SCIENCE NEWS

EDITED BY ARCHIE AND NAN CLOW

47

PENGUIN BOOKS

Penguin Books Ltd, Harmondsworth, Middlesex
U.S.A.: Penguin Books Inc., 3300 Clipper Mill Road, Baltimore 11, Md
AUSTRALIA: Penguin Books Pty Ltd, 762 Whitehorse Road, Mitcham, Victoria

—

First Published February 1958

Made and printed in Great Britain
by The Campfield Press,
St Albans

In 1958 there will be four issues of *Science News*, appearing in February, May, August, and November. Subscriptions are accepted at 11/– for four issues, post free. Numbers 5–8 and 21–46 inclusive are still available. If there is any difficulty in obtaining any of these from a bookseller they may be had by writing direct to the publisher

CONTENTS

10 *illustrations are between pages* 72 *and* 73

THE GIBBERELLINS
NEW PLANT GROWTH SUBSTANCES

A. M. MAYER

In Japan, and other Far Eastern countries, there has long been known a fungus disease of rice that causes unduly elongated plants. This disease is caused by the fungus *Fusarium moniliforme* and allied species. 'Bakanae' disease, as it is called, has been known for at least fifty years, and as long as thirty years ago it was suspected that it is caused by some substance produced by the fungus. Shortly before the second world war Japanese scientists began to work intensively on this fungal product. They isolated a substance which they termed gibberellin, after one of the stages in the development of the fungus.* The Japanese work continued throughout the war, but went unnoticed in America and Europe. After the war, however, scientists in America and England began to rediscover the Japanese work. They isolated the substance produced by *Fusarium moniliforme* and began to study its biological properties. But it took some time before real interest in the gibberellins was aroused. Plant physiologists began taking notice only in 1955–6. A fashion for research into the gibberellins suddenly began. Why this sudden interest? What are they chemically, and why should they so suddenly and so intensely interest the plant physiologist?

First, what are the gibberellins? The work of American, Japanese, and English workers has shown the existence of at least three compounds, termed gibberellic acid, gibberellin A1, and gibberellin A2. The structure of the last two is not yet clear although they are apparently closely related to gibberellic acid. A structure has been proposed for gibberellic acid by Grove and

* Before the classification of the fungus was clear one stage was termed *Gibberella*.

his co-workers in England. This shows gibberellic acid to be a derivative of fluorene. It is a hydrogenated, substituted fluorene having two additional rings, one of which is a lactone, the point of attachment of which is not entirely clear. The proposed structure is shown in Figure 1. Gibberellic acid has a molecular

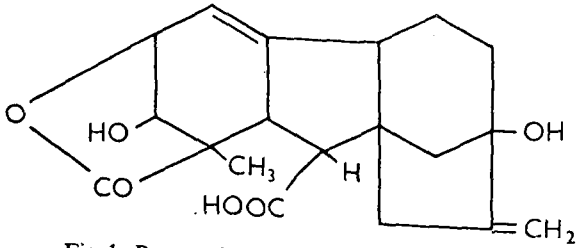


Fig. 1. Proposed structure for Gibberellic Acid.

weight of 346 and is a relatively large molecule which, sterically, is probably quite complex. Biologically the gibberellins appear to behave almost identically. For this reason we will not differentiate between them here. What is the importance of this rather peculiar fungal product to botanists?

Plant hormones, or plant growth substances as they are more generally called today, have long been known. For a long time a more or less exclusive role was assigned to heteroauxin (indolyl-3-acetic acid or IAA) in the regulation of growth and other processes in plants. IAA affects cell growth, and particularly cell elongation. It apparently regulates such diverse aspects of plant life as growth, fruit development, leaf fall, root formation, cambium division, and possibly flowering.

For many years the function of IAA as the only plant growth substance was not in dispute. As analytical methods improved, however, new growth substances were discovered. Some resembled IAA in their action. Others inhibited growth or modified the action of IAA. The control of plant growth appeared to be more complicated than at first suspected. No longer was it possible to relate it to one substance only. Moreover, no final solution to the mechanism of action of IAA was found. The

whole field was thus open for something new that might explain some of the undiscovered, unexplained processes.

It was into this opening that the gibberellins entered. Early work had shown that the gibberellins greatly resembled IAA in its action, particularly in the tremendous stimulation of growth in length of plant stems. Plants treated with a few micrograms of gibberellin became much longer, although they were somewhat thinner than untreated plants. However, while gibberellin acts on whole plants, IAA causes elongation chiefly in isolated plant parts. Further comparisons showed that in many of the biological tests that are used to prove the presence of IAA, the gibberellins are also active. This was enough to arouse immediate interest, and this interest has lately become even greater.

Formerly it was thought, as stated above, that the gibberellins are products of the fungus *Fusarium*. It has however emerged recently that higher plants produce substances very similar to, if not identical with, the fungal gibberellins. This at once increased the importance of these substances. Previously it was possible to dismiss them as curiosities produced by fungi and having some rather startling effect on plants, without considering them as natural plant metabolites. The new findings make this view untenable. Apparently we must include the gibberellins in the family of normal natural plant products. Not only this. We must also include them in the increasing number of compounds known to modify the behaviour and growth of plants.

What is known today about the effect of gibberellins? Considering the very short time that work has been under way the number of facts is enormous. As is to be expected there is a certain amount of discrepancy in the results already obtained. The first outstanding fact is their stimulation of growth. In one case at least this has recently been shown to be due to increased cell division in the sub-apical region of the plant, a species of henbane, *Hyoscyamus*. (Inset 9*a*, *b*, and *c*.)

Early hopes were directed to the possibility of increasing crop yields by using gibberellin. Although plants treated with gibberellin grow more rapidly, and get larger more quickly, the dry weight per plant is very nearly unchanged, and so, apparently, is

the yield per plot in large field experiments. Furthermore it has been shown that photosynthesis in plants treated with gibberellin is no different from that of untreated plants. Therefore, as a potential increaser of crop yields, we must discount the gibberellins at present.

There are many plants which, due to a certain hereditary make up – to certain genes – are dwarfs. Certain of these dwarf species can be converted to normal plants by treatment with gibberellins, maize for example. Others however cannot. This is of tremendous interest to the geneticist and the plant physiologist. It shows that dwarfness is not due to one single mechanism. On the contrary dwarfness can apparently be caused by many processes, some of which are reversible by gibberellic acid.

In general, gibberellic acid affects plant elongation only in the light. In the dark it is quite ineffective. It must be presumed that like other plant processes not all wavelengths are effective. Rather, a specific action spectrum is no doubt operative although it has not yet been determined. In this light dependence, gibberellin differs completely from IAA.

