notes chosen from the set of eight notes in a single octave. The number of melodies is the same as the number of 10-letter words that can be formed with eight distinct letters, allowing duplications. It is $8^{10} = 1,073,741,824$, and this without even considering varying rhythms which create, in effect, varying melodies. Most of these tunes will be dull (mi, mi, mi, mi, mi, mi, mi, mi, mi, mi for instance), but many will be extremely pleasing. Are there rules by which a computer or a person could pick out the pleasing combinations?

Attempts to formulate such rules and embody them in a mechanical device for composing tunes have a curious history that began in 1650 when Athanasius Kircher, a German Jesuit, published in Rome his *Musurgia universalis sive ars magna consoni et dissoni*. Kircher was an ardent disciple of Ramón Lull, the Spanish medieval mystic whose *Ars magna* derived from the crazy notion that significant new knowledge could be obtained in almost every field simply by exploring all combinations of a small number of basic elements. It was natural that Kircher, who later wrote a 500-page elaboration of Lull's "great art," would view musical composition as a combinatorial problem. In his music book he describes a Lullian technique of creating polyphony by sliding columns alongside one another, as with Napier's bones [see "Mathematical Games," Scientific American, March, 1973], and reading off rows to obtain various permutations and combinations. Like all of Kircher's huge tomes, the book is a fantastic mix of valuable information and total nonsense, illustrated with elaborate engravings of vocal cords, bones in the ears of various animals, birds and their songs, musical instruments, mechanical details of music boxes, water-operated organ pipes with animated figures of animals and people, and hundreds of other curious things.

The Lullian device described by Kircher was actually built, circa 1670, for the diarist Samuel Pepys, who owned a copy of Kircher's music book and much admired it. The original machine, called "musarithmica mirifica," is in the Pepys Museum at Pepys' alma mater, Magdalene College, Cambridge.

During the early 18th century many German music scholars became interested in mechanical methods of composition. Lorenz Christoph Mizler wrote a book in 1739 describing a system that produced figured bass for baroque ensemble music. In 1757 Bach's pupil, Johann Philipp Kirnberger, published in Berlin his *Ever-ready Composer of Polonaises and Minuets,* using a die for randomizing certain choices. In 1783 another book by Kirnberger extended his methods to symphonies and other forms of music.

Toward the close of the 18th century the practice of generating melodies with the aid of tables and randomizers such as dice or teetotums became a popular pastime. Maximilian Stadler, an Austrian composer, published in 1779 a set of musical bars and tables for producing minuets and trios with the help of dice. At about the same time a London music publisher, Welcker, issued a "tabular system whereby any person, without the least knowledge of music, may compose ten thousand minuets in the most pleasing and correct manner." Similar anonymous works were falsely attributed to well-known composers such as C. P. E. Bach (son of Johann Sebastian Bach) and Joseph Haydn. The "Haydn" work, "Gioco filarmonico" ("Philharmonic Joke"), Naples, 1790, was recently discovered by the Glasgow mathematician Thomas H. O'Beirne to be a plagiarism. Its bars and tables are identical with Stadler's.

The most popular work explaining how a pair of dice can be used "to compose without the least knowledge of music" as many German waltzes as one pleases was first published in Amsterdam and in Berlin in 1792, a year after Mozart's death. The work was attributed to Mozart. Most Mozart scholars say it is spurious, although Mozart was fond of mathematical puzzles and did leave handwritten notes showing his interest in musical permutations. (The same pamphlet was issued in Bonn a year later, with a similar work, also attributed to Mozart, for dice composition of country dances. The contredanses pamphlet was reprinted in 1957 by Heuwekemeijer in Amsterdam.)

Mozart's *Musikalisches Würfelspiel,* as the waltz pamphlet is usually called, has been reprinted many times in many languages. In 1806 it appeared in London as *Mozart's Musical Game, Fitted in an Elegant Box, Showing by Easy System How to Compose an Unlimited Number of Waltzes, Rondos, Hornpipes and Reels.* In New York in 1941 the Hungarian composer and concert pianist Alexander Laszlo brought it out under the title *The Dice Composer,* orchestrating the music so that it could be played by chamber groups and orchestras. The system popped up again in West Germany in 1956 in a score published by B. Schott. Photocopies of Schott's charts and musical bars appear in the instruction booklet for *The Melody Dicer,* issued early this year by Carousel Publishing Corporation, 27 Union Street, Brighton, Mass. This boxed set also includes a pair of dice and blank sheets of music paper.

The "Mozart" system consists of a set of short measures numbered 1 through 176. The two dice are thrown 16 times. With the aid of a chart listing 11 numbers in each of eight columns, the first eight throws determine the first eight bars of the waltz. A second chart is used for the second eight throws that complete the 16-bar piece. The charts are constructed so that the waltz opens with the tonic or keynote, modulates to the dominant, then finds its way back to the tonic on its final note. Because all bars listed in the eighth column of each chart are alike, the 11 choices (sums 2 through 12 on the dice) are available for only 14 bars. This allows the system to produce 11^{14} waltzes, all with a distinct Mozartean flavor. The number is so large that any waltz you generate with the dice and actually play is almost certainly a

Frontispiece of Kircher's* Musurgia universalis *(1650), with Pythagoras at lower left

waltz never heard before. If you fail to preserve it, it will be a waltz that will probably never be heard again.

The first commercial recording of "Mozart" dice pieces was made by O'Beirne. Both the randomizing of the bars and the actual playing of the melodies was done by Solidac, a small and slow experimental computer designed and built between 1959 and 1964 by the Glasgow firm of Barr and Stroud, where O'Beirne was then chief mathematician. It was the first computer built in Scotland. O'Beirne programmed Solidac to play the pieces in clarinetlike tones, and a long-playing recording of selected waltzes and contredanses was issued by Barr and Stroud in 1967. This recording is no longer available, but O'Beirne writes that he will send a standard cassette giving two half-hour tracks of varied Solidac dice music to any reader who sends $10 to his home address: 8 Rosslyn Terrace, Glasgow G12 9NB, Scotland. O'Beirne is the author of an excellent book on mathematical recreations (*Puzzles and Paradoxes*, Oxford University Press, 1965). He has been of invaluable help in the preparation of this account.

Other methods of producing tunes mechanically were invented in the early 19th century. Antonio Calegari, an Italian composer, used two dice for composing pieces for the pianoforte and harp. His book on the system was published in Venice in 1801, and later in a French translation. *The Melographicon*, an anonymous and undated book issued in London about 1805, is subtitled: "A new musical work, by which an interminable number of melodies may be produced, and young people who have a taste for poetry enabled to set their verses to music for the voice and pianoforte, without the necessity of a scientific knowledge of the art." The book has four parts, each providing music for poetry with a certain meter and rhyme scheme. Dice are not used. One simply selects any bar from group *A*, any from group *B* and so on to the last letter of the alphabet for that section.

A photograph of a boxed dice game appears in Plate 42 of *The Oxford Companion to Music*, but without mention of date, inventor or place of publication. Apparently it uses 32 dice, their sides marked to indicate tones, intervals, chords, modulations and so on. There also are ivory men whose purpose, the caption reads, is "difficult to fathom."

In 1822 a machine called the Kaleidacousticon was advertised in a Boston music magazine, *The Euterpiad*. By shuffling cards it could compose 214 million waltzes. The Componium, a pipe organ that played its own compositions, was invented by M. Winkel of Amsterdam and created a sensation when it was exhibited in Paris in 1824. Listeners could not believe that the machine actually constructed the melodies it played. Scientists from the French academy investigated.

"When this instrument has received a varied theme," their report stated, "which the inventor has had time to fix by a process of his own, it decomposes the variations of itself, and reproduces their different parts in all the orders of possible permutation.... None of the airs which it varies lasts above a minute; could it be supposed that but one of these airs was played without interruption, yet, the principle of variability which it possesses, it might, without ever resuming precisely the same combination, continue to play... during so immense a series of ages that, though figures might be brought to express them, common language could not."

The report, endorsed by physicist Jean Baptiste Biot, appeared in a British musical journal, *The Harmonicon*, Volume 2, pages 40-41, 1824. Winkel's machine inspired a Vienna inventor, Baron J. Giuliani, to build a similar device, the construction of which is given in detail on pages 198-200 of the same volume.

In 1865 a composing system called the Quadrille Melodist, invented by J. Clinton, was advertised in *The Euterpiad*. By shuffling a set of composing cards a pianist at a quadrille party could "keep the evening's pleasure going by means of a modest provision of 428,000,000 quadrilles."

Joseph Schillinger, a Columbia University mathematician who died in 1943, published his mathematical system of musical composition in a booklet, *Kaleidophone*, in 1940. George Gershwin is said to have used the system in writing *Porgy and Bess*. In 1946 Heitor Villa-Lobos, using the system, translated a silhouette of New York City's skyline into a piano composition. *The Schillinger System of Musical Composition* is a two-volume work by L. Dowling and A. Shaw, published by Carl Fischer in 1941. A footnote on page 673 of Schillinger's eccentric opus, *The Mathematical Basis of the Arts* (Philosophical Library, 1948), says that he left plans for music-composing machines, protected by patents, but nothing is said about their construction.

In the 1950's information theory was applied to musical composition by J. R. Pierce and others. In a pioneering article "Information Theory and Melody" (*Scientific American*, February, 1956), chemist Richard C. Pinkerton included a graph which he called the "banal tune-maker." By flipping a coin to determine paths along the network one can compose simple nursery tunes. Most of them are monotonous, but hardly more so, Pinkerton reminds us, than "A Tisket, a Tasket."

During the 1960's and early 1970's

SIMP table A	SIMP table B	SIMP table C
1. In particular,	1. a large portion of the interface coordination communication.	1. must utilize and be functionally interwoven with
2. On the other hand,	2. a constant flow of effective information	2. maximizes the probability of project success and minimizes the cost and time required for
3. However,	3. the characterization of specific criteria	3. adds explicit performance limits to
4. Similarly,	4. initiation of critical subsystem development	4. necessitates that urgent consideration be applied to
5. As a resultant implication,	5. the fully integrated test program	5. requires considerable systems analysis and trade off studies to arrive at
6. In this regard,	6. the product configuration baseline	6. is further compounded, when taking into account
7. Based on integral subsystem considerations,	7. any associated supporting element	7. presents extremely interesting challenges to
8. For example,	8. the incorporation of additional mission constraints	8. recognizes the importance of other systems and the necessity for
9. Thus,	9. the independent functional principle	9. effects a significant implementation of
0. In respect to specific goals,	0. a primary interrelationship between system and/or subsystem technologies	0. adds overriding performance constraints to

Honeywell's buzz-phrase generator for writing Simplified Integrated Modular Prose (SIMP) the proliferation of computers and the development of sophisticated electronic tone synthesizers opened a new era in machine composition of music. It is now possible to write computer programs that go far beyond the crude devices of earlier days. Suppose one wishes to compose a melody in imitation of one by Chopin. A computer analysis is made of all Chopin melodies so that the computer has in its memory a set of "transition probabilities." These give the probability that any set of one, two, three or more notes in a Chopin melody is followed by any other note. Of course, one must also take into account the type of melody one wishes to compose, the rhythms, the position of each note within the melody, the overall pattern and other things. In brief, the computer makes random choices within a specified general structure, but these choices are subject to rules and weighted by Chopin's transition preferences. The result is a "Markoff chain" melody, undistinguished but nevertheless sounding curiously like Chopin. The computer can quickly dash off several hundred such pieces, from which the most pleasing may be selected.

There is now a rapidly growing literature on computer composition, not only of music in traditional styles but also music that takes full advantage of the computer's ability to synthesize weird sounds that resemble none of the sounds made by familiar instruments. Microtones, strange timbres, unbelievably complex rhythms and harmonics are no problem. The computer is a universal musical instrument. In principle it can produce any kind of sound the human ear is capable of hearing. Moreover, a computer can be programmed to play one of its own compositions at the same time it is composing it.

Interested readers should consult "Computer Music," by Lejaren A. Hiller, Jr., *Scientific American*, December, 1959, and the book, *Experimental Music*, which Hiller wrote with Leonard M. Isaacson (McGraw-Hill, 1959). The book includes the music for their first major computer composition, *The Illiac Suite for String Quartet*. More recent books are *Music by Computers*, edited by Heinz von Foerster and James W. Beauchamp (Wiley, 1969) and *Formalized Music: Thought and Mathematics in Composition* (Indiana University Press, 1971) by a Greek-born Paris composer, Iannis Xenakis. *A Bibliography of Electronic Music*, compiled by Lowell M. Cross, listing 1,563 references, was published in 1967 by the University of Toronto Press.

How can we sum up? Computers certainly can compose mediocre music, frigid and forgettable, even though the music has the flavor of a great composer. No one, however, has yet found an algorithm for producing even a simple melody that will be as pleasing to most people of a culture as one of their traditional popular songs. We simply do not know what magic takes place inside the brain of a composer when he creates a superior tune. We do not even know to what extent a tune's merit is bound up with cultural conditioning or even with hereditary traits. About all that can be said is that a good melody is a mixture of predictable patterns and elements of surprise. What the proportions are and how the mixture is achieved, however, still eludes everybody, including composers.

O'Beirne has called my attention to how closely some systems of musical composition resemble the Buzz Phrase Generator [*see illustration on these two pages*]. This is a give-away of Honeywell Incorporated. Pick at random any four-digit number, such as 8751, then read off phrase 8 of module *A*, phrase 7 of module *B* and so on. The result is a SIMP (Simplified Integrated Modular Prose) sentence. "Add a few more four-digit numbers," the instructions say, "to make a SIMP paragraph. After you have mastered the basic technique, you can realize the full potential of SIMP by arranging the modules in *DACB* order, *BACD* order, or *ADCB* order. In these advanced configurations, some additional commas may be required."

SIMP sounds very much like authentic technical prose, but on closer inspection one discovers that something is lacking. Computer-made melodies are perhaps less inane, closer to the random abstract art of a kaleidoscope, but still something essential (nobody knows what) is missing. Indeed, a good simple tune is much harder to compose than an orchestral piece in the extreme avant-garde manner, so loaded with randomness and dissonance that one hesitates to say, as Mark Twain (or was it Bill Nye?) said of Wagner's music: It is better than it sounds.

When a computer generates a melody that becomes as popular as (think of the title of your favorite song), you will know that a colossal breakthrough has been made. Will it ever occur? If so, when? Experts disagree on the answers as much as they do on if and when a computer will write a good poem, paint a good picture or play a game of master chess.

SIMP table **D**

1. the sophisticated hardware
2. the anticipated fourth generation equipment
3. the subsystem compatibility testing
4. the structural design, based on system engineering concepts
5. the preliminary qualification limit
6. the evolution of specifications over a given time period
7. the philosophy of commonality and standardization
8. the greater fight-worthiness concept
9. any discrete configuration mode
0. the total system rationale

Last month's combinatorial problems are answered as follows:

1. Cards 1 through 9, of three suits, are to be arranged in a row so that for every card value k, just k cards are between the first and second cards of value k and between the second and third cards of value k. Not counting reversals, there are three solutions:

1, 8, 1, 9, 1, 5, 2, 6, 7, 2, 8, 5, 2, 9, 6, 4, 7, 5, 3, 8, 4, 6, 3, 9, 7, 4, 3

1, 9, 1, 2, 1, 8, 2, 4, 6, 2, 7, 9, 4, 5, 8, 6, 3, 4, 7, 5, 3, 9, 6, 8, 3, 5, 7

1, 9, 1, 6, 1, 8, 2, 5, 7, 2, 6, 9, 2, 5, 8, 4, 7, 6, 3, 5, 4, 9, 3, 8, 7, 4, 3

The third solution was found without computer aid in 1966 by Eugene Levine, now a mathematician at Adelphi University. It was published in "The Existence of Perfect 3-Sequences," *The Fibonacci Quarterly*, Volume 6, pages 108-112, November, 1968. Levine shows that a solution for triplets exists only when n, the highest value of a card, has a digital root of 1, 8 or 9 (that is, when n equals -1, 0 or 1, modulo 9), and that 9 is the smallest n that has a solution. Levine found solutions for the next higher cases of $n = 10$, 17, 18 and 19, and conjectured that there are solutions for all higher values meeting his proviso.