Another rather startling effect of gibberellic acid is on germination. Certain seeds, for example some lettuce varieties, require light for germination. Suitable concentrations of gibberellic acid, from 1 to 10 milligrams per 100 ml water, abolish this light requirement. Apparently gibberellic acid can in some way, as yet unknown, substitute for light in causing germination. But perhaps even more remarkable is its effect on temperature inhibited germination. Lettuce seeds which do not germinate at 30° C cannot be made to do so by light. Gibberellic acid, however, can to some extent, although not completely, reverse the temperature inhibition of germination. It appears, therefore, that either both light and temperature effect processes that can be modified by gibberellic acid, or that all three of them control some master reaction. It is also possible that gibberellic acid merely initiates some different process which bypasses the light and temperature dependent ones.

Great things were expected of gibberellic acid in relation to its effect on flowering. It has long been postulated that whether a plant flowers or not is determined by the presence of a substance called 'florigen'. Some plants flower if illuminated for more than eight hours per day – long-day plants. Others only flower if kept in darkness continuously for at least eight hours per day – short-day plants. Other plants, again, are indifferent to the day length in which they are kept. This phenomenon has been termed photoperiodism. The photoperiodic behaviour of plants is modified by temperature.

At various times various theories have been suggested in order to explain the experimentally observed facts. As already stated they all involve the existence of a growth substance called 'florigen', and its formation or destruction in light or darkness. Yet 'florigen' has never been extracted or isolated. So confused was the picture that some workers rejected its existence, and assumed that flowering was controlled by the balance between IAA and some antagonist to it. Others proposed the interaction of 'florigen' (of nature unknown) with IAA. Certainly in some experiments IAA did in fact change the flowering behaviour of plants in certain day-lengths. The first observations on gibberellin suggested that it could induce flowering in long-day plants not given the correct photoperiod for normal flowering. As more plants were tested this proved not to be general. Short-day plants did not respond at all to the gibberellins. Further observations uncovered other aspects of the problem. Some plants if given an unsuitable photoperiod remain in rosette form. If given treatment with a suitable temperature (vernilization) they become normal in growth but do not flower. If in addition they are given the correct photoperiod, they flower. It was found that the gibberellins could substitute in many cases for vernilization. It emerges, therefore, that although the gibberellins are undoubtedly involved in flowering they are not the long looked for 'florigen'. As to the latter it begins to become more and more likely that there is no one single florigen. Apparently the interaction of a number of growth substances is involved.

In its effect on flowering, as well as in its effect on germination, gibberellin again differs entirely from IAA. Similarly the effect

of gibberellin on tissue cultures differs from that of IAA and is somewhat confused. While the most marked IAA effects are observed by application to isolated plant parts, this, as already mentioned, is not the case for gibberellin. The latter has been reported to promote root growth in maize. In other species inhibition of stem tissue growth has been noted. Here it is worth remarking that stem elongation, at least, by gibberellin is caused only in the presence of IAA. It is possible, therefore, that the contradictory results obtained are due to different interactions of IAA and gibberellin in the tissues cultured.

If we try to summarize the present situation we are forced to a number of conclusions. The growth and flowering behaviour of plants are determined not by one single substance but by a large number of them. One of these is gibberellic acid or one of its derivatives. This substance, relatively complex in structure, is taken up by plants through leaves, stems, roots, and seeds, either in its original form or after some structural change.

Gibberellin-like substances appear to be in the nature of normal plant constituents, and not merely curiosities produced by fungi. Gibberellic acid has already been shown to effect a large number of physiological processes in plants. Undoubtedly many others will be discovered as research continues. The prospects for future work are exciting. The specific mechanism of action of gibberellic acid remains to be determined. It seems likely that what will be found is not some single reaction but rather a large number of interactions of gibberellic acid with other naturally occurring substances. IAA appears to be one of these, but almost certainly is not the only one.

Finally, it will be of great interest to determine the structural requirements for activity in the molecule of gibberellic acid. This alone opens up a fascinating study in the relationship between chemical structure and biological activity.

ELECTROLUMINESCENCE

H. K. HENISCH

FEW subjects can have a longer and more fascinating history than the art of making light. Allowing our imagination only a little free play, we may guess that light first came to the inventor as a bonus, as a highly desirable by-product of the quest for warmth and as an added reward for his skill in keeping fire under control. This association of light and warmth has remained strong to the present day, in as much that most sources of artificial light are still hot (incandescent) bodies. We have every reason to deplore this, considering how much more convenient and economical it would be to control lighting and heating independently. With this end in view, engineers devised the now familiar fluorescent lamp which derives its light indirectly from an electrical gas discharge through the intermediate action of a luminescent coating. The coating serves to convert ultraviolet radiation impinging on it into visible light. While such a lamp does not remain entirely cold when energized, no part of it even approaches the temperature range of incandescence, as the emitted light might otherwise suggest. Fluorescent lamps are, therefore, efficient and economical in operation, but they are no less fragile than conventional filament lamps and have disadvantages of their own – nor are they the only cold light sources.

Another well-known cold source is, of course, the firefly. Indeed, it is even 'colder' than the fluorescent lamp, but, as much as we may regret it, its role in illumination engineering must remain modest. Its light comes from a chemical process, namely, the oxidation of luciferin which the firefly secretes at well-chosen intervals. When substances react with one another and energy is released, it usually appears in the form of heat, but many instances are known in which the energy, or, at any

rate, a very large part of it, appears as light. This phenomenon is called *chemiluminescence* or, if biologically controlled, as it is in the firefly, *bioluminescence*. Both phenomena can be demonstrated simply in the laboratory but, varied and beautiful as some of these processes are, they have not so far proved suitable as the basis of a practical lamp.

The search for an efficient and preferably robust cold source has been greatly stimulated in recent years by the development of materials called *electroluminescent phosphors*. Their function is to emit light when an electric current is passed through them, without, at the same time, becoming incandescent. This mode of operation is thus very different from that of a filament lamp. In electroluminescence, we are evidently concerned with the direct conversion of electrical energy into visible radiation without passing through heat as an intermediate stage. Moreover, whereas the light emission of a lamp filament is a consequence of high temperature alone and cannot be otherwise controlled, cold luminescence in a solid is a structure-sensitive property. The phosphors used in practice are usually based on zinc sulphide to which various impurities, or 'activators', are added, depending on the desired characteristics. Luminescence is known to arise from electronic processes which take place within the immediate crystal environment of such activators, often referred to as luminescent centres. These electronic processes depend on a supply of excited (high energy) electrons and this can be provided in one of two ways: by irradiating the material with light of a suitably short wavelength (photoluminescence) or by allowing electrons to gain energy from an applied field (electroluminescence). In due course, electrons excited in either of these ways release their extra energy and return to their original state. When they do, the energy is normally given up in the form of heat, but in the phosphors the luminescent centres are localities in which the quantum relations are favourable for the release of energy in the form of visible radiation. The difference between photoluminescent materials, e.g., those used in fluorescent lamps, and most electroluminescent phosphors is thus primarily a difference in the mode of electron

excitation. The light emission processes are believed to be the same or, at any rate, very similar.

Apart from zinc sulphide, there are many other substances that exhibit light emission under electrical excitation, for instance, silicon, silicon carbide, cadmium sulphide, zinc oxide, gallium phosphide, boron nitride, calcium tungstate, and even a variety of organic compounds. All these materials can be made electroluminescent, though the mechanism of the energy conversion need not be the same in detail. The first observations of electroluminescence itself were probably those made by Lossew in 1923[1] and, in a different context, those of Destriau[2] in 1936. It has since been the subject of numerous investigations, some in the field of illumination technology and some in the course of attempts to analyse the fundamental processes involved, important as they obviously are for our understanding of matter in the solid state.[3]

ELECTROLUMINESCENCE OF POWDER AGGREGATES

To demonstrate electroluminescence, the activated phosphor in powder form is spread as a thin layer between two conducting electrodes, at least one of which must be transparent. A simple arrangement of this kind is shown in Figure 1. Techniques for the preparation of conducting layers on glass, e.g., by means of tin oxide, are well known, and the phosphor powder can conveniently be applied as a suspension in a nitrocellulose lacquer. A typical thickness would be 0·005 cm. It does not matter very much how finely divided the phosphor is, except that too coarse material makes it difficult to achieve a uniform layer. It is not essential for the operation of this 'cell' that the individual powder grains should be in contact with one another. In general they are separated by a thin film of lacquer, and the current passing is thus partly capacitive. On the same grounds the cell still functions if insulating layers are interposed between the electrodes and the phosphor suspension. It is reasonable to expect that the properties of the lacquer should have some influence on the performance of the cell. The ideal substance for this purpose should obviously be transparent and

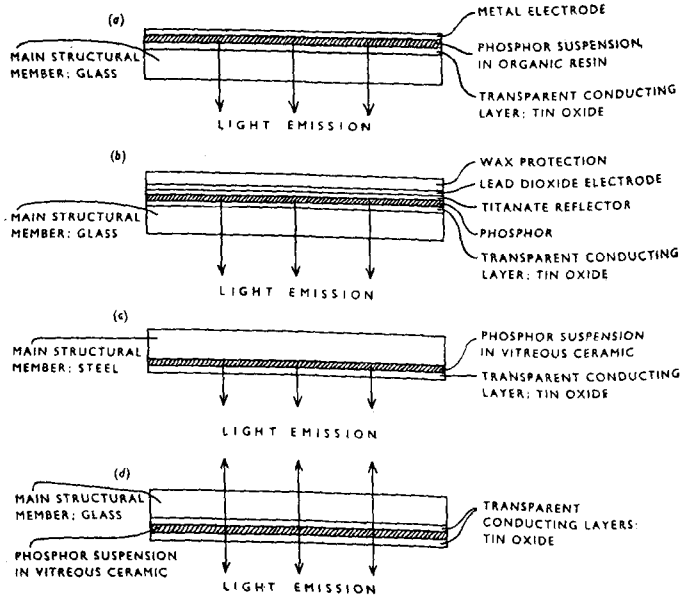


Fig. 1. Electroluminescent light sources, illustrating various types of construction. (Thickness not to scale.)

high dielectric constant, a high breakdown strength, and a low loss factor. The thin metal electrode can be subsequently applied by one of several well-known processes, e.g., vacuum evaporation. When an external voltage is connected to such a cell, the phosphor emits light which is visible through the transparent electrode and can thus be examined.

The detailed aspects of electroluminescence are varied and highly complicated, but the most important observations can be summarized briefly. The two properties of this new light source that are of principal interest are, of course, its average brightness and its colour. Typical results are given in Figure 2. It will be seen that the brightness increases only slowly below a certain threshold voltage, and much more rapidly when this voltage

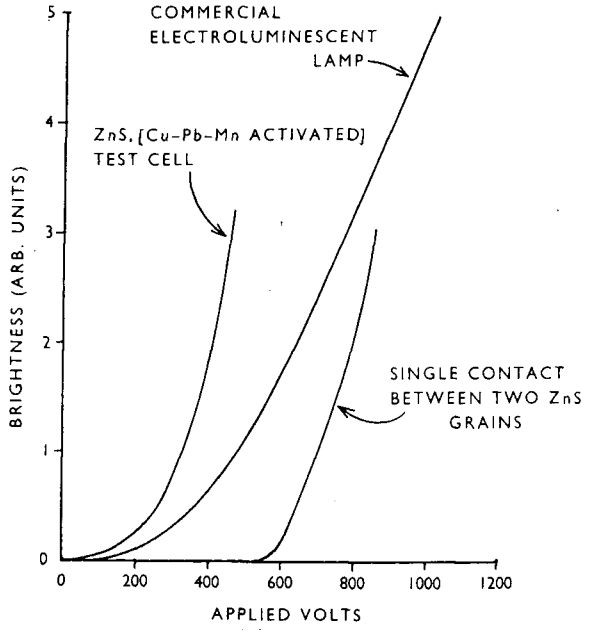


Fig. 2. Electroluminescent brightness as a function of applied voltage. Typical results obtained on powder specimens for low frequency excitation.

is exceeded. Also, whereas conventional light sources are largely independent of the frequency of electrical excitation, the brightness of an electroluminescent cell may depend on this parameter quite sensitively. Indeed, not only the brightness (for a given voltage) but also the colour of the emitted light depends on frequency. The colour range is determined largely by the activator present, and it can be controlled within wide limits. As the frequency is varied, the colour changes. A zinc sulphide phosphor, which emits mostly green light under low frequency excitation (e.g., 500 cycles per second), emits light with a much more prominent blue component at higher frequencies (e.g., 2,000 cycles per second).

The response of these cells is rapid and there is no appreciable warming-up period. When the instantaneous light emission is kept under continuous observation, it is found to be subject to periodic fluctuations which occur with twice the frequency of the applied voltage. These fluctuations are called 'brightness waves' and their complicated wave-shapes have been studied extensively.[4] They are known to depend on the structure of the cell and on the nature of the activator. The situation is further complicated by the fact that the light may spread over a wide range of the visible spectrum, different colours being emitted at different phases of the excitation cycle, and possibly also from slightly different places within each phosphor grain.

Electroluminescent Lamps and Indicators

Much remains to be done to investigate and explain the behaviour of these powder aggregates, but it is already clear that they can form the practical basis of several kinds of lamp.[5] Some of these are shown in Figures 1*b–d*. Figure 1*b* illustrates an improved version of the lamp already described. The titanate reflector increases the light efficiency of the device and the lead dioxide electrode acts as an automatic protection in case of overload. If any weak spot should arise in the phosphor layer, the dioxide which is normally a semiconductor would be subject to localized heating. Under these conditions the composition of the material changes slightly and the substance becomes non-conducting. It is thus impossible for a weak spot to burn out. This protective process is often referred to as *self-sealing*.[6] The lamp in Figure 1*c* is of rather different construction in that its main structural member is a metal plate. Such lamps are very robust and stable. Figure 1*d* shows a design that enables light to be emitted from both sides of the plate.