D. P. Roselle and T. C. Thomasson, Jr., writing "On Generalized Langford Sequences," *Journal of Combinatorial Theory*, Volume 11, pages 196-199, September, 1971, reported computer results that confirm there is no solution for $n = 8$, and came up with the same solution for $n = 9$ that Levine found. An exhaustive computer search for $n = 9$ and $n = 10$ was made by G. Baron, who reported his results at a conference on combinatorics held in Hungary in 1969. He found the three solutions for $n = 9$ and five solutions for $n = 10$. These results are given in his paper in *Combinatorial Theory and Its Applications: Volumes 1, 2 and 3*, edited by Paul Erdös, A. Rényi and Vera T. Sós (North-Holland, 1970). No solution has yet been found for this problem if there are more than three duplicates of each value.

2. The solution to David L. Silverman's problem of pairing each spade with a heart so that the pair-sum is square is: 1-8, 2-2, 3-*K*, 4-*Q*, 5-*J*, 6-10, 7-9, 8-1, 9-7, 10-6, *J*-5, *Q*-4, *K*-3. As Silverman observes, 9, 10 and *J* must be paired with 7, 6 and 5. This establishes six pairings. Since the 6's have been used, 3 pairs only with *K*. Since the 5's have been used, 4 pairs only with *Q*. The remaining three gaps are filled in only one way, proving the solution's uniqueness.

3. Tom Ransom's problem asked for the minimum number of cards, in a set of five, that must be turned over to answer the question: Are all colored-back cards jokers? Letter the cards in last month's illustration from *A* through *E*. Obviously *D* must be turned to see if it is a joker, and *E* must be turned to see if it has a colored back. This gives four possibilities:

(1) *D* is a joker, *E* has a black back.

(2) *D* is a joker, *E* has a colored back.

(3) *D* is not a joker, *E* has a black back.

(4) *D* is not a joker, *E* has a colored back.

For cases 2, 3 and 4 the answer to the question is no. No more cards need be turned. For case 1 the answer is yes, but it takes more thinking to realize that turning the other three cards cannot contradict this answer. *B* is irrelevant because it has a black back. Seeing the back of either joker is also irrelevant. If a joker's back is black, it is not involved in the question. If it is colored, the answer is still yes. Most people staring at an actual row of cards have such an overwhelming desire to see the backs of the jokers that they usually answer: *A, C, D, E*.

One might conclude, therefore, that turning *D* and *E* is sufficient to answer the question. It is not! Recall last month's story about the cautious logician who observed a black sheep in Scotland and concluded that at least one sheep in Scotland is black on at least one side? When someone thinks he has solved the problem, Ransom turns over card *B* to reveal that its other side is a colored back! This, of course, contradicts a yes answer. The correct solution, therefore, is that card *B* as well as cards *D* and *E* must be turned.

Ransom has a second "kicker," suggested by his friend P. Howard Lyons. So that a person working on the problem will not forget the exact phrasing of the question ("Are all colored-back cards jokers?"), Ransom writes it on a file card which is placed above the row of cards. This card also must be turned to determine whether its back is colored or black!

THE AMATEUR SCIENTIST

Hang gliding, or sky surfing, with a high-performance low-speed wing

Conducted by C. L. Stong

In the past few years amateur glider pilots in growing numbers have joined sea gulls and pelicans in soaring flight a few feet above the sloping beaches of U.S. coasts. Inland they skim the windward surfaces of sand dunes, ski slopes and other unobstructed hills. The rapidly evolving sport, which is known as sky surfing or hang gliding, makes about equal demands on the enthusiast's skills as a pilot and as an aerodynamicist.

The aircraft on which the sport is based vary widely in details of design. In general they are inexpensive and frequently homemade gliders of ultralight construction. The pilot is suspended by a harness.

Many of the gliders have as a common ancestor a triangular kite that was patented in 1951 by Francis M. Rogallo and his wife Gertrude. The kite has a glide ratio of better than **4 : 1**, meaning that in flight it goes more than four feet horizontally for each foot of drop. Most sky surfers first learn to fly on a Rogallo kite, which is often called simply a "wing." Many of them continue to fly the kite in preference to the hang gliders that were subsequently developed. The hang gliders have a higher level of performance than the kites, but they cost more and are more hazardous.

The Rogallo kite was conceived as part of a multimillion-dollar research project seeking a replacement for the parachutes that lower spacecraft to the earth's surface after they return to the atmosphere. As developed by the Rogallos the kite was a flexible delta wing. It was deployed, and its shape was maintained, by a set of tension lines, much as the shape of a parachute is maintained by its shroud lines.

Although the kite was abandoned by the National Aeronautics and Space Administration as a device for landing space capsules, primarily because it was difficult to stow and to deploy, a number of people interested in aerodynamics quickly conceived of modifications that would convert the craft into a serviceable hang glider. One of them was Michael A. Markowski of Marlboro, Mass., an aerospace engineer who had participated in the design of the Douglas Aircraft Company's DC-10 and subsequently did research for Sikorsky Aircraft. Markowski substituted an airframe of tubular aluminum for some of the tension lines of the Rogallo kite and fitted the structure with a control bar of aluminum tubing in the form of an inverted **Y**. A harness suspended the pilot in either a sitting or a prone position, from which he could grasp the control bar and exert force against it, thus shifting his weight to control the craft in pitch, roll and yaw [*see illustration on page 140*]. Markowski describes the evolution of his project.

"My interest in foot-launched gliders began as a spare-time activity, strictly for fun. After designing several rigid-wing biplanes and a monoplane of the flying-wing type I settled on the Rogallo wing as the most practical means of acquiring experience in the art of hang gliding. By taking advantage of several technical reports about the Rogallo wing that had been compiled by NASA I designed a full-scale model that by good fortune flew 'right off the drawing board.' Skimming the ground of a local ski slope turned out to provide great fun and excitement.

"I quickly mastered the art of taking off and gliding to a landing on foot without benefit of an instructor. It is an astonishing sensation to run along the

Michael A. Markowski taking off in his EAGLE-III ground for a few steps and rise effortlessly into the air. I scarcely realized that I was attached to a kite, because the craft represents only 20 percent of the gross weight of the pilot-glider system. Suspended prone below the glider, the pilot has a sensation of consummate freedom, perhaps because of his bird's-eye view of the terrain.

"My first hang glider had a sail consisting of .004-inch polyethylene. The machine weighed 40 pounds empty. It served me well for many ground-skimming flights until the sail began stretching, which degraded the glide ratio. As the temperature dropped during the fall of 1971 the plastic started to crack. After consulting a parachute rigger, who was familiar with sail materials, I replaced the polyethylene with rip-stop nylon. This material served for about a year before I retired it. Nylon is too elastic to maintain the curve of a good sail. I now use Dacron sailcloth.

"That prototype glider turned out to be only the first of a series of models. It was followed by both radio-controlled and free-flight models at reduced scale for investigating still other characteristics of the Rogallo wing. Fortunately this series of experiments alerted me to a startling fact. In the jargon of aerodynamics Rogallo kites are 'pitch-down divergent.' When the nose is lowered, the craft speeds up and the sail begins to luff, or flutter, at the trailing edge. The flutter advances toward the apex and the craft dives at an ever increasing angle until it strikes the ground!

"Having observed this disconcerting behavior in the small models, I decided to learn by experiment if the phenomenon would also occur in a full-scale glider. It did. The initial solution that came to mind was the addition of a horizontal stabilizer. I tried one. It worked, but it was awkward and was easily damaged in normal use. The final solution was simply suspending a weight some distance below the keel of the glider. This device worked well and introduced no structural problem.

"Sky surfing is an art. It can be mastered only by diligent practice. Of necessity the hang-glider pilot must be self-taught. The craft do not have dual controls for instruction, as conventional airplanes and gliders do. You start by learning a few additional details about your physical prowess and the reaction of the kite to the wind on level ground before you venture even briefly into the air. Like all young birds, you try your wings many times before leaving the 'nest.'

"The Rogallo kite with a control bar and a safety harness can be flown almost anywhere that suitably sloped terrain faces the prevailing winds. The novice is urged to begin his training at the foot of a small grass-covered hill or a sand dune with a slope of about 25 degrees. There should be a steady, uphill breeze of about 10 miles per hour. Be certain that the breeze is free of gusts that might dump you or abruptly lift you 20 feet or more.

"When you have assembled the kite near the bottom of the slope, put it down with its nose in contact with the ground and pointed directly into the wind. Walk completely around the craft and check the integrity of all its metal components, making sure that they are attached properly and show no bends, breaks or cracks. Simultaneously check the fabric for holes, cracks or rips. Assure yourself that all points of attachment of the sail are sound and that no grommets have pulled away from the fabric. All rigging should be tight enough to twang when it is plucked.

"When you have inspected the glider, grasp the upright members of the control frame, lift the kite above you and run forward with the nose pointed directly into the wind. As you run, tip the nose alternately upward and downward at increasing angles to sense the effect. Raising the nose will cause the sail to inflate,

Flying in the prone position

catch more wind and pull upward, thereby reducing the speed at which you can run. Lowering the nose has the reverse effect; it decreases the wind resistance and enables you to run faster. Note, however, that lowering the nose excessively causes the sail to luff, with the result that the wind acts on the top of the fabric and pushes the craft down.

"Continue practicing on level ground until you can unerringly predict and 'feel' exactly how the wing will react to every angle at which you hold the kite. Then strap yourself into the harness. Continue practicing on level ground until the harness feels natural.

"At this stage you can begin to work your way up the slope. At first, however, go up to an elevation of only two or three feet. *Always point the nose of the kite directly into the wind,* even when you carry the kite uphill. As you gain proficiency, the wind will even help you carry the craft up the slope. When you run downhill, hold the control frame near the bottom, so that the uprights pass close to your shoulders.

"Finally, from an elevation of 10 or 12 feet, begin running downhill with the sail barely inflated. As you pick up speed, push the control frame away from you somewhat, thus pitching the nose upward. If you have reached sufficient speed, you and the kite will rise into the air. If not, your forward motion will simply be retarded. In that case keep trying until you acquire the correct combination of forward speed and pitch angle for flight. Soon you will be skimming the ground.

"Once you have become airborne, maintain your fore-and-aft balance by shifting your weight. Push against the control bar to move your weight backward, which will increase the angle of pitch and thus reduce your speed. Pull the bar to shift your weight forward, thus decreasing the angle of pitch and increasing your airspeed. Shift to the right by exerting lateral force on the control bar to make a right turn, and do the opposite for a left turn.

"Remember, when you are sky surfing only two or three feet above the surface, it is better to err on the side of keeping the nose high and flying too slow than to pull the nose down sharply, which will make the craft dive and expose you to the risk of hitting the ground harder than you would like. Continue practicing near the base of the slope until ground skimming becomes second nature. Then work your way gradually to the summit.

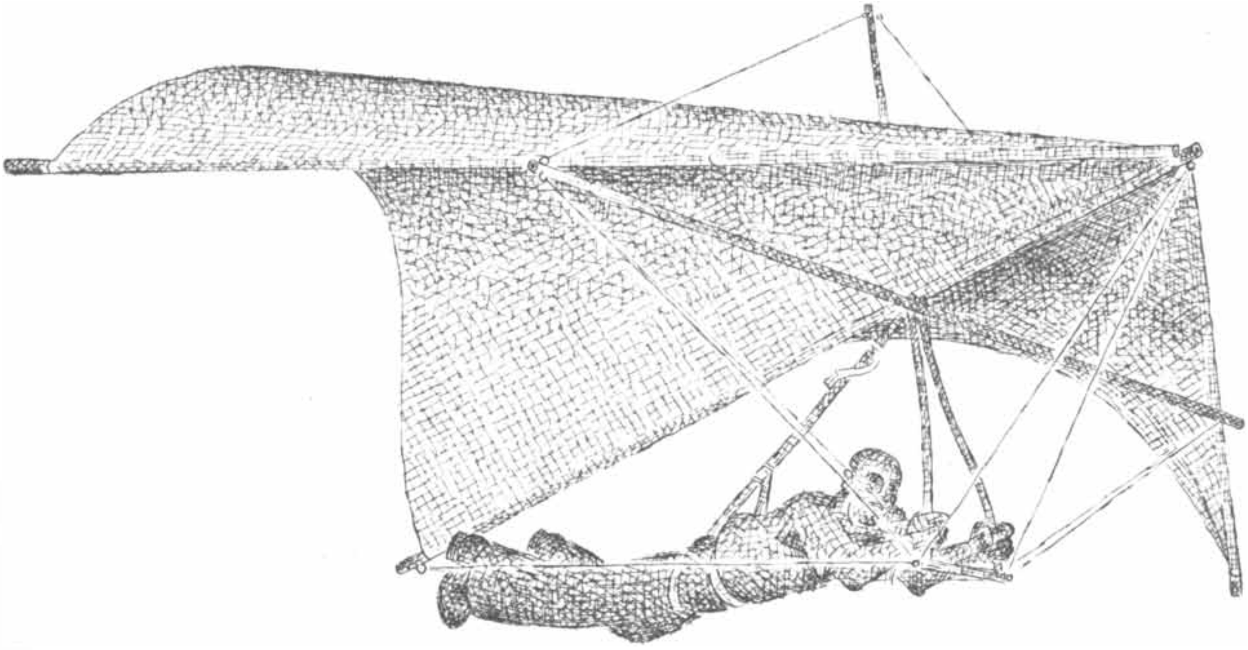
"Concentrate first on learning to maintain fore-and-aft balance, that is, pitch. If your craft starts drifting sideways, slow down and land as quickly as possible to avoid being blown into the side of the slope. As you gain skill in maintaining balance in pitch, begin to practice turns. Your first turns should be gentle, smooth and wide. Shift your weight by pressing sideways on the control bar very gently.

"All turns cause the kite to sink at a faster rate than when it flies in a straight line. You can compensate for this tendency somewhat by increasing the angle of pitch slightly during the turn. Avoid pushing the nose up to the angle at which the glider would stall, lose flying speed and drop toward the ground. The optimum angle of pitch can be sensed only with experience.

"As you progress to higher altitudes you will have occasion to make sharper turns and to maneuver in the three dimensions of space, but do not rush your learning process. Word has somehow spread that the Rogallo kite is an exceptionally safe craft. In actuality it is only relatively safe. Aerodynamically the wing is characterized by a very gentle stall, meaning that when the kite begins to lose flying speed, it tends to settle rapidly and to nose down slowly instead of going abruptly into a nose dive. Its performance is governed by the same laws of physics that affect other flying machines.

"The Rogallo kite is basically easier to fly than other aircraft because the pilot controls it by shifting his weight, which is a more or less instinctive action. However, weight shifting as a control technique has its limitations. The forces of the controlling moments remain constant, whereas the disturbing forces are squared with speed. Aerodynamic controls of the kind developed by the Wright Brothers enable one to fly safely in winds that would be unsafe for a hang glider of the Rogallo type. In other words, the Rogallo wing is no toy, and it can be a killer. Most beginners do not know how to land one safely from a stall at a height of, say, 40 feet. This maneuver can and must be learned by patient practice.

"Both wing loading and airspeed are important factors in the performance of



The modified Rogallo kite

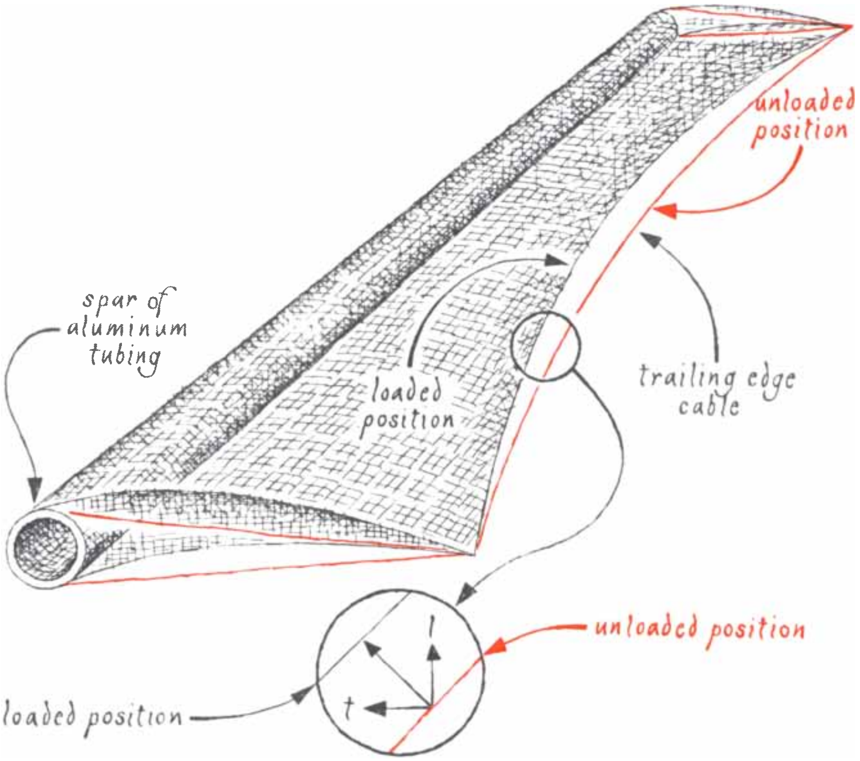
all aircraft. With these constraints in mind I designed a series of gliders based on the Rogallo wing for pilots weighing from 100 to 210 pounds. The gliders could be flown in wind speeds ranging from less than seven miles per hour to more than 20 miles per hour. In general the resulting craft have a maximum glide ratio of about 4.5 : 1 and a minimum sink rate of 450 feet per minute.