From the practical point of view, the long-term stability and operational life of these devices is of great importance. Electroluminescent lamps differ from conventional light sources in that their working life is not normally terminated by a sudden and complete failure. Instead, there is a steady deterioration during use until the brightness becomes unacceptably low. Under

normal continuous operation, a working life of 25–40,000 hours is a reasonable expectation. The practical efficiencies now obtainable are in the neighbourhood of 5–10 lumens per watt, as compared with 10–15 lumens per watt for conventional filament lamps. For a given light output present electroluminescent lamps thus dissipate more power but, in view of their large surface area, their operating temperature is much lower.

To remedy the shortcoming of relatively low surface brightness is perhaps the most important aim of present research and development. However, there are existing and potential applications for which high brightness is not essential and sometimes not even desirable. In such situations electroluminescent panels have some attractive features to offer. The panels can be made to luminesce in many different colours, depending on the phosphor used. Colour control on a given panel by the adjustment of frequency is hardly practicable, since the frequency is rarely variable in power systems, and also because frequency and brightness are not independent. On the other hand, it may be possible to superimpose two or more translucent panels based on different phosphors and to produce resultant colour variations by energizing the component layers to different relative intensities. Electroluminescent lamps can now be made with a high degree of reliability and they are mechanically more robust than traditional light sources. In special applications, substantial advantages are associated with the fact that these lamps are almost two-dimensional and can be produced in any size or shape. Lamps from a quarter square inch to 18 square feet have been made. As the area increases one would expect it to become increasingly difficult to ensure uniformity of the phosphor layer. For some applications, it is of particular interest that the electrodes on a given panel can be shaped more or less arbitrarily. This makes it possible to display several alternative sets of information on the same panel when the device is used as a luminous indicator rather than as a lamp.

LIGHT AMPLIFIERS AND OPTRONS

There is every reason to believe that in future the exceptional

advantages of electroluminescent lamps will be exploited further and will lead to the large-scale introduction of these devices into our everyday environment. Moreover, if artificial light originally had the status of a 'bonus', the development of electroluminescent panels illustrates that the bonus system in research is as important as ever, because, apart from their functions as lamps and indicators, electroluminescent panels have several other applications of potentially great technical significance. Most of these employ the panels in conjunction with certain other substances, known as *photoconductors*. Normally these materials are poor conductors of electricity, but when illuminated become more or less highly conducting. They can thus serve as light detectors and their numerous applications in other contexts are well known. For present purposes the photoconductor can be used either as a separate unit or as an integral part of an electroluminescent display panel. Consider, for instance, the arrangement shown on Figure 3*a*. In darkness, the photoconductor, which is connected in series with the electroluminescent panel, has a high resistance, and the voltage across the panel is therefore low – too low to stimulate light emission. When light from some other source falls upon the photoconductor (Figure 3*b*), its resistance 'collapses' and a much larger share of the total applied voltage appears across the panel – sufficient to cause electroluminescence of high intensity. By suitable design and choice of materials it is possible to make the light radiated by the panel greater than that incident upon the photoconductor by a factor of more than 100. We thus have an arrangement which is capable not only of transmitting a light signal but of amplifying it.[7]

The simple light amplifier here described could have many direct applications and it is also the basis of a 'bi-stable device' known as an *optron*.[8] In the optron, the photoconductor is arranged in such a way that part of the light (if any) emitted by the electroluminescent panel falls upon it (Figure 3*c*). The resistance of the photoconductor then remains low even after the original light stimulus has been withdrawn. Accordingly, once the device has been stimulated by an external source (which may

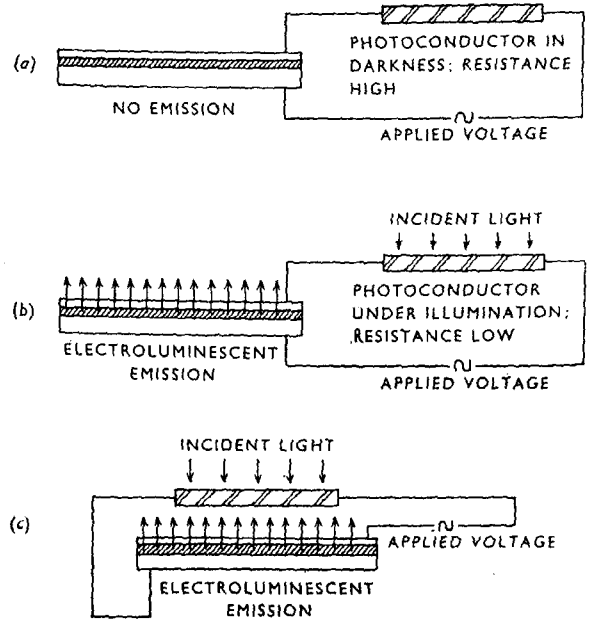


Fig. 3. Electroluminescent cells operating in conjunction with photoconductors. *a* and *b* light amplifier, *c* 'bi-stable' device, the optron.

or may not be electroluminescent), the panel will remain permanently energized, irrespective of the ambient illumination. The device thus has a stable 'on' and, obviously, a stable 'off' position. In that sense it has certain important similarities to a switch, except that it has no moving parts and that it is activated optically instead of mechanically. No doubt, practical applications will be found for devices of this kind, operating alone or in conjunction with other solid state circuit elements.

Picture Display Panels and Image Converters

The light amplifier described above is a rather inarticulate detector of light, since the final display is characterized by light

intensity alone. It would obviously be of greater interest to construct a device that would not only amplify a simple light signal, but a whole picture of wide tone range. This can, in fact, be done along the same principles.[9] What is needed is a two-dimensional array of minute light amplifiers so that each small area of the incident picture is handled by its own amplifier. For this purpose it is best to combine the photoconductor and the electroluminescent cell into one integral unit. The simplest assembly of this kind is shown on Figure 4*a*. The regions in which the electroluminescent emission is bright will be those on which primary light of high intensity is incident. Conversely, the regions on which no primary light is incident will not luminesce. A picture projected on to the top surface is thus reproduced in intensified form on the bottom surface. Since the photoconductor is immediately adjacent to the electroluminescent layer (compare Figure 3*c*), the luminescent display remains stable even after removal of the incident light. This is no more than an instance of the familiar principle of 'positive feedback'. For many purposes this feedback is a disadvantage and this relatively simple device has a variety of other shortcomings as well. Most of these arise from the thickness of the photoconductor that is needed to control the electroluminescent layer. The photoconductor is necessarily opaque, since light cannot interact with a material without being absorbed by it. The illumination of a thick photoconductive layer is thus bound to be inefficient since only the immediate surface regions can be affected by the light. More complicated forms of display have therefore had to be devised. One of these is shown on Figure 4*b*. The photoconductor is now applied as a grooved layer and is situated upon a semi-conducting film which serves to diffuse the currents and prevents the appearance of striations on the picture display. The top electrode takes the form of a thin layer of silver which covers only the apex of each ridge. A further opaque and thin insulating layer (lampblack in a thermo-setting resin) is interposed to protect the photoconductor from the electroluminescent emission and thus to prevent the uncontrolled stabilization of the image referred to above. As one would expect, the device is difficult to produce,