"Although one can have a lot of fun with a glider of this performance, the duration of flights is necessarily limited to minutes rather than hours. Sky surfing with a Rogallo wing is comparable in this respect to riding breakers at the beach with a surfboard. You enjoy a succession of one-way trips.

"I wanted a craft of higher performance that would enable me to glide continuously along windward slopes in both directions by making 180-degree turns at the ends. I also wanted to enjoy this high performance with a minimum of fuss. That meant developing a machine that one person could not only assemble and disassemble but also fold and load for transportation atop an automobile.

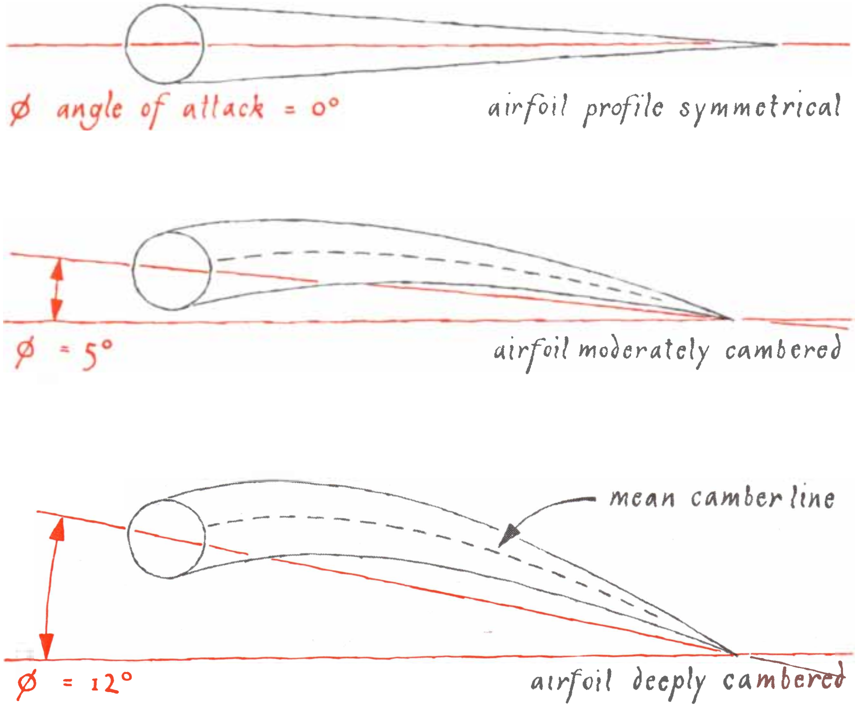
"An invention that promised to make possible such a glider came to mind when I recalled a college course in low-speed aerodynamics. During the course a 'sail wing' that had been developed at Princeton University in 1948 was described. The device had been developed as an advanced sail for boats. In 1952 it had been adapted for possible operation as an auxiliary lifting surface on ground-effect machines. Basically the sail wing consists of a tubular spar that supports the leading edge of a fabric envelope and a set of short, rigid booms at the tip and foot of the spar between which a slender cable is stretched to form the trailing edge of the wing [*see top illustration at right*]. The structure can be easily folded and stowed.

"The aerodynamics of the sail wing are both simple and impressive. The performance approaches that of conventional ribbed airfoils in terms of lift and drag. At zero angle of attack, when the plane of the wing lies in the plane of its motion, the sail wing assumes a symmetrical cross section that generates no lift [*see bottom illustration at right*]. At an increased angle of attack the surfaces of the wing form a cambered airfoil that does develop lift. Moreover, the camber deepens with increasing angles of attack, an aerodynamic effect equivalent to the effect of an automatic wing flap. A graph that depicts the wing's resulting force of lift slopes upward more steeply than that of a conventional 'hard' wing [*see top illustration on next page*].



The Princeton "sail wing"

"The overriding factor determining the glide ratio of an aircraft is the relation between the length of the wing and its width. The relation is termed the aspect ratio. It would be incorrect to state that doubling the aspect ratio of a wing would double the glide ratio of the aircraft, but increasing the aspect ratio greatly improves the glide ratio. (The glide ratio, incidentally, is numerically equal to the lift-to-drag ratio. The coefficient of induced drag is equal to $C_D = C^2_L/\pi AR$, in which C_D is the coefficient of drag, C_L is the coefficient of lift and



Response of sail wing to angle of attack

AR is the aspect ratio, which is equal to the length of the wing divided by its average chord, or width.)

"After reviewing the tabulated aerodynamic characteristics of several hundred airfoils I came to the conclusion that with a low-speed airfoil the coefficient of lift should increase at the rate of about 7.5 percent of the angle of attack; that the stall should become evident at a lift coefficient of about 1.6; that the larger the diameter of the leading edge the gentler the stall and the higher the coefficient of lift; that the deeper the camber the higher the coefficient of lift, and that the high point of the camber should be more than one third of the distance from the leading edge of the airfoil to the trailing edge. I observed with some surprise that these characteristics described exactly the results of experimental tests made on the sail wing, as reported by Princeton and by NASA. It was clear that the sail wing combined high performance in a simple, foldable structure that should function as the lifting surface of the ultimate hang glider.

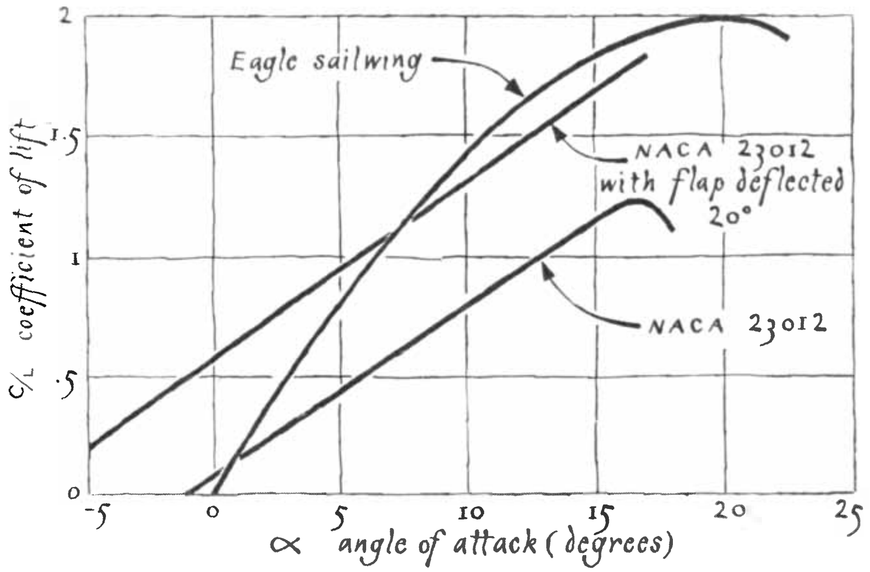
"I promptly built and flew a series of scale models, some of which were radio controlled. The first prototype, which I named EAGLE-I, had a wingspan of 40 feet and weighed 70 pounds. The wing lay in a single plane, that is, it had no upswept dihedral angle. The aspect ratio was 8 : 1, the area was 200 square feet and there was no sweepback. The tail had a conventional rudder, an elevator and a horizontal stabilizer. For lateral control I substituted 'spoilers' for ailerons. The spoilers were small flaps that could be lifted by control cables to create drag near either tip of the wing.

"Initial ground and 'sail inflation' tests were made by the same procedure that applies in learning to fly a Rogallo wing. When the machine was lifted into the slightest breeze, the sail wing assumed exactly the predicted contour. Flight testing was begun on low, shallow sand dunes. The first few ground skims indicated that the elevator gave perfect control of pitch; they also helped me to shift the suspension harness to the proper balance point. The rudder proved to be of some value in controlling yaw, but the spoilers were ineffective. The fabric of the control surfaces became grossly distorted under load, which necessitated redesign.

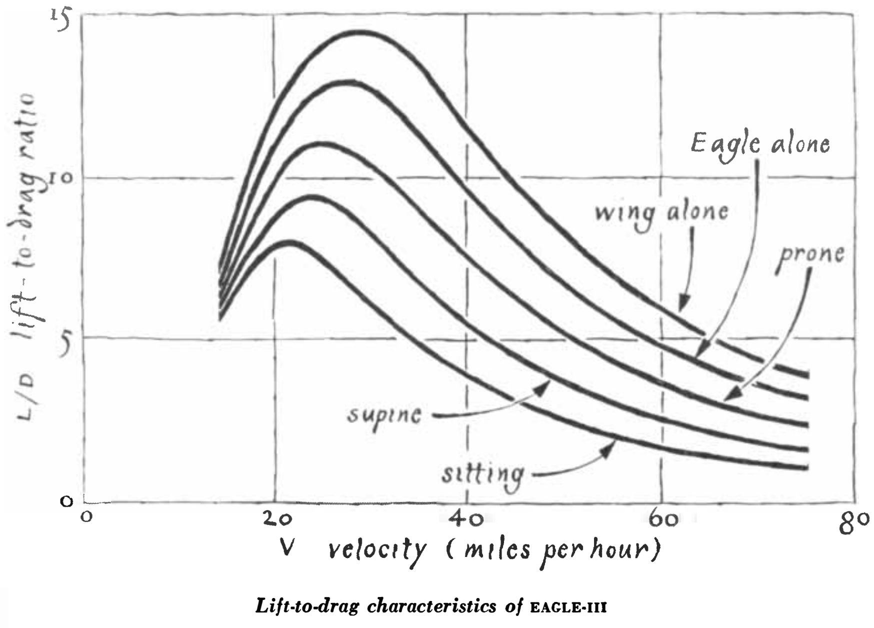
"On the other hand, EAGLE-I was fully stable in flight and had a glide ratio of close to 10 : 1. Although I enlarged the spoilers 50 percent, the modification was never tested. A critical inspection of the craft indicated that a number of design simplifications could be made in the construction of the airframe. Moreover, I decided to achieve lateral control by the Wright Brothers' system of wing warping. The result was a much cleaner design that weighed 63 pounds.

"Flight tests proved that the machine was a high-performance glider. I took one step in a breeze of 10 miles per hour, pulled back the control stick and was lifted almost straight up. It was a fantastic sensation. The lateral control system, however, was still ineffective. Moreover, it was apparent that the span and area of the wing were much too large to handle in anything more than a flat calm.

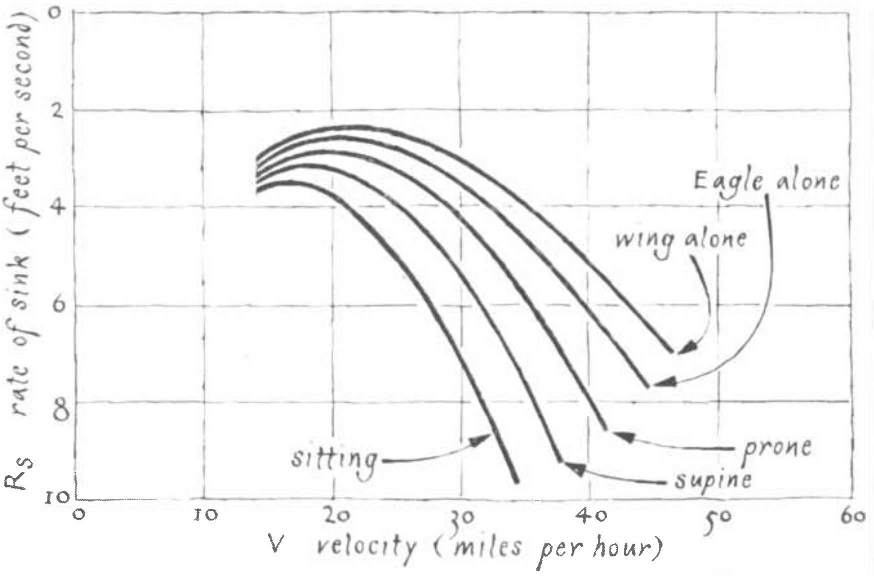
"The additional tests and modifications led to the construction of EAGLE-II. This machine weighed 75 pounds and had a wingspan of 34 feet and a wing area of 158 square feet. The removal of a few bolts made it possible to fold the craft easily for transportation on a car top. EAGLE-II had a set of pulleys and control cables and a control stick that enabled the average kite pilot to make an easy transition to the high-performance craft. The wing, with an aspect ratio of 7.25 : 1, was nonswept. On test it developed a lift coefficient in excess of two, with a gentle stall. Indeed, the 'stall' would be more aptly described as a 'mush,' that is, a parachutelike settling rather than a dive. When the control stick was pulled backward slowly, the angle of attack and the camber of the wing increased simultaneously. The action was followed by a steep but slow descent. Response to the controls re-



Aerodynamic characteristics of various airfoils



Lift-to-drag characteristics of EAGLE-III



Sink ratio of EAGLE-III

mained good up to stall but was sluggish in yaw in the 'mush' condition.

"EAGLE-II had a wing taper of 3 : 1 and a dihedral angle of eight degrees. At first I rigged the craft for a dihedral angle of only two degrees, but it tended to skid too much in yaw. I also rigged two degrees of washout in the tips of the wings to guard against tip stall. (The tips of the wings are twisted to reduce the angle by two degrees.) The structure was stressed by a test load equivalent to six times the force of gravity with a safety factor of 1.5.

"The tail structure consisted of a frame of aluminum tubing that supported fabric of rip-stop nylon. In time it became apparent that this material stretched excessively. The resulting exaggerated camber degraded the glide ratio. As I have mentioned, I now use stabilized Dacron sailcloth.

"Both the rudder and the elevator were balanced aerodynamically, that is, they were hinged slightly forward of the quarter-chord line of the control surface. The horizontal stabilizer and the elevator were removable as a unit from the keel of the airframe. Forces developed by the rudder and the dihedral-angle structure of the wing combined to produce a roll-yaw couple that helped make turns easy.

"Hinged structures of covered aluminum tubing that formed the wing tips both warped the wings and served as ailerons. I named them 'warperons.' Coupled directly to the rudder, they deflected the wing tips differentially just enough to produce coordinated turns.

"Exhaustive flight tests have now been completed on EAGLE-III. In theory the maximum glide ratio of the machine alone is substantially better than 10 : 1. This performance is of course degraded by the presence of the pilot, who creates forces of drag but no lift. The amount of drag introduced by the pilot depends on his position. Computations indicate that when the pilot flies in the prone position, the glide ratio of EAGLE-III approaches 11 : 1. In the sitting position it is almost 8 : 1. These figures assume an optimum airspeed of about 24 miles per hour. The sink rate also varies with the amount of drag induced by the pilot; it is about 200 feet per minute. The performance of EAGLE-III is therefore compatible not only with sustained flight on the windward side of sloping beaches and comparable terrain but also with cross-country gliding.

"Essentially EAGLE-III is an ultralight, high-performance monoplane with a variable-camber wing. It is possible for the pilot to stall the craft in flight, but the stall is gentle compared to that of a standard airfoil, such as the NACA 23012. Beginners should not attempt to fly EAGLE-III, but pilots who have mastered hang gliding with a Rogallo kite make an easy transition to this high-performance craft.

"Both the Rogallo series of sky-surfing kites and EAGLE-III are available commercially as a kit of raw materials and also as a prefabricated kit that includes machined fittings and certain components that are relatively difficult to make, such as the sails. The kits, together with detailed working drawings and complete instructions for fabrication and assembly, are supplied by Man-Flight Systems, Inc. (P.O. Box 872, Worcester, Mass. 01613)."