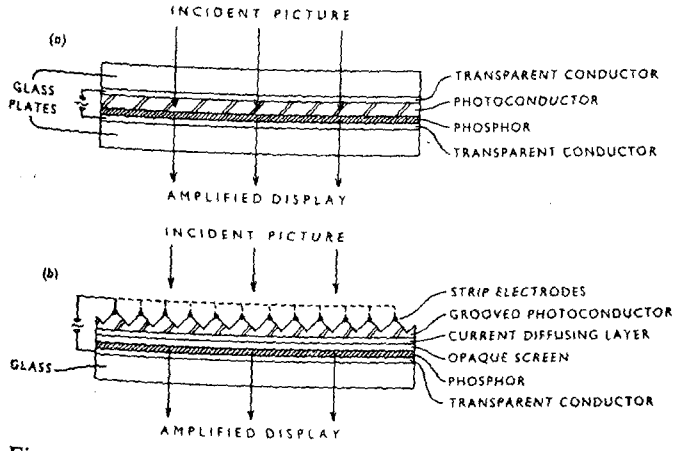


Fig. 4. Electroluminescent display panels: *a* with positive feedback, *b* with positive feedback prevention. (Thickness not to scale.)

but satisfactory plates of 12 × 12 inches area have in fact been made. Photoconductors are known which respond to radiation in the X-ray and infra-red regions of the spectrum. By using such materials in display panels, these devices become image converters with many potential applications, e.g., in medical radiography.

Display panels of the kind illustrated on Figure 4 can serve to reproduce an amplified picture and even a moving one, provided, of course, that a less intense picture is already available. In this respect their operation differs profoundly from that of a television-type cathode ray tube, in which the picture is produced by periodic scanning and intensity modulation. The question arises whether an electroluminescent display panel can be adapted for the same purposes, i.e., to compose a picture from a series of electrical signals. There are hopes that this will eventually be done. For this purpose it is necessary to envisage a display panel which consists of two sets of parallel strip electrodes aligned at right angles, one on each side of the phosphor (Figure 5). By means of a suitable switching system, voltages would be

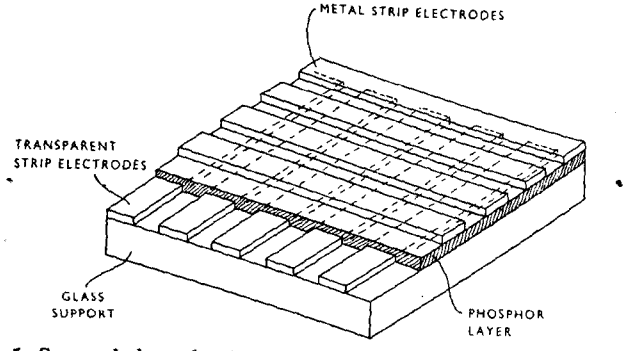


Fig. 5. Scanned electroluminescent display device: schematic representation of simplest case. (Thicknesses not to scale.)

applied to the strips sequentially, in such a way that the cross-over, at which electroluminescence will appear, scans the whole picture area. In addition, the voltages applied would have to be modulated in accordance with a video signal so as to reproduce the various tones of the picture. In principle all this would be relatively simple, but very serious complications arise in practice from the stringent requirements of frequency response, intensity, and tone contrast. It may be found possible to overcome these difficulties in the course of further research and development. If this were so, we could look forward to having a television screen in the form of a flat picture surface in a frame, without bulky apparatus behind it, and connected to its electronic master only by a thin set of wires in a cord.

EXPERIMENTS ON SINGLE CRYSTALS

To improve the efficiency of the above devices, a good deal can probably be done by design and empirical development alone but, in the last resort, it will be necessary to gain a much deeper insight into the fundamental mechanism of the electroluminescent process. In this context, display panels based on powder aggregates are unsuitable from the experimental point of view, since the results are too complicated to permit unambiguous

interpretation. The general rule is, therefore (as, indeed, it is in other branches of solid state physics), to avoid powders and to experiment on single crystals. It is in this field that the scientific interest of the subject is mostly concentrated. It is possible to grow single crystals in the laboratory, either directly from the vapour of the material concerned, or by allowing vapours of the constituent elements to react in a controlled environment. Efforts are always being made to increase the size of the crystals grown in this way, and to achieve control over their degree of structural perfection and impurity content. The linear dimensions of the crystals now at our disposal range from several millimetres for zinc sulphide to several centimetres for silicon. Even at the lower end of this scale the crystals are large enough to permit the application of electrodes and to give scope for optical and electrical observations by delicate techniques. Unfortunately the crystals are hardly ever perfect; most of them have structural defects, known as 'stacking faults', which may affect their electroluminescent behaviour.

The experimental results obtained on single crystals are also complex, but not as hopelessly entangled as those obtained from powders. In contrast to powders, single crystals luminesce even when the current passing is uni-directional and constant. A point of immediate interest is the location of the light within the crystal, and although the problem appears simple, different conclusions have been reached by different experimenters. On superficial inspection, the light generally seems to be concentrated in the neighbourhood of the electrodes, but it is now appreciated that these appearances may be highly misleading, at any rate in small crystals of materials like zinc sulphide which have high refractive indices. A high refractive index gives rise to multiple internal reflections, and when these are suppressed by observing the luminescent crystal while immersed in a medium of similar refractive index, the light is seen to be uniformly (but not always continuously) distributed throughout the volume.[10] Thus, electroluminescence *can* be distributed in this way, but this may not always be so, and there are instances in which the emission is genuinely localized.

Electroluminescence under direct constant current excitation presents an important aspect of the problem, but, as an experient, it is not by itself highly informative. In nearly all experients, therefore, the applied electrical stimulus is made to vary ith time. From the theoretical point of view, the next simplest rocedure would be to apply a pulsed uni-directional field so iat the inherent delays of the luminescent response and subequent decay can be studied. However, for various practical easons this does not yet appear to have been done on any extenive scale, and most of the results so far available have been btained under sinusoidal excitation, either in its simplest form r as a rectified half-wave. This last method was chosen in a eries of experiments by Frankl[10] in order to take advantage of he 'rest period' provided during the suppressed half-cycle Figure 6). During this half-cycle no field is applied, and the pecimen can thus return to something like its normal quiescent tate before the new excitation begins. When such a rest period s not provided, e.g., when a continuous sinusoidal excitation is pplied, it is found that the electroluminescence does not at any ime decay to zero. Conditions are then evidently more complicated and thus less susceptible to interpretation.

Results of the kind obtained by Frankl are illustrated in Figure 6. It will be seen that during every half-period of excitation there are two instances of light emission. It is found that the two peaks depend on the applied voltage in different ways. There is disagreement among observers as to whether the two peaks are of the same colour or not. A further important observation (though not shown in Figure 6) is that there is no trace of any reverse (depolarizing) current during the quiescent half-period.