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION (Act of August 12, 1970: Section 3685, Title 39, United States Code). *1.* Title of publication: Scientific American. *2.* Date of filing: October 7, 1974. *3.* Frequency of issue: monthly. *4.* Location of known office of publication: 415 Madison Avenue, New York, N.Y. 10017. *5.* Location of the headquarters or general business offices of the publishers: 415 Madison Avenue, New York, N.Y. 10017. *6.* Names and addresses of publisher, editor and managing editor: Publisher, Gerard Piel, 415 Madison Avenue, New York, N.Y. 10017. Editor, Dennis Flanagan, 415 Madison Avenue, New York, N.Y. 10017. Managing Editor, none. *7.* Owner: Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Atwell & Co. (for Edwin H. Baker Pratt), United States Trust Co. of New York, 45 Wall Street, New York, N.Y. 10005; Janet R. Becker, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Buckley & Co. (for Elinor Gottlieb Mannucci), c/o Fiduciary Trust Co. of New York, P.O. Box 924, Wall Street Station, New York, N.Y. 10005; Cudd & Co. for (a) Helen R. Snellenburg, Morris Duane and Raymond J. Armstrong, Trustees u/a 2/17/66 F.B.O. Helen R. Snellenburg and (b) Fred Grossman, Successor Trustee u/a 10/15/51 F.B.O. Robert L. Rosenwald, The Chase Manhattan Bank, 1 Chase Manhattan Plaza, New York, N.Y. 10015; Martin M. Davidson, 11 River Lane, Westport, Conn. 06880; Richard E. Deutsch, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Bayard Ewing, c/o Industrial National Bank, Trust Department, 100 Westminster Street, Providence, R.I. 02903; Dennis Flanagan, c/o Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Sidney R. Garfield, M.D., Kaiser Foundation Hospitals, 1924 Broadway, Oakland, Calif. 94612; William T. Golden, Room 4201, 40 Wall Street, New York, N.Y. 10005; Kidder, Peabody & Co., Inc., 10 Hanover Square, New York, N.Y. 10005; Nathan W. Levin, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Frasier W. McCann, Room 726, 10 Rockefeller Plaza, New York, N.Y. 10020; Donald H. Miller, Jr., c/o Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, c/o Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Barbara K. Rosenwald, P.O. Box 3235, Norfolk, Va. 23514; Julius Rosenwald II, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Lessing J. Rosenwald, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Rowe & Co. (for Rhode Island Charities Trust), c/o Industrial National Bank of Rhode Island, Trust Department, 100 Westminster Street, Providence, R.I. 02903; Stanley Schaefer, c/o W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104; I. M. Scott, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Joan R. Scott, c/o The Chase Manhattan Bank, P.O. Box 1508, Church Street Station, New York, N.Y. 10008; Dorothy D. Spivack, Room 804, 921 Bergen Avenue, Jersey City, N.J. 07306; John Hay Whitney, Room 4600, 110 West 51 Street, New York, N.Y. 10020. *8.* Known bondholders, mortgagees, and other security holders owning or holding 1 percent or more of total amount of bonds, mortgages or other securities: none. *9.* For optional completion by publishers mailing at the regular rates (Section 132.121, Postal Service Manual). 39 U.S.C. 3626 provides in pertinent part: "No person who would have been entitled to mail matter under former section 4359 of this title shall mail such matter at the rates provided under this subsection unless he files annually with the Postal Service a written request for permission to mail matter at such rates." In accordance with the provisions of this statute, I hereby request permission to mail the publication named in item 1 at the reduced postage rates presently authorized by 39 U.S.C. 3626. (Signature and title of editor, publisher, business manager, or owner) Donald H. Miller, Jr., Vice-President and General Manager. *10.* Not applicable to this organization. *11.* Extent and nature of circulation: *A.* Total number of copies printed (net press run): average number of copies each issue during preceding 12 months, 638,956; actual number of copies of single issue published nearest to filing date, 656,114. *B.* Paid circulation: *1.* Sales through dealers and carriers, street vendors and counter sales: average number of copies each issue during preceding 12 months, 94,515; actual number of copies of single issue published nearest to filing date, 94,000. *2.* Mail subscriptions: average number of copies each issue during preceding 12 months, 480,715; actual number of copies of single issue published nearest to filing date, 500,005. *C.* Total paid circulation: average number of copies each issue during preceding 12 months, 575,230; actual number of copies of single issue published nearest to filing date, 594,005. *D.* Free distribution by mail, carrier or other means: *1.* Samples, complimentary and other free copies: average number of copies each issue during preceding 12 months, 23,859; actual number of copies of single issue published nearest to filing date, 16,291. *2.* Copies distributed to news agents but not sold: average number of copies each issue during preceding 12 months, 32,422; actual number of copies of single issue published nearest to filing date, 36,045. *E.* Total distribution (sum of *C* and *D*): average number of copies each issue during preceding 12 months, 631,511; actual number of copies of single issue published nearest to filing date, 646,341. *F.* Office use, leftover, unaccounted, spoiled after printing: average number of copies each issue during preceding 12 months, 7,445; actual number of copies of single issue published nearest to filing date, 9,773. *G.* Total (sum of *E* and *F*—should equal net press run shown in *A*): average number of copies each issue during preceding 12 months, 638,956; actual number of copies of single issue published nearest to filing date, 656,114. I certify that the statements made by me above are correct and complete. Donald H. Miller, Jr., Vice-President and General Manager.

BOOKS

An annual review of children's books for the Christmas season

by Philip and Phylis Morrison

This year even more than last we note a real dearth of titles in the physical sciences. The reasons for this shift are evident in the adult world, but we submit to authors and publishers that children are an unending procession. Not only are they free from our waves of taste; they need and deserve a sample of all the modes in which their predecessors tried to understand the world.

Technologies

The Unknown Leonardo, edited by Ladislao Reti. Designed by Emil M. Bührer. McGraw-Hill Book Company ($39.95; $34.95 if ordered before January 1). During the years when Columbus planned and carried out his voyages to the Indies, the artist-engineer Leonardo da Vinci served the Duke of Milan. The wonderful résumé the 30-year-old Florentine sent to the Duke opens this volume. We read nine proposals for warlike devices: "I will make covered chariots"; "I have kinds of mortars"; until finally: "In time of peace [I am] the equal of any other in architecture.... I can carry out sculpture ... and also I can do in painting whatever be done, as well as any other, be he who he may." He could too.

This fat volume, with 800 illustrations, 116 in full color, is the general reader's share of the excitement produced by the chance find a few years ago of two small notebooks of Leonardo's in his own mirror hand, for 135 years misshelved in the National Library of Madrid. The find has increased by a fifth our existing store of documents from Leonardo; we can estimate that we now know a quarter of all he wrote or drew. The 10 scholars whose chapters make up this exceptionally rich book discuss the whole of the master's life and work, with special concern for what is newly seen in the Madrid Codices. The pages are so chock-full of big reproductions of paintings (the *Last Supper* and the *Mona Lisa* included) and a crowd of sketches, studies, single letters, maps, instruments, churches, fortresses and models that the volume is a celebration, as it is meant to be, not only of the enigmatic Leonardo but also of the life of the Renaissance. It will ornament any library where young people might browse through or pore over its images and pages.

Too varied to summarize, the Madrid find illumines one task of Leonardo—unfinished, like so many—that is a marvel for the eye and the mind. It was the project for a bronze horse bearing the Duke's noble father, Francesco Sforza, in armor. But what a horse: enormous, four times life-size! Leonardo was seized with the scheme, devised for it entirely new means of casting in one piece, derived from his artillery experience, drew many plans of clever devices for carrying out this remarkable feat and finished a full-scale model in some earthen material, perhaps a molding stucco. The model was unveiled at the wedding of Bianca Maria Sforza to the Emperor Maximilian in 1493. The horse demanded 79 tons of bronze, but the bronze set aside for the work was used instead for cannon (while Leonardo was painting the *Last Supper*). Even so Milan fell to the French; the model too was lost. No such large bronze has ever been cast since. Leonardo wrote in the Madrid Codex: "Epitaph/If I could not make/If I."

There is here a 15th-century bicycle design (maybe) and a true worm gear among the clear, connected analytic account of mechanisms that fills one Madrid Codex. There is a set of deep aphorisms, an analysis—even two—of his paintings, and more. The aura of unstinting creativity that pervades this new telling of an old story glows as much with melancholy as it does with joy. All of it is a surefire treat, for teen-aged readers and older.

Solar Energy: Technology and Applications, by J. Richard Williams. Ann Arbor Science Publishers Inc. ($9.95). With very few formulas but a wealth of tables, graphs, flow charts and conceptual system diagrams, this Georgia Tech engineer presents in a small compass, clearly and pleasantly, the story of ways and means of getting power from the sun. He rightly begins with the availability of sunshine, times of clear sky and angles at various sites, and he goes on to collectors in general, their use for space- and water-heating, then to electric power, both thermal and photovoltaic, to furnaces and on a larger scale to hay burning, windmills and sea generators. Farthest out of all are satellites in geosynchronous orbit with microwave power transmission to the earth. More of the text and more of the attractive diagrams are devoted to the uses nearest at hand. In Japan, we read, 2.5 million solar hot-water heaters are currently in use. A scheme employing a farm full of moving heliostat mirrors, all aimed at a black boiler mounted high on a central tower, is evolving in the Southwest on the pilot-plant scale. The Coast Guard now uses a photovoltaic system for some solar-powered buoys (saving costly trips to sea to replace fuel stores), which is on the market as a set of 1.5-watt modules, each easily giving in sunlight some half an ampere at a couple of volts, to charge batteries. ("Wires are used to prevent seagulls from landing, but nothing is done about snow.") Dust, dirt and snow seem to be no large problem.

This is in no way a how-to book. It is a sensible engineering survey, with enough data to make the problems clear and the proposed solutions real and understandable. It is an optimistic but not a wishful piece of work; it should be a useful and enjoyable book for readers at high school age and above who like to understand design and to look a little ahead.

Bulldozers, Loaders, and Spreaders: A Book about Roadbuilding Machines, by the 1973 Ninth Grade English Class of the Green Vale School, with Nancy Arnauot, teacher. Doubleday & Company, Inc. ($4.95). Tight money and too much concrete already in place surely reduce opportunities, but there must still be families or schools where some youngster has not long ago passed by an active construction site. What do those heavy machines do? Here they are at work in photographs mainly supplied by the proud manufacturers. Here is a powerful bulldozer plodding through deep mud. A wheel-loader carries five great logs as though its hand held a small bundle of sticks. Next are the strong prehensile back hoe and the long low scraper. A grader is seen, and dump trucks flock to haul away the soil it has left behind. The road itself, that great ribbon of concrete, is being placed by the slip-form paver, which spreads, smooths, levels, cuts the edges squarely and even finishes the surface. It is a pleasure to see the corporate authors, 12 students and a teacher, who together planned, wrote and assembled the images that make up the book. That class was a talented and effective group of young girls.

The Dome People, by Ruth Howell. Photographs by Arline Strong. Atheneum ($5.25). Buckminster Fuller's geodesic domes are practical tools in the kit of the expert builder. Easy to build in modest sizes, they are also the objects of a benign cult of enthusiasts. This pleasant little book, with pictures taken throughout the task, relates how the young people of a community near some big city built their own dome as a center for teen-agers. It is a forgiving design, although it has its own logic. They scrounged design and materials, talked a small piece of park out of the parks department, pitched tents and started to work. The site by law had to be guarded; they brought sleeping bags and camped. The foundation was mixed, poured and leveled. A floor frame grew. Then the dome started up. Hubs were bolted to the two-by-four struts. A mistake necessitated a restart, but they kept at it. "The Dome is more than just struts. It's all of us. It's poetry," one girl said. Mountain-climbing experts led the plywood panel crews, way up there on the dome top. Calking, doors, Plexiglas hexagonal windows, ditches, piping, wires and plenty of paint came next. (A usable geodesic dome is more than struts.) One year after the start more than 100 people gathered to "dedicate this building to ourselves." The dome is alive; in it people meet and make music and theatre. The dome people maintain it too, they way they want it to be. ("Building the Dome had been the hardest and most rewarding thing in their lives.")

Green Grass and White Milk, written and illustrated by Aliki. Thomas Y. Crowell Company ($3.95). For readers in their first reading years this cheerful and colorfully illustrated book presents what amounts to a flowchart. Green grass; golden hay; four stomachs in a cow to process her cud; two milkings a day from the udder, a bag with four teats; daily pickup from farm cooler into the refrigerated truck. Off to the dairy, where a flowchart takes an entire page and runs from truck to pasteurizers to cooler to homogenizer to carton-filling line. The half-pint carton appears in school at lunchtime. "Homogenized" means "the same all the way through." (More's the pity.) Goats and sheep get their due, but off in some European pasture. Then you find graphic recipes for instant butter from cream, and for your own yogurt from milk and a half pint of store-bought yogurt for the culture. (No explanation of yogurt is given, however.) This is a tentative and gentle look into causes, meant for a wide-eyed audience.

HANG GLIDING: THE BASIC HANDBOOK OF SKYSURFING, by Dan Poynter. Published by the author, Box 4232, Santa Barbara, Calif. ($5.95). Hanging by the armpits from a big kite, the enthusiasts of this elegant but quite hazardous sport (a dozen deaths per year) sail off sea cliffs and sand dunes of both coasts and in between, even at Kitty Hawk itself. They show "serious intent to commit aviation," as the author of this, their first book of the century, begins. They usually fly at about 20 miles per hour in structures of aluminum tubing and 150 or 200 square feet of fabric or film that weigh around 30 pounds. Control is classically by shift of weight alone (although this is changing), and launch and landing are powered solely by the pilot's legs. (One pioneer is ski-borne.) Such elemental gliders can soar as well as glide slowly down the slopes, but so far they frequent the beach and ridge, and only occasionally gain altitudes much beyond the summit of the launch ground.

There are competitions and clubs, and a small industry of design and kit suppliers. This paperback catalogues three dozen manufacturer's models. Most of these are limp delta sails, patterned on a military parawing development, but some are biplane boxes looking like the Wrights', and a few are high-performance monoplanes with control surfaces. There are chapters on launch and flight, on the history of the art, on construction, on such hardware, fabrics, aluminum tubes and fittings and on the law. (No license is needed.) There is also a brief introduction to design. It is exciting reading, even for a born spectator. The pioneer Otto Lilienthal began flight in 1891. He died in 1896, falling after he had lost control of his latest monoplane at 50 feet. His last words ring: "Sacrifices must be made." The present sport needs no more sacrifices; it demands instead rigorous good sense and detailed engineering, both of which are present in this small book, which is all the same to be opened with care!

THE JUGGLING BOOK, by Carlo. Vintage Books, a Division of Random House ($2.95). Meticulously, step by step, the first six lessons, each consisting of several pages with diagrams, begin with the choice of balls. (Lacrosse balls are solid, bounce well, and weigh enough to force your hand to close partway when the ball strikes your palm.) They unfold through basics, body position, up a carefully self-checking sequence of moves until the three balls arc fluidly between your hands in that unending cascade, with its implications of infinity, that is the foundation stone of this ancient circus art. The lessons continue; the cascade is consolidated and elaborately varied. The feat is extended to hoops (up to 10 or 11) and clubs and then to the use of five balls. "One person in a hundred is a 'natural' three-ball juggler. I have never seen a 'natural' five-ball juggler." This remarkable book (for 12-year-olds up to young adults) by an ex-geologist, now a writer and a philosophical adept at circus performance and teaching, has the high aim of taking juggling "from the circus ring to the front yard." Here the ruling image is not "the wiseacre . . . in tuxedo or tails equipped with gimmickry and frills" but the simplicity of basic shapes, direct personal involvement, hard work, dedication, the spirit of a demanding craft, not a form of guileful deception by a master of the hidden. Anyone can do it who will. The work is a demonstrated and popular success in its endeavor, and around it and other efforts there is growing a spontaneous public pleasure in mime, street theater and similar circuslike performing arts. (Lacrosse balls are in short supply!) There is a deep kinship between the world seen by science and the world of the juggler, who learns by thought and hand how to master the cascade by conformation to its simple dynamics, a timed interplay of hand-eye coordination and free fall. We have the word of Joseph Needham that for most people long ago, before science and technology had much impinged as special occupations on everyday life, "there was not much difference between jugglers, alchemists, mechanicians, star-clerks, leeches and all other practitioners of magic and gramarye."

Looking

SHADOWPLAY, by George Mendoza, with Prasanna Rao. Photographs by Marc Mainguy. Holt, Rinehart and Winston ($5.95). "As a sickly child, I was forced to stay in my bed. At night, when everyone was asleep, I practiced my shadows with a hurricane lamp." So recalls Prasanna Rao of Mysore, whose eloquent positioned hands, not careless in the tone of a single joint, are here projected into more than two dozen remarkable shadows, the rabbit and bird of every deft entertainer raised to a high pitch of subtlety of form, and extended remarkably to portray without props Napoleon, a taxicab, a rhinoceros, a man laughing and two people in the park. For every shadow there is shown in full light on a facing page, often from a slightly different aspect, the artist's hands in the proper position. Motion is lacking to us (although not to the nationwide audience of children's television this season). The details that give the nuances of reality to Rao's shadow pictures are not at all easy to emulate. The sickly child notwithstanding, one forms the clear impression that Rao's hand control and insight bear some relation to the high tradition of Indian dance, where hand gesture enters so strongly and subtly. Two pages showing "hand exercises"—no shadows—add evidence for that view. Mendoza, who put the book together, is a film writer who has already made splendid books of photographs out of the related art of the great mime Marcel Marceau.

WHAT IS IT? A BOOK OF PHOTOGRAPHIC PUZZLERS, by Joan Loss. Doubleday & Company, Inc. ($4.50). ZOOMING IN: PHOTOGRAPHIC DISCOVERIES UNDER THE MICROSCOPE, by Barbara J. Wolberg. Photographs by Lewis R. Wolberg. Harcourt Brace Jovanovich ($7.75). Since the days of Galileo the small world has been viewed through the microscope. Big full-page black-and-white photographs add drama to the reproduction of that world and allow an approach to the real task of lens viewing. These two books make such an approach agreeable. The puzzler book just crosses the threshold of the strange. In it the photographer has shown us 20 blown-up fields of household commonplaces, hard to recognize once we become as Lilliputian viewers, seeing our own surroundings 10 to 30 times bigger. Is that line of rods a comb? That pile of rugs the pages of a book? These are witty and attractive views achieved with a single-lens-reflex camera and bellows; in each case the object is seen overleaf from a normal distance, suddenly all too familiar.

Once across that threshold, you can enter a far stranger world through the microscope, at magnifications from a couple of hundred times up to a couple of thousand. In *Zooming In*, for readers more experienced with the abstract, nearly every one of the 60-odd pages displays one or more magnified views, well chosen and handsomely photographed and reproduced. The lower powers make the familiar world strange; at higher powers we can seek undreamed-of wonders: paramecia, chromosomes in tissue cells, the conduits of a corn root section, liquid crystals and the grains of etched brass. Although the book is only one of many places one can find such photographs, its clarity and the care in selection of the samples are rewarding. For example, one sees not merely the textbook pattern of chromatin in a single cell but an entire cluster of cells, wall and all, each one carrying that imperious ideograph within.

SEE WHAT I AM, by Roger Duvoisin. Lothrop, Lee & Shepard Company ($5.95). Max is a good-looking kitten, if he does say so himself. That is not clear, however, when you see him all yellow, blue, red or even black. In white he disappears save for his eyes and claws, and in the binary green of blue plus yellow, or the purple of red plus blue, he is in danger of becoming a laughingstock among mice. Once in the full harmony of four-color superposition on the clean white page, Max is himself: "white with yellowish-reddish, black spots and blue eyes." You *are* good-looking, Max, and you present lightly and well, with a bit of sly sermon, how a few colors form the richer images we see. For families with new readers or small listeners, here is a pleasant and meaningful book.

Sky, Earth, Sea

SUMMER GOLD: A CAMPER'S GUIDE TO AMATEUR PROSPECTING, by John N. Dwyer. Charles Scribner's Sons ($5.95). Gold is where you find it. Finding it has ignited the imagination of men for centuries, as the history of California and Alaska will witness. John Dwyer is a gold hunter, but he does not intend to leave his bones to whiten in the desert, nor even to take the unneeded risk of using mercury intimately in cleaning his show of fine dust. He is an amateur; he takes only the equipment he can carry on his back, and he hopes someday to find that glory hole with a nugget too heavy to lift. Until then he enjoys the search along the mountain stream, the little vial full of glittering dust, the unending hope and "a bonanza, if not of gold, then of fun . . . in places where the average tourist is seldom seen."

The book gives a summary history of U.S. mining, describes the sites, mainly Western but some even in the Appalachians (where placer mining is hopeful), and guides the novice (with photographs) to the right use of that classical density separator, the simple gold pan. The author is a thoughtful man. To learn to use the gold pan, practice on a mix made up with gravel and small lead shot. Never cook in your gold pan; traces of grease may cause fine gold to float away. Don't forget to check for the signs of old stream beds; you want the eroded results of an enduring flow of water, not the swiftest currents of today. You can use electronic metal detectors, mainly for lost and buried coins and jewelry but also for placer gold. The miner's main tools are the shovel and the pick. He tells you where to buy maps, tools, detectors and supplies, who will assay your finds and how to stake a claim, or at least how to make the first moves in the complex game of coming to own a claim. A list of famous mines is one first guide to where to go; these were lode mines, not placers, but streams in those districts ripple with hope. Who would not be off to the Gold King mine in the Eureka district of Animas County in Colorado? Don't talk about foreseen treasure finds, even to your best friend. You have a rendezvous with Lady Luck; may the color be plentiful in your pan.

Language and Mathematics

HANDTALK: AN ABC OF FINGER SPELLING AND SIGN LANGUAGE, by Remy Charlip and Mary Beth. Photographs by George Ancona. Parents' Magazine Press ($4.95). Mary Beth is a young woman, an actor in the National Theatre of the Deaf, whose hands and face, now cheerful, now grave, but always deft and serious, speak out expressively from the colorful pages of this handsome, unique alphabet book. This is not mime, graphic though Mary Beth's gestures are. It is a structured language into which we are taken a very little way. What we see on these two dozen close-ups of Mary Beth and a couple of friends are formal words of signing, or American Sign Language (what a few bright chimpanzees such as Washoe have been learning lately, by the way). Each sign is a word, whose initial letters build up an alphabet ABC as usual. It is not easy reading, but it is fun; help is provided, of course, in the form of small inset photographs, each of which shows a letter, spelling out the word, not in the rapid ideographs of signing but in the slow, easily decoded symbols of alphabetic finger spelling, for which a key is given in the endpapers. This is a novelty and a beauty, a tribute to the talented collaborators—actress, designer and photographer—who made it so happily and well. There is much good humor and gaming; the center fold shows the two palms (one smeared with jelly and one with peanut butter); "when you close these pages these hands make the sign for sandwich." Many delighted readers will, as we have, learn the little shorthand sign made with three fingers: I L Y, Mary Beth. Kids will find in this book games, secrets, puzzles, almost a dance. There are more than 5,000 signs in this language and hundreds of thousands of people who know them. If you want to learn more, ask a deaf person!

THE GOLDEN MEAN: MATHEMATICS AND THE FINE ARTS, by Charles F. Linn. Doubleday & Company, Inc. ($4.95). Cast up on the wilder shores of the ocean of mathematics is some pretty strange flotsam. Not much is as odd as the "aesthetic measure" of the genuinely distinguished mathematician G. D. Birkhoff. Just 50 years ago he devised a scheme for "the measurement of beauty" that, for example, rated the square highest among 90 forms he first tried. The measure was the ratio of order to complexity, order being judged by five factors: symmetries, equilibrium and so on (with a negative score for a fudge factor called unsatisfactory form). Birkhoff worked, however arbitrarily, in a most ancient tradition. The Greeks found integer beauty first perhaps in the musical scale, and then geometrical beauty in certain self-similarities of spatial form related to the "golden section." The Parthenon gives some evidence for the virtues of proportions such as 1.62 : 1. Classical architects went beyond that and introduced those refinements—or deceptions—of slight bends here and there, by no means accidents of construction. Here in evidence is a fine photograph of a broken line in the plan of the choir of the cathedral of Siena. Nor can one doubt the group theory manifest in the wonderful plane-filling patterns of much art, its highest point perhaps the tiled walls and gratings of western Islam.

All these subjects, and more modern ones such as the logical illusions of Maurits Escher and the flood of computer graphics, link art to mathematics by a chain of inference far short of proof but full of suggestion. In this ingratiating little book, with plenty of pictures and hardly a single formula, a disarmingly cheerful argument, musing and skeptical, introduces a young reader to the history and ideas of this tantalizing frontier. The author, an experienced writer and teacher of mathematics, never strains credulity or insists on what is plainly uncertain. His work is a model of exposition of tentative, enticing, imaginative ideas, just what the subject matter requires. It is a fine first book for anyone on either edge of that boundary who wants to consider setting out from it some day. Perhaps some young reader will follow these leads to build a real highway.

CIRCLES, TRIANGLES AND SQUARES, by Tana Hoban. Macmillan Publishing Co., Inc. ($4.95). It is a more than plausible theory which holds that the mind's power of abstraction, made concrete by the words we use, forms delightedly around a nucleus of real examples of a new-found class. That theory is most persuasive when it is applied to the young, and this artist of the lens has given children at the edge of reading a little wordless treasury of the concrete, so sharply and wittily seen that the classifying urge should be beyond resisting. Circles? They are pipe bundles, eyeglasses, hoops, typewriter keys, baby buggy wheels. Triangles? They are a cocked paper hat, a director's chair, an elevated train structure, a cookie, a park swing, a house of cards. Here are real sets, vivid, unexpected and yet exactly as we know them. More than two dozen big black-and-white photographs fill this book with images of convincing meaning. It lies squarely (may we say) on the frontier of language and mathematics, without a word of text.

Human Beings: Their Biology and Cultures

HUNTERS OF THE WHALE: AN ADVENTURE IN NORTHWEST COAST ARCHAEOLOGY, by Ruth Kirk, with Richard D. Daugherty. Photographs by Ruth and Louis Kirk. William Morrow and Company ($5.95). Let the rain drip from the big spruces and the slow roll of the sea fling breaker and foam endlessly across the beach. This book is a sunny story just the same. Professor Daugherty grew up along the shore of the Olympic Peninsula and became an archaeologist at Washington State University. At the very corner of the great forested island, where it stands farthest out to sea, are the Makah Indians, living now as they have always lived. By good luck, hard work and understanding he and they are working together at excavating, before the sea claims it all, a beach settlement. For long centuries Makah people lived by the riches of their sea world, chief among them the whales—all kinds of whales. Here is a photograph of a whaling canoe making fast about 1900; here too a photograph of the heavy whale skulls and vertebrae hauled for some reason (Was it pride?) centuries ago up to the house they are now excavating and reassembling piece by piece. Everyone has had a hand in this work. Makah high school girls work alongside the students and scientists from the university, while the senior citizens try to puzzle out artifacts from what they can recall of the old days. Marine helicopters and Coast Guard surveyors play an indispensable role in the long years of meticulous work against the powers of time, mud slump and storm wave.

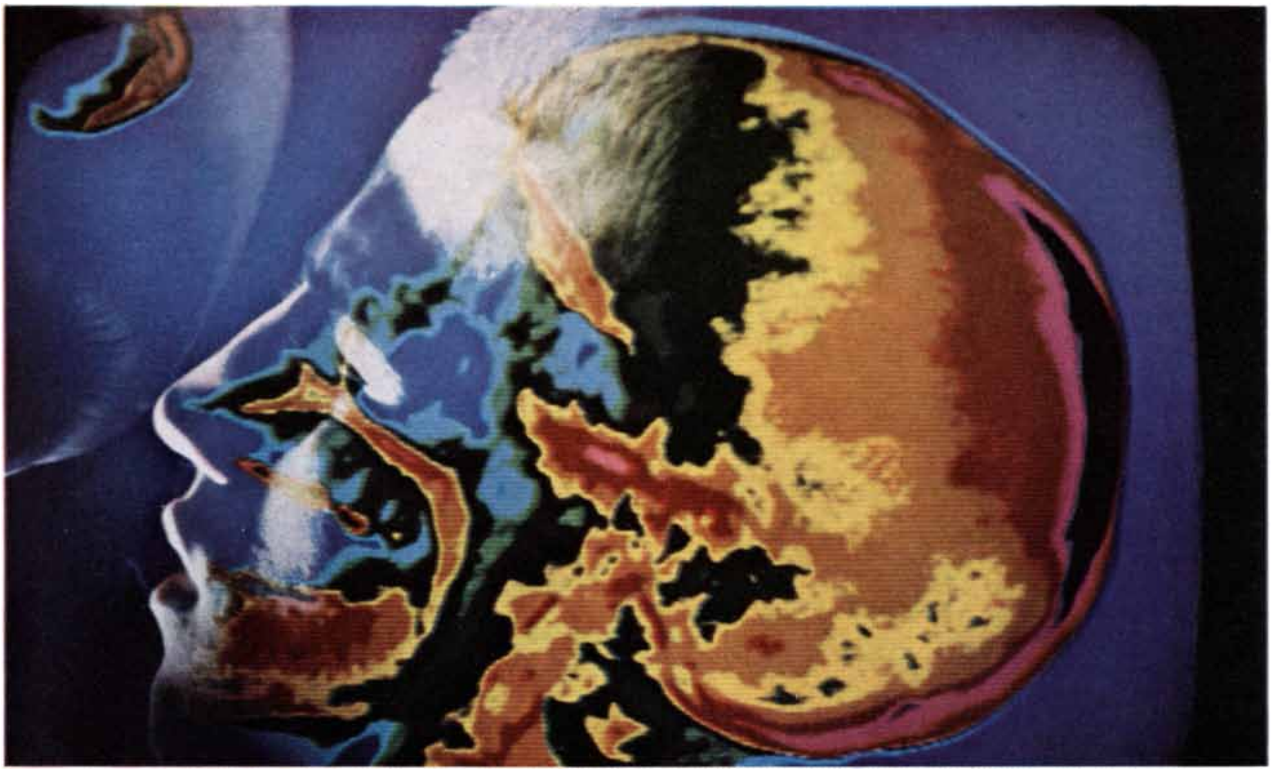
The village of Ozette on its half-sheltered beach was last occupied in the 1930's, but the wood house the searchers are now examining (several more lie buried still, an archaeological savings account) dates to the era of Columbus, before even a few trade beads from Europe had found their way in exchange to the Makah. There is both high art (mostly wood carvings and engravings) and more commonplace treasure of the past. No archaeological site in the U.S. has cost more money and effort, and few will repay so handsomely in knowledge. This is a fortunate, proud and joyful book, with a foreword by the Makah Tribal Council describing their plans for a nearby museum of their own past "built by Makahs and staffed by Makahs."

There are engraving tools made from a single beaver's tooth, artfully mounted, that fit precisely the grooves in some carvings. There are some iron tools too, not always well preserved. These remain a puzzle; perhaps they were recovered and reworked by craftsmen long ago from iron in drifting wrecks of Chinese junks. These are questions still seeking answers. The Tribal Council writes: "There is just one earth, and we are all on it together."

A HISTORY OF POLAR EXPLORATION, by David Mountfield. The Dial Press ($17.50; $14.95 if ordered before Christmas). When Sir John Franklin was lost in 1847 somewhere in the Northwest Passage, they spared no ingenuity in attempting his rescue. Here is a period engraving of a balloon, 30 feet across and designed for 12 hours of flight in the Canadian Arctic, one of a set launched to scatter leaflets from little parachutes released hour after hour by a slow-burning fuse! Maybe Franklin would find one and know help was on the way. On the other hand, when Fridtjof Nansen was icebound in 1895 in his little ship *Fram,* he and another crewman decided to try for the Pole by sledge. They set out boldly, but they had to turn back and then to winter alone on deserted Franz Josef Land, building themselves a cozy stone hut. That frozen island had been visited by man only twice before, 15 or 20 years earlier. When summer came, the sturdy Norwegians had left on a coastal lap of their intended paddle by kayak a couple of hundred sea miles to Spitzbergen when they heard dogs barking. Soon Nansen "thought I heard a shout... and as I drew nearer I thought I recognised Mr. Jackson...." "Aren't you Nansen?" "Yes, I am." This Arctic Stanley-Livingston encounter (Frederick Jackson was exploring Franz Josef, and was in no way seeking Nansen) ought to be better known; here is an engraving of the young and vigorous Nansen made from Jackson's photograph as they met. Poor Nansen was suffering a little from his ordeal; he had grown rather fat on a diet of polar-bear meat!