Theory of Electroluminescence in Single ZnS Crystals

As already mentioned, the aspects that are peculiar to electroluminescence are those that concern the process by which electrons are accelerated to high energies. It is also necessary to formulate a mechanism whereby a luminescent centre in which light emission has just taken place is reconditioned for the next

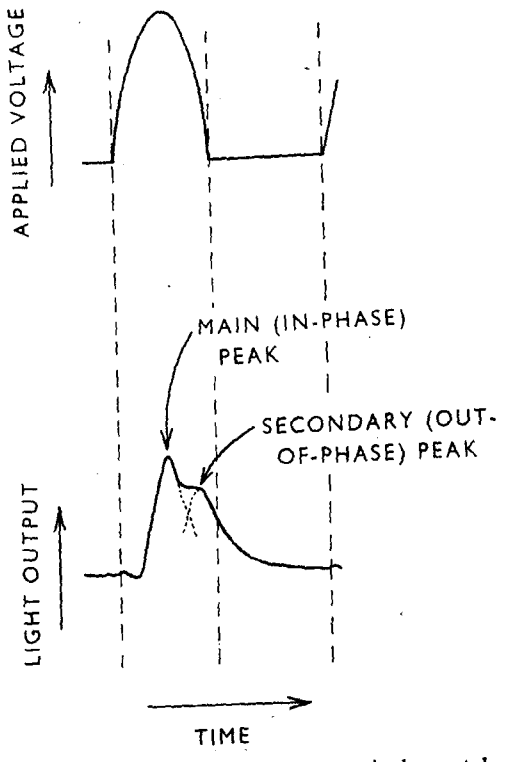


Fig. 6. Typical results obtained on a single crystal of zinc sulphide, ZnS. (*After D. R. Frankl.*[10])

emission. Moreover, to account for the second light peak in each half-cycle, one must envisage a delaying process between excitation and the return of some of the electrons to the normal state. We refer to this phenomenon as 'trapping'. The theoretical principles that govern all these processes by themselves are reasonably well understood. The remaining problem is thus, not so much how electroluminescence can arise, but precisely how the mechanisms that are known to be theoretically possible can

function (and perhaps interact) in particular crystal environments and lead to the observed experimental results. On this subject there is still a good deal of controversy.

Before a luminescent centre can emit radiation it must be excited and this can be done in several ways. In principle, it could be done by an electric field alone, provided it is sufficiently powerful to allow an electron to escape from the centre (field ionization). Alternatively, excitation could be achieved by the impact of some high energy electron with the centre and this may or may not result in the physical removal of an electron from the centre itself (impact ionization). This demands not only a high field, but a field that remains high over an appreciable distance so that the electrons concerned can acquire the minimum energy required to ionize a centre. For various reasons, including some connected with considerations of efficiency, impact ionization is more widely held to be the process responsible, at any rate in zinc sulphide. In either case it is necessary to explain how a sufficiently high field strength can arise. The fields required are much stronger than the *average* fields calculated from the applied voltage and the specimen dimensions. It is thus necessary to assume that the internal field is concentrated in certain localities. It would follow that the electroluminescent emission should be likewise confined to these localities and their immediate neighbourhood. Experimenters who observe their electroluminescence at contacts have thus favoured theories that depend on the high electric field known to be normally associated with a contact 'barrier'.[11] On the other hand, when electroluminescence is observed uniformly distributed throughout the bulk material, it is necessary to assume that the crystal contains a series of closely spaced internal barriers between highly localized regions of different impurity content. A good deal is known about the properties of such barriers from experience with other semiconductors, particularly germanium and silicon. It is possible that such barriers are in some way associated with the stacking faults already mentioned, though this has not yet been reliably confirmed.

Given a high electric field, luminescent centres would be excited by impact with high energy conduction electrons and some of the centres would return to their normal state immediately, emitting radiation which we observe as the main (in-phase) peak. Others, apparently, return only after a delay. In principle, such a delay could come about if the excited electrons were separated in space from their centres and needed time to return. This has indeed been postulated, but difficulties arise from the fact that no depolarization currents are observed corresponding to such a return movement. Alternatively, the excited electrons may be temporarily trapped in suitable neighbouring localities and the probability of release may be governed by the field itself.[10] These problems and many detailed aspects of the subject are still unresolved. It has been said that some years ago electroluminescence was a complex subject that we did not know how to take apart, whereas it is now a series of relatively simple subjects that we do not know how to put together. Progress can thus take peculiar forms, and there may be surprises in store for us yet. In any case, a transitional stage of this kind offers great scope for imaginative experiment and judicious interpretation. There is every reason to believe that this field of research, which combines practical, theoretical, and, not least, aesthetic merits in such ample measure, is entering a phase of rapid and highly promising advance.

* * *

This paper has appeared in Hebrew in *Mada* (*Science*), The Weizmann Science Press of Israel, Jerusalem.

REFERENCES

1 Lossew, O. W.: *Telegrafia i Telefonia,* 1923 No. 18 61, see also Lossew, O. W.: *Phil. Mag.,* 1928 **6** 1024.
2 Destriau, G.: *J. Chim. Phys.,* 1936 **33** 620.
3 Destriau, G., and Ivey, H. F.: *Proc. I.R.E.,* 1955 **43** 1911. A comprehensive bibliography is given by these authors.
4 Destriau, G.: *Brit. J. Appl. Phys.,* 1955 Supplement No. 4 49.
5 Payne, E. C., Mager, E. L., and Jerome, C. W.: *Sylvania Technologist,* 1951 (January) 2.
6 Henderson, S. T., and Mash, D. M.: Personal communication.

7 Orthuber, R. K., and Ullery, L. R.: *J. Opt. Soc. Amer.*, 1955 **44** 297.
Cusano, D. A.: *Phys. Rev.*, 1955 **98** 546.
Williams, F. E.: *Phys. Rev.*, 1955 **98** 547.
8. Loebner, E. E.: *Proc. I.R.E.*, 1955 **43** 1897.
9 Kazan, B., and Nicoll, F. H.: *Proc. I.R.E.*, 1955 **43** 1888.
10 Frankl, D. R.: *Phys. Rev.*, 1955 **100** 1105.
11 Piper, W. W., and Williams, F. E.: *Brit. J. Appl. Phys.*, 1955 Supplement No. 4, 39.

PSYCHOPHARMACOLOGY

D. S. TROUTON

PSYCHOPHARMACOLOGY is not a new science, but it is a convenient term under which to describe recent advances in controlling, by means of drugs, psychological functions such as mood and behaviour. Of course, agents such as alcohol, opium, and a large number of naturally occurring substances, especially alkaloids, have been used from time immemorial on account of their effects on the mind. Their use has, however, been largely empirical, and there has been little understanding of their mode of action. Recently the advances have been such as to have encouraged the hope that a new era in psychiatry has begun which may eventually justify the belief, rather surprisingly held by Freud, that drug therapy will eventually oust all other forms of treatment of psychiatric disorders.

The two groups of drugs on which most attention has been focused recently have been:
(*a*) those that tend to make normal people appear to be psychotic (hallucinogens, psychotomimetic drugs or phantastica), and (*b*) those that are mostly intended to have the opposite effect, i.e., to make disturbed psychotic (or sometimes neurotic) patients more normal (tranquillizers or ataractic drugs).
The practical importance of the second group is obvious (and even drugs in the first group have been claimed to have some therapeutic value), but it is chiefly with their theoretical significance that we are concerned here.