The text of this book surveys the entire story of polar explorations, north and south, from the Greeks to 1969, when supertanker *Manhattan* cut the icy Northwest Passage and the mounties finally announced they were abandoning huskies and going over to motors. It is a well-told, laconic and pretty complete tale by an English professional writer with a special interest in the history of exploration by sea. Its more than 200 photographs (two dozen in color) are remarkably fresh and well selected, distinguishing the book. The photographs of Robert Scott's ship, seen through an ice grotto or across a strange iceberg, are brilliant and expressive Antarctic compositions by the expedition photographer, Herbert Ponting; killer-whale noses poking up in an opening in the ice represent the excellent color work of the moderns. Hard to beat, however, is a woodcut said to be reproduced from a contemporary volume on Frobisher's strange voyage, in which the Elizabethan artists caught an Eskimo harpooner in his kayak to the life. Two complaints: the endpaper maps require supplementing for any serious reader, and the picture credits explain what agency sup-

Why jump into a specialized field until you've studied all the possibilities?



If you're a new physician, you might not have decided on a field of specialization yet. That's easy to understand. But while you're deciding, you can be getting valuable experience and expert guidance towards that decision as a physician in the U.S. Air Force. Air Force physicians are exposed to many fields of medicine. They practice with highly professional people in modern facilities. And all Air Force physicians have seemingly unlimited opportunities for furthering their education and professional growth. Air Force graduate medical education programs are available. And because Air Force physicians participate in Group Practice (a collection of specialists in different fields complementing each other by expanding total capability) they all expand their individual ability and horizons with each new patient.

As an Air Force physician (and officer) you can apply for worldwide work assignments. The salary is excellent. And you get a 30-day paid vacation each and every year.

If you're still deciding on a medical specialty...or a future...look into the Air Force before you make a final decision. We think you'll like what you see. For more detailed information about Air Force opportunities in medicine please fill out and mail the coupon.

C-SA-124

Air Force Opportunities
P.O. Box AF
Peoria, IL 61614

Please send me more information on the Air Force Physician Program. I understand there is no obligation.

Name ______________________ Sex ☐ M ☐ F
(Please Print)

Address ______________________

City ______________________

State ________ Zip ________ Phone ________

Soc. Sec. # ________ Date of Birth ________

Health Care at its best.
Air Force.

plied the picture, not where the reader can look it up in its original form.

TREASURE KEEPERS, by John FitzMaurice Mills. Doubleday & Company, Inc. ($7.95). In the year 1943 Hermann Goering paid half a million dollars for a fine painting by Jan Vermeer. The Dutch artist who sold it to the greedy Reichsmarschall knew, however, that he was betraying no national legacy to the oppressor of the Netherlands. The vendor had painted that Vermeer himself! A sunny real Vermeer in full color ornaments those pages of this fresh and fascinating book that tell the story of the dour Han Van Meegeren, forger extraordinary, who sold in the space of half a dozen years six fake Vermeers and a couple of other "old Dutch masters," all made with his special resin-hardened paints and ancient pigments. His work was so good that it earned the equivalent of three million sound dollars. Indeed, so well had he taken in the experts that many refused to deny these unknown old masters until the forger had painted a new Vermeer before their very eyes.

It is the art museum itself—the treasure keeper—that is the subject of this volume by a well-known British restorer, painter and critic. Works of art are physical systems of some complexity. They must be safeguarded against injury and the tooth of time: ultraviolet, oxidation, humidity, change, the shocks of infrequent travel, sulfur dioxide, theft, defacement and more. They require study for authentication, which has brought out the best in microanalysis, even scanning-electron-beam techniques. Often the restorers must be put to work; in one marvelous example turning a lyre from Ur of the Chaldees back from a corroded mass of silver chloride into a solid silver instrument with a polished sheen, like new. Then there are disasters like the terrible 1966 flood of the Arno against which they still work every day in the big tiled barracks behind the railroad station repairing paintings. (Restoring Florence's manuscripts will take two more decades at least; each book is taken apart and its oily pages are one by one cleaned, washed, disinfected against mold, resized, dried and rebound.)

The great hall of the National Coach Museum in Lisbon, like better-known rooms in the Louvre, the Uffizi or the Hermitage, are shown in photographs. There is a wide selection of their treasures, generally displayed to make a point, from the Pietà under a madman's attack to new display styles for Benin bronzes or Greek vases. In sum, this is a simply written book for everyone who has a fondness for museums of art, for what they hold and what they hope!

TONKA, THE CAVE BOY, by Ross Hutchins. Illustrated by Tak Murakami. Rand McNally & Company ($4.95). The personalization of a "cave boy," with a synthetic name and synthetic experiences, is a cliché of our epoch, fit for comic strips and tired television comedy. Ross Hutchins, a writer for young people, has fully redeemed the pattern by his honesty and specificity. For 20 years people have been excavating in Russell Cave, in northern Alabama, to establish that it has been settled continuously for more than 8,000 years. At one level they have found "the bones of a dog in a grave that had been carefully lined with stones, a spearhead near him." The story of Tonka is an effort to reconstruct the experiences and the emotions of a boy of those old people who might once have sadly and carefully buried his beloved dog. There was no one boy like Tonka, but his imagined tale is a look into our hearts and our human history. The line drawings of real artifacts strongly aid the story, which is illustrated with paintings of the imagined people as well.

ESKIMO SONGS AND STORIES, collected by Knud Rasmussen on the Fifth Thule Expedition, selected and translated by Edward Field. With illustrations by Kiakshuk and Pudlo. Delacorte Press/Seymour Lawrence ($5.95). Every link of the chain a poet! Without so magical a chain, this book could not have been placed in your hands. First of all were the Netsilik themselves, isolated Eskimo of the Canadian Arctic shore, "cut off from the surrounding world by ice-filled seas." Singers and poets such as Orpingilik told their work to Knud Rasmussen. Rasmussen, who was part Eskimo and had grown up in Greenland, spoke the language well. He traveled from the west coast of Greenland all the way across the Arctic to the Bering Straits in the 1920's, listening to poets. He sought out the Netsilik as people who had kept the old ways and the old songs. Rasmussen must have been part poet too. He recorded their lines in carefully literal English renderings in his official reports. Edward Field is an American poet and reader of poems. He rewrote Rasmussen's distillation from the Eskimo into an up-to-the-minute colloquial speech. Then the book was word-ready but not finished. Two men, one the old singer Kiakshuk, now dead, and the other the hunter Pudlo, linked the strong forms of their stencils and stone prints to the hard strength of this verse. They were two artists of Cape Dorset; their men and women, beasts and spirits, fly across the pages in big areas of color.

"In those times when just saying a word could make something happen, there was no light on earth yet." And the Word was made Light? There it was only the fox and the hare; the nighttime of the fox still follows the daytime of the hare. "The weather with its storms and snows was once an orphan child." The stories display the tragic lot of the old Netsilik; orphans and old people fared badly.

The Netsilik are not remote these days; Pelly Bay is on film in the schoolroom, the families at the stone weir. This version of their songs and stories grows out of the relationships formed during recent years. Cape Dorset village and its West Baffin Artists' Cooperative are also known all over the world.

FEVER! THE HUNT FOR A NEW KILLER VIRUS, by John G. Fuller. Reader's Digest Press, distributed by E. P. Dutton & Company ($8.95). John Fuller is an experienced film-maker and journalist. He found himself in the Nigerian city of Lagos in 1971 producing a film about the tour of some famous American basketball players through West Africa. A chance conversation awakened his interest in a medical drama he had hardly noticed. It would not let go. This first-class thriller is the outcome, all true. Its main strength is the vividness with which he describes the remarkably diverse characters and locations of this continuing life-and-death series.

The story begins among the American medical missionaries who run the little hospital and clinic in Lassa, a small plateau town tucked away in the corner of Nigeria not far from Lake Chad. The time was late 1969; the ambience was civil war. The missions begin and close the story, those little tightly organized groups of devoted people—able, hardworking, enormously assured by their faith. Light-plane pilots, radio operators (each mission hangs on the radio link of its homebuilt Heathkit), physicians and nurses are here. In the hospital at Lassa a new virus appears. Two experienced hospital nurses die of it, so that the next case is sent off to New York by air, a terrifying trip by light plane and jet, with a lonely stay in the Lagos Pest House, otherwise mainly given over to children ill with the normal infant diseases.

The next scene shows Columbia Presbyterian Hospital and its isolation ward. Since the two previous cases—few others were known—were fatal, the disease had to be held "frightening." The isolation-ward staff, usually urbane and cool with their careful immunizations, began to be edgy within unfamiliar gloves and masks; this disease was new. The scene shifted to the Arbovirus Research Unit at the Yale School of Medicine. (Arbovirus stands for arthropod-borne virus.) There the professional virologists would bring the infectious organism alive into the world of the known; the job could be done by their electron microscopes and complement-fixation tests in animal colonies. The nurse from Nigeria, desperately ill with a fever peaking at 107 degrees, recovered in a couple of months, but a veteran Yale research worker came down with the acute fever. He was given an immune serum from the first patient who had recovered, and he too lived. The virus was too hot to handle normally; the National Center for Disease Control in Atlanta, equipped with a brand-new specially engineered hot lab, was the only place for such a hazard.

In the next year, out in the field, trying to patch together the chains of contact that must have brought the disease to the mission hospitals, the experts—American, Canadian and Nigerian—sought ex-patients and their relatives along the highways and among the teeming marketplaces of Nigeria. They caught small mammals (analogy pointed to such hosts) and they paid close attention to every report from a bush hospital.

Antibody samples showed that Lassa virus extends over a wide area of western Africa. It next surfaced acutely in Sierra Leone early in 1973; this time they had plenty of samples, including eight wild rats trapped in the house of the first patient to die with the classic virus once the Atlanta team had arrived there in response to a call of distress. (This time the Air Force was ready to evacuate any sick Americans; commercial air travel was out of the question.) There is no doubt now; the reservoir of virus is a small rat, rarely seen in a house until the usual house rats abandon it. Control is now possible with serum and with management of the rodent. Lassa is serious but tamed. The campaign continues.

Fear, heat, human devotion, religious faith, careful scientific argument and the warmth and cordiality of Nigerians crowd this exciting book for readers of any age.

ABOUT HANDICAPS: AN OPEN FAMILY BOOK FOR PARENTS AND CHILDREN TOGETHER, by Sara Bonnett Stein. Photography by Dick Frank. Walker and Company ($4.50). Two interleaved texts form this small book. The pictures—one on each right-hand page, with several in color—illustrate the big-letter text, type a quarter-inch high. The few words tell a story of two small boys, Joe and Matthew. Matthew pushes Joe. Something is wrong. Once the man with the hook for a hand comes by, the issue is joined. Daddy and Matthew and the man talk frankly. Then they talk about Joe and his troubles. Matthew has secret fears, which come out and float away. The two boys make up. Since Matthew's feet work and Joe's hands work, they decide to build something together. And the friendship they build is really wonderful!

Meanwhile the small-print text runs along for the parents, in a serious, straightforward account of what moves people and what fears they can face and share, because "that is how people—children and adults—become easier with what is strange to them." A number of experts have some finger in this well-baked book, which is one of an Open Family series devoted to common problems (hospital, babies, dying) families should face together. The parental text helps for continuation of the discussion, preparing parents for questions and consequences of the experiences they share with their children. The Center for Preventive Psychiatry in White Plains, N.Y., is the organizational focus of the series. There are good sense, talent and the assurance of professional experience in the making of this book.

A Bestiary

OSTRICH, written and photographed by Wallace Litwin. Coward, McCann & Geoghegan, Inc. ($4.29). Litwin is a world-traveling professional photographer from New York. Between assignments in Central Africa he learned that an ostrich family was preparing to nest nearby. "I got an old Jeep and spent the daylight hours of the following month and a half standing on the motor hood with my cameras," he recalls. The male ostrich from time to time objected; he kicked the fenders to ribbons, but he could reach no higher. Only once did he jump to the hood; Litwin hid under the Jeep, very promptly.

The pictures are as real as that sounds. The book consists of some 30 striking photographs, with a few sentences of simple, well-written text for each. Young readers will enjoy it, and anyone else too. Ostrich life and love are curious. The 300-pound male courts by song and dance. He chooses three hens and scoops out an eight-foot nest in the sand, in which each hen lays eight eggs or so. These are eggs of consequence: a pound and a half to three pounds each. Fifteen or 20 should hatch. The male guards the egg laying. The dominant hen (in whose choice the male takes no part) drives the other two away. She and the proud father act as the sole parents. (Mother looks good-natured, but father is fierce.) After six weeks, incubated by the male at night and the female during the day, the chick hatches with a shrill peep, its "call of triumph." Some egg, some shell, some chick! Here is the brood of fuzzy chicks accompanying their father to the waterhole for a drink; they never bathe only in dust. By next breeding season the lucky chick joins the flock. In a couple of more years the chick becomes a parent in turn.

WHAT'S WRONG WITH BEING A SKUNK? by Miriam Schlein. Illustrated by Ray Cruz. Four Winds Press ($3.95). Why nothing. Skunks are playful, really very good-natured. They walk about relaxed, even with bears around. They have a special weapon: stink. A skunk uses it only to save his life, never to be mean or spiteful. Even then, he gives plenty of warning. If you see that tip of the tail go up, it is too late. You have had it, at 10 feet, maybe in the eye. Sometimes a hungry bobcat will go after a skunk. Then the weapon fails; it is not the skunk's way to run. The skunk is strikingly well marked; in Africa there is even an animal that looks like a skunk. It is not a skunk; it is a zorilla. In the U.S. there is a spotted skunk. It walks on its front paws before firing. Do not admire the acrobat; just leave. ("Very quietly.") This is an easy and droll book, with lots of big ink drawings having the same properties. It is a delight to encounter such a light-hearted, humorous and informed work, for young readers and for reading out loud.

THE PORPOISE WATCHER: A NATURALIST'S EXPERIENCES WITH PORPOISES AND WHALES, by Kenneth S. Norris. W. W. Norton & Company, Inc. ($7.95). Twenty-five years ago the author of this cheerful, personal, reflective recollection was a graduate student at the Scripps Institution of Oceanography, working for a "human whirlwind," Professor Carl Hubbs, a man curious about everything. Young Norris counted gray whales (courting his future wife on their shared watch) and collected warm-water fish with verve. Somehow he took the job of curator for the first big-scale marine exhibit to open on the Pacific Coast. The curator's first task was to hire a good fisherman as collector. They found Frank Brocato "top boat" along the Los Angeles waterfront. The next chapters of this book, and of Norris' life, recount the fine times they all had learning to capture porpoises and pilot whales for exhibit and training. Here is a delightful day-by-day study of Kathy, the porpoise with whose help it was first demonstrated by controlled experiments that porpoises practice a subtle echo-location. (There were, of course, plenty of hints and conjectures already.) We meet Alice, in the pools at the University of California at Los Angeles, and realize how difficult it is to sort out a discipline of clicks. Then Hawaii calls, and pools give way to open-sea experiments, with the porpoises free to perform set tasks and in the end to return on recall or to rejoin the wild world. It is still too soon to envision porpoise couriers or porpoise marine sheep dogs; the bond between men and porpoises is not yet as sure and strong as that.

Norris has followed wild porpoises from an underwater conning tower mounted in a small special craft, the Semisubmersible Seasick Machine. He is still at it, trying to understand the clicks and echoes of porpoise and of cachalot, now out of the University of California at Santa Cruz. The last pages touch deep concerns: how can we exploiting humans manage a world we never made? Norris' case study is the dilemma of the net tuna industry. Porpoise-based tuna fisheries may drown, essentially by accident, about a quarter of a million porpoises a year in the eastern Pacific alone. At the same time catching the tuna without the porpoises to mark the schools for the light plane above or for the distant lookout is only a fifth as efficient as it is with them.

THE MAGNIFICENT BIRDS OF PREY, by Philip S. Callahan. Holiday House ($6.95). Falconry is "the sport of kings." It probably arose on the steppes of central Asia, and it first reached Europe in Roman times. This book, by a research biologist who is himself a Florida falconer, combines a knowing survey of the biology of the raptors, or birds of prey, with a brief but convincing chapter on how to reach that royal state of "a hawk on the hand." The wing loading of these birds, so harmoniously formed for flight, varies from swift heavy falcons to light soaring hawks by a factor of nearly three. All of them are lighter on the wing than the lightest man-carrying glider by a similar factor. They are kites, as some are named. It is hard to credit the story told here of a cruel human raptor, a

Texan who flew his light plane decades ago in the Davis Mountains, who boasted of killing golden eagles with a sawed-off shotgun as he flew beside them, almost 1,000 birds a year for 10 years. After all, the raptors kill to eat. A hawk's eye may have three times the cone density of our own; they have two foveas in each eye, excellent binocular spacing, unusually effective muscles for lens accommodation, and eyes as big as human eyes mounted in a body weighing only a few kilograms or less.

The book is a gem of knowledge and sensitivity, as much to human nature and history as to bird qualities. A young person—not too young—who wants to watch the hawks fly by (maps of flyways are here) or even longs to take one downy eyas from the nest of a Cooper's hawk can start here very confidently. Lists of books and articles, the nature of the laws, how to meet fellow falconers and where to seek equipment are also given.

SOLO: THE STORY OF AN AFRICAN WILD DOG, by Hugo van Lawick, with an introduction by Jane van Lawick-Goodall and drawings by David Bygott. Color photographs by the author. Houghton Mifflin Company ($6.95). Packs of wild dogs, perhaps a dozen together, roam the wide plains of the Serengeti as they have for a very long time, pulling down for their food gazelles and zebras out of the grazing herds that stretch to the horizon. With Land-Rover, radio and camera the author and his colleagues have studied the life of the dog packs for more than five years. A justly famous television film and a first-rate book have already given an account of the research. This volume, honest and detailed, is the simply written account, not of the observations as a whole, but of the life of one puppy. Solo had a tough time. She was a runt, her mother was mistreated by the pack, her littermates were dead. Once the pack began its long trek to refuge from the drought, Solo finally could not keep up, although the others did not abandon her lightly. The humans who followed every move "finally came to the arbitrary decision to look after Solo for a full month." Then they placed her, rested, well fed, newly strong, with a dog family (the family of her "aunt") that had itself moved away from the pack. She was accepted "as though it was an everyday occurrence, instead of an experience which can never before have happened in the history of a wild dog pack anywhere." Solo is, they expect, a mother herself now in the Serengeti, four years old. "Maybe we shall find her.... Certainly we shall not stop looking."

FLEAS, by Helen Hoke and Valerie Pitt. Franklin Watts, Inc. ($3.90). A first book on fleas, with many line drawings by James Robins and some period engravings and photographs. This is a very appealing work on a delicate subject. The authors explicitly thank the English author Brendan Lehane, whose witty book *The Compleat Flea* started them off. They have matched his catholicity well, producing a story that mixes a good deal of flea biology (not only, although mainly, about human fleas), some general ideas of parasitology and the itchy human past and its problems. Of fleas in literature only Swift's rhyme remains. (The definitely grown-up Lehane book dealt heavily with the flea as a metaphor in prurient and pornographic literature, of which there is a considerable library!) Two fascinating topics are amply considered. One is the flea circus, that triumph of showmanship and props over common sense. Here you can see pretty well in photographs the only flea circus in Britain, that of "Professor" Len Tomlin of Manchester. New York, the theater capital of this country, has not had a working flea circus for years; only a few memorabilia of vanished glory remain in Times Square. William Heckler has passed on to a trouper's reward, his flea Paddy with him, veteran, he said, of "Broadway's longest run," a completely imaginary 52,850 performances.

The flea is not only comedy but also tragedy. There is an excellent brief account of the bubonic plague in human history, with the last pandemic about 1910. It is shown to be a mutual disease, death alike to rat, flea and humankind.

A SNAIL'S PACE, by Lilo Hess. Charles Scribner's Sons ($4.95). Snails are known for their slow pace, as Shakespeare can witness. This author-photographer has clocked her pet at some 17 miles per year, and she shows you how to time yours quite effectively. If one listens quietly enough, one can hear the rasping sound of the radula-file of the snail as it feeds. You can demonstrate this in a classroom with a cardboard-box sound studio and a recorder microphone. A description of the entire snail life cycle and a good deal more of such imaginative husbandry accompany the sharp close-ups of this always lucid animal photographer. There is more: live cowries and whelks, held in fish tanks, and some beautiful empty shells.

Snails are neither he nor she but rather both; their slow egalitarian mating is covered in a series of photographs. We can admire here a giant land snail with a shell about four inches across; the remark in the text that such beasts can reach 12 inches arouses wonder.

DANCERS ON THE BEACH: THE STORY OF THE GRUNION, by Edward R. Ricciuti. Illustrated by Richard Cuffari. Thomas Y. Crowell Company ($3.95). Ours is the planet of water. Deep in the ocean there be monsters, and on the surface the plankton make a watery pasture. Special forms of life have even developed to live, or at least to begin life, on the changing beach where the surf beats forever. One such form is the grunion, a member of the silversides family, a six-inch fish that lays eggs on all the beaches south from Point Conception, halfway down Baja California. These fish spend their adult lives near the coastline, less than 40 feet deep.

In 1919 William F. Thompson of the California State Fisheries Laboratory, walking along the surf line at Long Beach, first showed the remarkable timing of the grunion. The adult fish leave the sea briefly to lay and fertilize their eggs high on the beach only during the three or four nights each month of the season after spring tides. The ensuing nights bring a lowered water level, so that the eggs are not flooded away. Indeed, the fish spawn only at night, because on that coast the higher of the two daily tidal peaks is the one that comes at night. (The related species in the Gulf of California spawn in daytime, the time of highest tide in that landlocked bay.) After a couple of years the new grunions themselves spawn, although not in the same place where they were hatched. In the summer months after the highest tides people "dart and run after the silvery fish," which they scoop up by hand. Nets and traps are not allowed; the grunion are very good to eat.

This small, serious book, with its broad and detailed ink drawings, conveys an unusual sense of the unity of nature and of the depth of biological issues that should reach grade school readers.

Biology in General

GREEN FUN, by Maryanne Gjersvik. The Chatham Press, Inc. ($1.95). FROM SEED TO JACK-O'-LANTERN, by Hannah Lyons Johnson. Photographs by Daniel Dorn. Lothrop, Lee & Shepard Company ($5.50). GROWING UP GREEN, by Alice Skelsey & Gloria Huckaby. Workman Publishing Company ($8.95). GREEN TREASURES: ADVENTURES IN THE DISCOVERY OF EDIBLE PLANTS, by Charles Morrow Wilson. Macrae Smith Company ($5.25). We all live, city and country people, on the yield of agriculture. The headlines remind us of that every day, should we unthinkingly eat our packaged meals. These four books, which extend from a gentle, simple and childlike one of little text to one without any decoration at all, although still with easily read narrative pages, concern themselves with plants for man. We do not, however, live by bread alone. *Green Fun*, a square booklet the size of your hand, lists and pictures 30-odd things to do with the "common weeds, seeds, leaves and flowering things." It is a garden of happy people, an old man with a maple-leaf crown pinned together with the leaf stems, a charming girl deeply quiet in her daisy wreath and a small child seeking that "loud whistling squawk" made by blowing across a blade of grass. (One of your reviewers produced a fine grass whistle for the first time in more than five unfulfilled decades, and he salutes the clear directions of the small paragraph that goes with every proposal.) The poppy seed self-shaker and the basket of burrs filled with blackberries suggest the variety of this book, which is as sunny as a summer day. If these childish artifacts are not something out of the dawn of our species, one would be surprised. They feel so; the cornhusk dolly shown here, not so childish a feat, is one item in evidence. The entire book is a sheer delight and a strength for families and groups of little kids everywhere, at least in summer.

From Seed to Jack-O'-Lantern, still brief of text and lavish with black-and-white photographs, is the connected narrative of a crop, from the tractor pulling the plow across the flat Jersey acres in the stubble of early spring to the golden harvest of October, when "from miles around, people come to buy"—pumpkins. Mostly they go to make jack-o'-lanterns, for the pleasure and awe of children, although spicy pumpkin pies and custards are not to be ignored. Close-ups of flower, bee and fruit, long shots of the fields, the whole story is here, ending with carving directions and how to roast the seeds. The cast we see is mainly a set of engaged kids. The final page, an important one, explains how to grow your own pumpkin, outdoors or even in a pot, if you do not live near farmers like the Laurinos of Shrewsbury, whose husbandry the book celebrates. Indoors you will have the responsibility and the pride of using a paintbrush to take over for those pollinating bees.

A practical gardening book, *Growing Up Green*, is intended for parents and their children gardening together. It gives enough detail to get under way on a variety of scales, from seedlings in a bottle or one big amaryllis bulb up to a garden five feet by five feet or even bigger. The text treats with good sense and respect more than horticulture; it also deals with pedagogy and with attitudes. Mulches and layering are as sensible as the advice to avoid trying to involve your child in gardening with the idea that he will learn "responsibility." The waiting part is hard for children. If you think failure will teach the child a lesson, spare the effort. All he will learn is to say, years later, "I just don't have a green thumb." Gardening is caught, not taught, the authors hold. What is in this book is pretty infectious, albeit a trifle complacent.

Green Treasures is by a Vermonter, a journalist and a longtime tropical-plant hunter. It is a very lightly written series of narratives of famous plant hunters, full of fresh and important histories. Timothy grass was named by a young New Hampshire sheep farmer, Timothy Hanson, who brought it back in the 1740's from rich pasturelands near Bordeaux. Hanson came to travel the colonies, from seaboard New England out to the Kentucky River, to collect grasses and to sell his good hay seed. He never knew of our Great Plains, but until "Bermuda" grass came into prominence at the start of this century Timothy's grass (*Phleum pratense*) was the chief hay and meadow crop in our country. In his journal the old man wrote late in life: "Grass is of God's lastiness."

Plant hunters are everywhere still, nowhere more numerous and more respected than now in China. It is good to know that the greatest of agricultural nations has enough of a wild-plant population to nourish and reward organized plant hunting. Indeed, China now grows more than 600 different food crops; here we boast just over 50. But the hungry world needs more crops, perhaps yielding from both root and fruit, like the giant turnip of New Caledonia, a banana relative with five times the yield of rice in calories per acre. Starchy cassava holds the present record, 5 or 10 percent more than that turnip. This is a fascinating book, good reading for junior high school students and up, even though it presses pretty hard at recruiting plant hunters! It is a pity that no bibliography is included. There is more to read.

CHAMELEONS AND OTHER QUICK-CHANGE ARTISTS, by Hilda Simon. Illustrated in color by the author. Dodd, Mead & Company ($7.95). Aristotle did it; no reptile is better known than the wonderful chameleon, the "dwarf lion" of the Greek (it is not clear whence the name), a small tree-dweller found all along the southern coasts of the Mediterranean. He is fascinating in many respects, with his ratchet motions, independent eyes, dinosaur look and projectile tongue. He does not, however, change color to match his background. That skill is held by many forms, say the common flounder, shown here colorfully matching first a uniformly yellow sand, second a mottled substrate with red and blue pebbles, and third, heroically, black-and-white checkerboard! (Color, size, and mean spacing are matched, but the squares are rounded off.)

The chameleon can change color, all right. Here he is, painted from the life, in a tweedy red-brown, a smooth green and in between, with tasteful dark spots. Even the circus vendor's "chameleon," a small Mexican lizard properly called the anole, changes wonderfully. Here is a *Romeo and Juliet* scene, the ardent anole Romeo a fine golden green, his scarlet throat fan puffed out, while in the little plastic palm six inches above is Juliet, demurely brown. During mating she remains brown, save "for a small piece of skin on her neck which the male grasped. . . : that small area was bright green." The main inducers of color change in lizards are behavioral: mood, emotion, changes in environment, general health. These lizards have a coded color language, and protective coloration is for them quite secondary. So it is with frogs and cuttlefish, "cleaning" fish and crabs.

This remarkable book discusses, very often from direct observation, many more examples of color change. Each one is painted by the artist in her meticulous and sensitive style. The mechanisms are teased out as far as we know them, and redrawn microscopic views help us understand. When we see that a chromatophore cell can unfold from a pea shape to that of a spider chrysanthemum, we begin to understand. The anole provides an example of relatively simple color-change equipment: directly under the thin, clear outermost layer of its skin there is a static layer of droplets of yellow pigment, filling the space between the cells. Then follows a thick layer of cube-shaped cells filled with purine crystals, whose structural iridescence provides much angle-dependent blue and violet scattering, by no means simple. Below these lies a deep layer of dark melanin-bearing cells, whose intricately branched processes reach up through the layers above to the outermost skin itself. Add hormonal and direct nervous control, and you can begin to explain what the anole does, even a sick one that in distress will display a black spot behind each eye!

The author's own wide observation and experiment, careful study in the literature and plenty of paintings distinguish this work. To read it—or better, to try to extend it with anole or frog, if not squid or cuttlefish—will be a delight for any skillful reader who cares about living things.

SEE WHAT I CAUGHT!, by Ann Thomas Piecewicz. Illustrated by Perf Coxeter. Prentice-Hall, Inc. ($4.95). Grade school children who can reach grassy lawns, open fields, streams or ponds, or wooded places can catch live pets, without much danger, one imagines, to either species. This informal, modest, practical book, with amusing drawings that themselves drop quite a few hints, tells how. It is a guide also to simple husbandry; it is wrong to take a pet and then neglect it. Nets and critter carriers, cages and their cleaning, cage furnishings and food —all are here for a variety of common but fascinating cold-blooded little pets. The animals treated are turtles, salamanders, various frog forms, crickets and fireflies. There are wise warnings against the use of fresh tap water and against soap. The hunt for spring peepers is no sinecure, but it sounds like a rewarding challenge. You use a flashlight on the branches. What you are looking for is a white bubble: the peepers blow up their throats to make the sound, and the inflated throat looks like a white bubble on the tree branch! Happy hunting, and feed your frog with live food or a dangled morsel of hamburger every other day. On rainy days he may peep all day long from his leafy jar.

THE INSECT SOCIETIES, by Edward O. Wilson. Harvard University Press ($7.95). No book on biology in the past 20 years has been as satisfying as this treatise on ants, bees, wasps and termites (and maybe a few spiders). Its first publication was in 1971; this cheaper paperback edition makes it more accessible than ever. It is written with clarity and verve, but what distinguishes it particularly is its catholic mastery of all of biology, from paleontology to formal genetics, from ethology to biochemistry. Nothing less can be an adequate basis for the study of our social colleagues on this earth, and nothing less has its courageous and energetic author settled for. Biology is a whole science, and here it is wholly seen. Any young person who is a serious student of biology at the high school level and higher can grow from working at this book. It goes without saying that it is a grown-up treatise, no juvenile edition, but it is so honest and yet so rich that it attracts and holds by the scent of understanding.

Master Builders of the Animal World, by David Hancocks. Harper & Row, Publishers ($8.95). The author is by training an architect. Caught in the problem of design of structures for the London and the Bristol zoos, he found his way a few years ago to write a novel book—reviewed in these pages in 1971—that looked at the entire problem of the structures human beings have built for use by animals over the centuries. Now he has gone still further into biology and has produced a lively and very well-illustrated book (with a number of color plates) on the structures that for a much longer time the other animals have built very well for themselves.

His reach is a catholic one: webs and eggs, bird's nests and termitaries, gerbil burrows and decorated bowers all find a clear account here. Construction techniques and details of form are particularly treated; outside the technical literature one will not find the knots tied by the weaverbirds or the underwater nest forms of the various sticklebacks. Our primate cousins are not much at shelter; they do not need it in the tropical forests. The dormouse seems, after man, to be the best mammalian builder: it weaves a spherical nest of long summer grasses, lined softly and often camouflaged. Through the neat opening peers the dormouse mother, very snug indeed. It is hard to forget the bowerbirds, artists and decorators. Their males are drably feathered, but they exert themselves and display their tastes in decorating their nuptial bowers with colorful stones, bottle tops, petals, broken pieces of red plastic and blue glass bottles. Jay Gatsby tried no harder to overcome the impoverishment of low birth!