HALLUCINOGENS

Mescaline

Mescaline is one of the alkaloids that can be obtained from a cactus, *Lophophorus* (*Anhalonium*), which grows in the north of Mexico and on either side of the Rio Grande. The sliced off

tops of the cactus are called mescal buttons or peyote, and have for a long time been consumed by certain of the American Indians of that region as part of their religious rituals. These rituals have now become blended with some of the features of Christianity, and members of this cult believe that the Holy Spirit is present in the mescal,* which is consumed sacramentally. The 'National American Church of the United States', dating from c. 1870, attaches much importance to peyotism.

Since 1886 when the German pharmacologist, Lewin, isolated the alkaloid mescaline, there have been innumerable reports of its effects. Among the first to describe their experiences were Havelock Ellis, in England, and Weir Mitchell, in the United States. Some hours after taking the drug, Weir Mitchell found himself 'deliciously at languid ease'. He later went into a dark room where 'The display which for an enchanted two hours followed was such as I find it hopeless to describe in language which shall convey to others the beauty and splendour of what I saw.' He generously provided William James with a supply of mescal buttons, but after taking one William James was violently sick for twenty-four hours. Havelock Ellis wrote that the chief characteristic of the visions he observed under mescaline was their 'indescribableness', but this has not prevented a large number of people from attempting to describe them. During the next half century reports on its effects continued to appear and it was sometimes claimed that the drug was a valuable tool for the psychiatrist. However, the minute self-descriptions and other data that were accumulated, mostly by Continental psychiatrists, only yielded a meagre contribution to science.

This failure may be partly attributable to the fact that many of the earlier studies were largely descriptive, being concerned with an attempt to collect facts†; the need for testable hypotheses tended to be overlooked. There was thus an excessive, though implicit, reliance on the Baconian method; the chief disadvantage of this was succinctly expressed by the logician Susan Stebbing,* when she wrote of Francis Bacon: 'He completely failed to see that experimental observation must be guided and controlled by hypothesis', and it appears that Bacon was not the last to exhibit this failure.

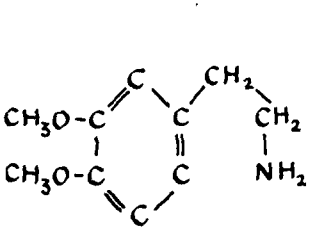
* 'Mescal' is rather a confusing name as it is also the name of a distilled liquor which is drunk in Mexico. It is not derived from the cactus but from *Agave americana* or Century plant (which may take as long as sixty years before flowering – hence its name).

† Even the 'facts' are open to certain criticisms. If mescaline impairs judgement, this may affect the ability to report experiences. In any case, the objective recording of behaviour has usually proved to be more conducive to the advance of psychology than the attempt to derive laws from subjective reports. Moreover, many of those who were investigated under mescaline were physicians who already knew what to expect. There was no attempt to obtain a random sample.

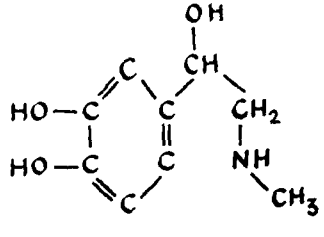
Another factor that may have contributed to the delay in formulating the hypothesis discussed below was the lack of liaison between different sciences. Eventually a hypothesis was put forward, based on the following:

(1) Certain of the symptoms of mescaline intoxication have long been known to resemble those found in schizophrenia, although the differences have, perhaps, received more emphasis.

(2) The chemical structure of mescaline is similar to that of adrenaline (a hormone which is widespread in the body and is liberated from the medulla of the suprarenal gland when the defences of the body are mobilized in 'fright or flight').



Mescaline



Adrenaline

* Stebbing, Susan: *A Modern Introduction to Logic* (London, Methuen), Third edition, 1942, p. 491.

It was a long time before (1) and (2) were related but, given the additional information that a toxic substance produced within the body has often been postulated as a cause of schizophrenia, the reader may well be able to formulate a hypothesis about the possible cause of schizophrenia; namely, that owing to some disorder of metabolism of adrenaline or its related compounds, a mescaline-like substance is produced within the body which plays a role in producing the symptoms of schizophrenia. It is surprising that no one had attempted to work along these lines until about five years ago, but, as Osmond pointed out concerning mescaline, 'Psychiatrists appear to have been unaware of its biochemical structure, biochemists to have been uninterested in its production of an artificial psychosis.'

A pioneer in the production of experimental psychoses in animals, de Jong, reported in 1930 that by administering large doses of mescaline to various different kinds of animal he had been able to induce symptoms having a marked similarity to those of schizophrenia. His report appears to have attracted little scientific attention. Ten years later, Stockings argued that the causative agent of the psychoses (which he considered to be all variants of the same disease process) was probably a substance similar to mescaline. He was awarded a medal for his work but its implications were not pursued. It was not until 1952 that two psychiatrists, Osmond and Smythies, working in London, put forward their hypothesis that a mescaline-like substance was responsible for the genesis of schizophrenia. This time the suggestion was not ignored. Interest in hallucinogens had recently been revived by the discovery of lysergic acid diethylamide (which will be described later). Moreover, Osmond and Smythies made a strong case for the existence of a mescaline-like ('*M*') substance in schizophrenia.

They argued that mescaline intoxication should not be compared with chronic schizophrenia (as those who emphasized the differences between the two conditions had done), but with *acute* schizophrenia. Table I (which is a much simplified version of theirs) shows some of the similarities between the two conditions. Both may be associated with very diverse symptoms and,

TABLE I. *A Comparison between the Psychological Effects of Mescaline and the Symptoms of Acute Schizophrenia**

	Mescaline		*Acute Schizophrenia*	
Sensory Disorders	*Illusions*	*Hallucinations*	*Illusions*	*Hallucinations*
Vision	+ + +	+ +	+ +	+ +
Hearing	+	+	+ +	+ +
Synaesthesia	+ + +	−	?	−
Motor Disorders				
Catatonia	+ +		+ + +	
Behaviour Disorders				
Negativism	+ +		+ + +	
Withdrawal	+ +		+ + +	
Thought Disorders				
Pressure	+ +		+ +	
Disturbed association	+ + +		+ + +	
Blocking	+ +		+ + +	
Conceptual thought replaced by visual images	+ + +		+ + +	
Neologisms	+		+ +	
Disorders of Interpretation				
Ideas of influence	+ +		+ + +	
Paranoid ideas	+ + +		+ + +	
Heightened significance of objects	+ + +		+ +	
Delusions	+ +		+ + +	
Depersonalization	+ +		+ +	
Mood Disorders				
Fear and terror	+ + +		+ + +	
Depression	+		+	
Indifference and apathy	+ +		+ + +	
Manic symptoms	+		+	
Insight	Sometimes absent		Sometimes present	

+ = occurs: + + = marked when it occurs: + + + = marked and frequent: − = not relevant. * From Osmond, H., and Smythies, J.: *J. ment. Sci.*, 1952 **98** 309.

of course, no one could show *all* the symptoms at once. There are obviously some differences; for example, under the influence of mescaline, hearing a noise may be accompanied by corresponding changes in one of the other senses, most commonly vision. This phenomenon, which is termed synaesthesia, does not appear to occur in acute schizophrenia. But the seemingly nihilistic point of view of those who stress the differences is less likely to be favourable to the growth of science than is basing a hypothesis on the similarities. Whether or not the hypothesis is eventually confirmed or rejected is comparatively unimportant. What *is* important is that a vast field of research has been opened up, with numerous other hypotheses to test.