The Rand McNally Atlas of World Wildlife, edited by Martyn Bramwell, with a foreword by Sir Julian

Huxley. Rand McNally and Company ($25). This big, flat volume has a fine old-fashioned look in that it presents page after page chock-full of little colorful boxes displaying birds and beasts and plants in satisfying variety. But it is truly up-to-date, in that its pictures are a logical mix of excellent color photographs, paintings from thumbnail size to two-page spreads, colorful diagrams, and maps in profusion. It is a corporate product, with credit given to a dozen firms of artists and a score of expert biologists, mainly British but including an Aussie, a Muscovite and a New Yorker or two. The endless details of layout, research and typography have been well managed by a London company. Though we lose the sense of authorship, we gain a plum pudding of a volume for reference, browsing and dreams.

The core of the work is about a dozen chapters, similar in form but not identical. Each deals with a region of the world's living things, three northern continental regions, three tropical ones, then Australia and New Zealand, the polar realms, the oceans and their islands. Take the chapter on northern Asia. First we encounter a two-page painted panorama. Brown bear, Arctic fox and lynx hunt calmly in a pine forest at its snowy edge, a scene as crowded with birds and stouts and voles as any zoo. The next pages hold a key to the panorama's animals and a full-page relief map of the region. Then five two-page spreads each consider, with many a box and inset, one special habitat within the region: the Siberian taiga, the greatest forest on the earth; the windswept steppe, a sea of grass; the cold deserts from the Caspian Sea to the Great Wall of China; the life on the roof of the world, and finally the great river valleys of China. For each locale we see a brief text and some dozens of varied images, mostly in color, telling related facets of its natural history. Here are, among much else, Lake Baikal and its landlocked seal, the ground-nesting steppe eagle, the Bactrian two-humped camel and the brush-tailed jerboa (its skull sealed to retain humid air), the yak and the wonderfully spiral-horned markhor, the ginseng and the Japanese crane.

The last section of the book deals with humans and wildlife; on the one hand there is a catalogue of endangered species, with fine black-and-white drawings and maps and on the other a map and brief account of the world's national parks and reserves. There are good indexes; it makes a lasting indulgence of a book for schools and libraries.

INDEX OF ADVERTISERS

DECEMBER 1974

ANNUAL INDEX

The following index lists all the authors and articles that appeared in SCIENTIFIC AMERICAN *during 1974. Also indexed are "Mathematical Games" and "The Amateur Scientist."*

AUTHORS

ARTICLES

MATHEMATICAL GAMES

THE AMATEUR SCIENTIST

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

EYEWITNESS TESTIMONY

ON THE WITNESS STAND: ESSAYS ON PSYCHOLOGY AND CRIME. Hugo Münsterberg. Doubleday, Page and Company, 1915.

EYEWITNESS IDENTIFICATION IN CRIMINAL CASES. Patrick M. Wall. Charles C Thomas, Publisher, 1965.

EXPERIMENTER EFFECTS IN BEHAVIORAL RESEARCH. Robert Rosenthal. Appleton-Century-Crofts, 1966.

THROUGH A BAG, DARKLY. Robert Buckhout in *American Psychologist,* Vol. 23, No. 11, pages 832–833; November, 1968.

THE SEARCH FOR BLACK HOLES

BEYOND THE BLACK HOLE. John A. Wheeler in *Science Year: The World Book Science Annual,* pages 76–89; 1973.

GRAVITATION. Charles W. Misner, Kip S. Thorne and John Archibald Wheeler. W. H. Freeman and Company, 1973.

BLACK HOLE EXPLOSIONS? S. W. Hawking in *Nature,* Vol. 248, No. 5443, pages 30–31; March 1, 1974.

HYDRA AS A MODEL FOR THE DEVELOPMENT OF BIOLOGICAL FORM

THE CHEMICAL BASIS OF MORPHOGENESIS. A. M. Turing in *Philosophical Transactions of the Royal Society of London: Series B,* Vol. 237, pages 37–72; August 14, 1952.

POSITIONAL INFORMATION AND PATTERN FORMATION. Lewis Wolpert in *Current Topics in Developmental Biology: Vol. 6,* edited by A. A. Moscona and Alberto Monroy. Academic Press, 1971.

A THEORY OF BIOLOGICAL PATTERN FORMATION. A. Gierer and H. Meinhardt in *Kybernetik,* Vol. 12, No. 1, pages 30–39; 1972.

REGENERATION OF HYDRA FROM REAGGREGATED CELLS. A. Gierer, S. Berking, H. Bode, C. N. David, K. Flick, G. Hansmann, H. Schaller and E. Trenkner in *Nature: New Biology, Vol.* 239, No. 91, page 98; September 27, 1972.

DISTRIBUTION OF THE HEAD-ACTIVATING SUBSTANCE IN HYDRA AND ITS LOCALIZATION IN MEMBRANOUS PARTICLES IN NERVE CELLS. H. Schaller and A. Gierer in *Journal of Embryology and Experimental Morphology,* Vol. 29, No. 1, pages 39–52; February, 1973.

THE PRIMARY EVENTS OF PHOTOSYNTHESIS

PHOTOSYNTHESIS. Eugene B. Rabinowitch and Govindjee. John Wiley & Sons, Inc., 1969.

LIGHT AND LIVING MATTER: VOLS. 1 AND 2. R. K. Clayton. McGraw-Hill Book Co., 1970 and 1971.

PHOTOSYNTHESIS: STUDIES IN BIOLOGY. D. O. Hall and K. K. Rao. E. Arnold, Ltd., London, 1972.

CHLOROPHYLL FLUORESCENCE AND PHOTOSYNTHESIS. Govindjee, G. Papageorgiou and E. Rabinowitch in *Practical Fluorescence: Theory, Methods and Techniques,* edited by George G. Guilbault. Marcel Dekker, 1973.

THE SOLIDIFICATION OF CASTINGS

SOLIDIFICATION: PAPERS PRESENTED AT A SEMINAR OF THE AMERICAN SOCIETY FOR METALS, OCTOBER 11–12, 1969. American Society for Metals, 1971.

SOLIDIFICATION PROCESSING. Merton C. Flemings. McGraw-Hill Book Co., 1974.

SOLIDIFICATION TECHNOLOGY: PROCEEDINGS OF THE FIRST ARMY MATERIALS TECHNOLOGY CONFERENCE. Edited by John J. Burke, Merton C. Flemings and Alvin E. Gorum. Brook Hill Publishing Company, 1974.

THE MYSTERY OF PIGEON HOMING

ORIENTATION BY PIGEONS: IS THE SUN NECESSARY? William T. Keeton in *Science,* Vol. 165, pages 922–928; 1969.

MAGNETS INTERFERE WITH PIGEON HOMING. William T. Keeton in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 68, No. 1, pages 102–106; January, 1971.

HOMING IN PIGEONS WITH IMPAIRED VISION. K. Schmidt-Koenig and H. J. Schlichte in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 69, No. 9, pages 2446–2447; September, 1972.

RELEASE-SITE BIAS AS A POSSIBLE GUIDE TO THE "MAP" COMPONENT IN PIGEON HOMING. William T. Keeton in *Journal of Comparative Physiology,* Vol. 86, pages 1–16; 1973.

ORIENTATION OF HOMING PIGEONS ALTERED BY A CHANGE IN THE DIRECTION OF AN APPLIED MAGNETIC FIELD. Charles Walcott and Robert P. Green in *Science*, Vol. 184, No. 4133, pages 180–182; 1974.

THE ORIENTATIONAL AND NAVIGATIONAL BASIS OF HOMING IN BIRDS. William T. Keeton in *Advances in the Study of Behavior: Vol. 5*, edited by D. S. Lehrman, R. Hinde and E. Shaw. Academic Press, 1974.

THE DETECTION OF NEUTRAL WEAK CURRENTS

NEUTRINO PHYSICS. Leon M. Lederman in *High Energy Physics: Vol. 2*, edited by E. H. Burhop. Academic Press, 1967.

INTRODUCTION TO HIGH ENERGY PHYSICS. Donald H. Perkins. Addison-Wesley Publishing Co., Inc., 1972.

EXPERIMENTS WITH NEUTRINO BEAMS. Barry C. Barish in *Scientific American*, Vol. 229, No. 2, pages 30–38; August, 1973.

COUNTERFEITING IN ROMAN BRITAIN

ROMAN COINS. Richard Reece. Ernest Benn Ltd., London, 1970.

COUNTERFEIT COINS IN ROMAN BRITAIN. George C. Boon in *Coins and the Archaeologist: British Archaeological Reports*, No. 4, pages 95–171; 1974.

INFLATION UNDER THE ROMAN EMPIRE. Arnold Hugh Martin Jones in *The Roman Economy: Studies in Ancient Economic and Administrative History*, edited by Peter Astbury Brunt. Basil Blackwell, Oxford, 1974.

MATHEMATICAL GAMES

CYBERNETIC SERENDIPITY: THE COMPUTER AND THE ARTS. Edited by Jasia Reichardt. Frederick A. Praeger, 1969.

COMPOSITION SYSTEMS AND MECHANISMS. Percy A. Scholes in *The Oxford Companion to Music*, edited by John O. Ward. Oxford University Press, 1970.

FROM MOZART TO THE BAGPIPE WITH A SMALL COMPUTER. T. H. O'Beirne in *Bulletin of the Institute of Mathematics and its Applications*, Vol. 7, No. 1, pages 3–8; January, 1971.

THE AMATEUR SCIENTIST

HANG GLIDING: THE BASIC HANDBOOK OF SKYSURFING. Dan Poynter. Published by the author, Box 4232, Santa Barbara, Calif., 1974.

The energy challenge. Our people take it personally.

Come behind the scenes for a moment.

Combustion Engineering is a corporation that has served and supplied energy industries for over 60 years.

But C-E is also individuals—over 35,000 of them—who know the energy situation personally, from outside and in.

Each of them is an energy consumer, buying gas for the family car, paying utility bills and reading headlines about Mid-East oil fields. But for each of them, energy is also the job and the professional challenge.

Over 5,500 people, more than 15 percent of C-E's total personnel, are scientists and engineers. Over the years they've come up with many ways to transform energy or produce power more efficiently.

Like a better method for firing fossil-fueled boilers, which is now used by power plants and industries on 6 continents.

Or a special system for burning the low-sulfur coals that are plentiful in our Western states and could go a long way towards closing the energy gap.

Everybody at C-E knows that one of the biggest energy challenges is how to double America's electric power every ten years. And we know that the solution is to find a new balance of power. To develop ways of generating more power economically without depleting our oil and coal resources and without spoiling the environment.

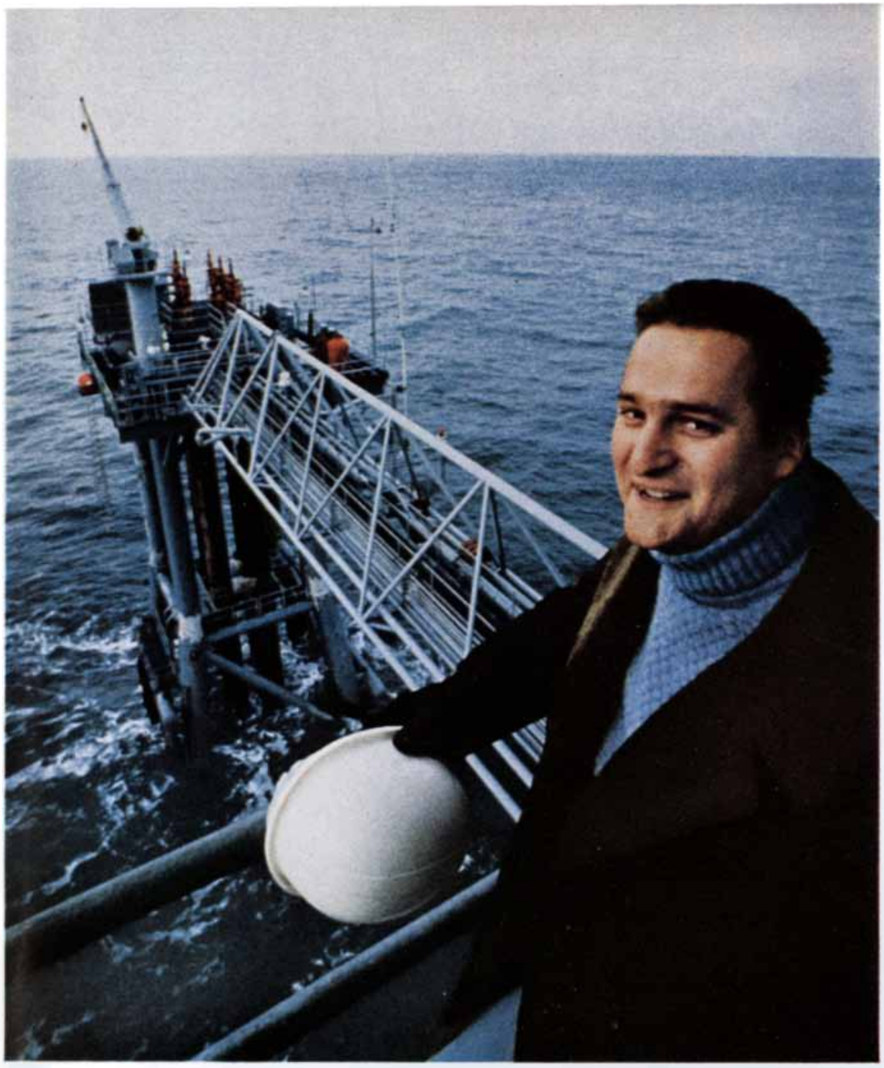
Is nuclear power one of them? Absolutely. C-E personnel have been working with nuclear power for over 30 years, so they know firsthand how abundant, clean and safe it can be. Today, they are building nuclear systems, training the men who run them, and designing advanced systems for the power needs of the future. Such as breeder reactors which will produce more fuel than they consume.

In our labs, factories and in the field, C-E people are working with the energy technologies of today and tomorrow—coal, oil, gas, nuclear power. That's how the Company was planned and built: by staying close to a variety of energy technologies.

And when you're that close, you know that energy challenges can be met. By the ultimate energy source—men and women who think long and work hard.

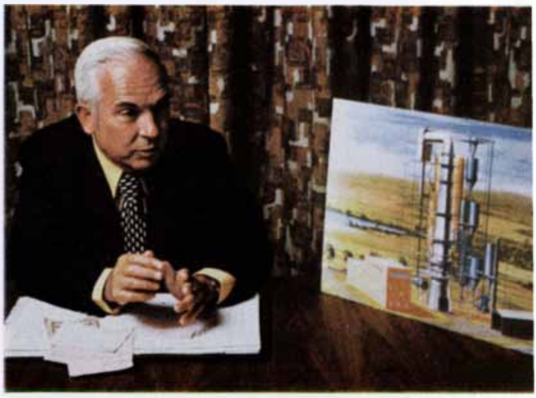
And who take their work personally.

Nuclear power is his business. Nuclear power may account for half of America's electrical energy by the end of the century. Helping to meet that goal are John West, Vice President-Nuclear Power Systems, and C-E's new System 80 design which has achieved wide acceptance by utilities since its introduction a year ago.



He coaxes oil and gas from the sea. Fritz Boschitsch is a C-E Natco sales engineer whose work relates to one of the world's major new oil exploration locations in the North Sea. In early 1974, most oil being produced in the North Sea was flowing through Natco equipment, such as shown on this offshore platform in Denmark's Dan field.

He fights pollution, and wins. Bill Murray, Industrial Products Regional Manager for C-E Air Preheater, with a fume incinerator that keeps offensive vapors from reaching the air. More than 80 of these units, part of a complete line of pollution control systems, are now in use for industrial complexes.



He's making gas out of coal. By 1985, some of the gas we use will be made from coal. Bob Patterson, Product Development Manager, Fossil Power Systems, leads efforts now underway on a program that will ultimately produce low Btu, clean burning gas for utilities . . . helping ease demand on more critical petroleum fuels.

Clean fuel from coal is his challenge. Dr. Morgan Sze, head of the C-E Lummus Engineering Development Center in Bloomfield, New Jersey, observes the performance of a prototype system for coal liquefaction. C-E Lummus' R&D efforts lead the way in the development of process technology with emphasis on coal conversion to liquid clean fuels and other coal chemicals.

If you would like more information on today's energy challenge, or are interested in a career with The Energy Systems Company, write Combustion Engineering, Inc., Dept. D, 900 Long Ridge Road, Stamford, Connecticut 06902.

COMBUSTION ENGINEERING

The Energy Systems Company