Various workers have joined in the search for the '*M*' substance and the effects of a variety of compounds related to adrenaline have been investigated. Adrenochrome, adrenoxine, and adrenolutin are among the compounds for which psychotomimetic properties have been claimed. Numerous other lines of research have been followed. For example, Lea argued that there should be abnormalities of pigmentation in schizophrenics if there is an abnormality in the metabolism of adrenaline because the substance melanin, the pigment in hair and skin, is, like adrenaline, derived from tyrosine. Over a thousand cases of schizophrenia and two controlled series were studied and it was 'found that there was a highly significant excess of dark hair in the schizophrenics', but this was limited to the age-group 15–19 years. Other lines of research that are being followed include the search for substances that antagonize the hallucinogenic drugs, and the investigation of their effects in cases of naturally occurring psychoses.

It is indisputable that biochemical abnormalities commonly occur in schizophrenic patients. In fact the trouble is that such varied abnormalities have been observed in these patients that it is difficult to regard any one of them as fundamental. There is some evidence that schizophrenia has a constitutional basis, so that there may be an inherited tendency towards certain physical (including biochemical, and neurophysiological) and psychological patterns of reactions. As these are intimately

interdependent it is difficult, if not futile, to attempt a general answer to the question whether biochemical or psychological influences initiate the process.* Both could form part of the vicious spiral leading towards the disorder. Hallucinations can be produced experimentally in man by psychological methods, which involve depriving the volunteer of sensory stimulation for a prolonged period. How this is related to the pharmacological method of producing hallucinations is still unknown, but the knowledge that hallucinations can be produced by psychological means does help to prevent an undue bias being placed on physical factors. Nevertheless it must be admitted that it is rare to encounter the conditions that appear to be needed to induce hallucinations psychologically. Moreover, naturally occurring hallucinations are more susceptible to the effects of certain drugs than to psychotherapy.

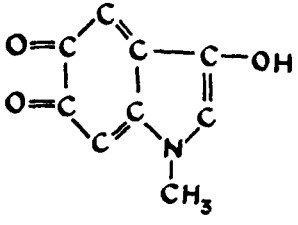
Lysergic Acid Diethylamide (LSD)

The discovery that lysergic acid diethylamide (LSD), a semisynthetic derivative of ergot, was a potent hallucinogen did much to revive interest in this type of drug. Its psychological effects were accidentally discovered in 1943, by Hoffmann, a chemist working on the derivatives of ergot. He must have absorbed some of the drug, possibly by inhalation, without being aware of it, and feeling unable to work he went home to bed, and when he drew the curtains he fell into a peculiar state of drunkenness characterized by an exaggerated imagination. When he closed his eyes, fantastic pictures of extraordinary plasticity and intense kaleidoscopic colourfulness seemed to surge towards him. Subsequently he took 0·25 milligrams of

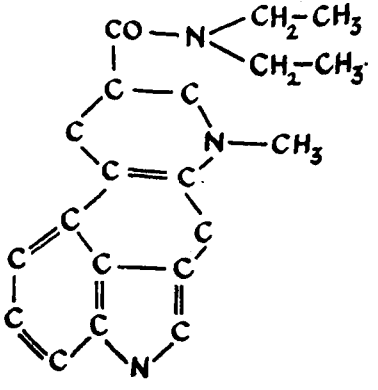
* In the case of phenylketonuria, a rare form of mental deficiency due to an inherited inability to break down phenylalanine, the cause is biochemical. However, it is now possible to prevent mental deficiency from developing, provided that a special diet is rigidly adhered to, otherwise symptoms of intoxication occur; if the forbidden foods are eaten the patient may experience effects similar to those caused by alcohol in normal people. Thus, psychological factors cannot be ignored even in this disorder.

LSD to investigate its effects; he experienced visual disturbances, restlessness, a suffocating sensation, and an occasional feeling that he was outside his body.

The effects of LSD have since been studied by numerous workers and found to be rather similar to those caused by mescaline. LSD is effective in minutely small doses of about 50 micrograms, in contrast to mescaline which must be given in a very much larger dose. The two compounds are chemically different: LSD contains an indole nucleus as do various other alkaloids such as yohimbine and harmine which may also cause hallucinations. Both adrenolutin and adrenochrome have an



Adrenochrome



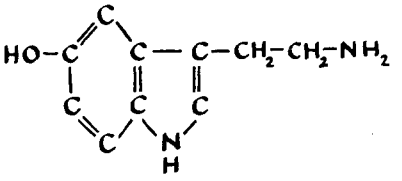
Lysergic acid diethylamide (LSD)

indole nucleus. However it is possible that the side chain of mescaline may be converted into an indole nucleus.* It has been suggested by Fischer that mescaline is transformed into an LSD-like substance before exerting its action.

A great deal of work has been done on lysergic acid diethylamide to discover its mode of action, and theories abound. Apropos of these, Sourkes writes: 'One is struck by the wide area of research they cover and by the large amount of data

* The formula of mescaline was arranged to bring out the similarity with an indole nucleus.

which is being amassed and collated in reference to them.' One theory, according to Rinkel, is that LSD acts by interfering with adrenaline metabolism. Another important theory is that it acts by antagonizing a substance known as serotonin (5 hydroxytryptamine or 5HT). This also has an indole nucleus.



Serotonin

Its role in the brain is by no means clear, but there is evidence that the other substances containing indole nuclei such as LSD are antagonistic to it, at least, peripherally. Thus it has been suggested that a psychosis may arise from a deficiency of serotonin in the brain. This provides a hypothetical link* with one of the tranquillizers, reserpine (to be discussed later), because it has been shown that it releases serotonin from the brain; this may help to explain both the effect of reserpine in psychiatric disorders and also the fact that it counteracts the effects of LSD.

TRANQUILLIZERS

The two main drugs in this group are chlorpromazine and reserpine, but the drug houses have recently introduced over twenty-five drugs that they describe as tranquillizers. As yet there has been little opportunity even to discover whether the drugs have the alleged properties. A tranquillizer differs from a sedative or hypnotic, such as a barbiturate, in that it is supposed to calm without causing confusion and it neither causes sleep directly nor the characteristic changes in electrical activity of the brain

* The relationships of these substances are probably more complex than this; for example, chlorpromazine has a reserpine-like effect but apparently does not liberate 5HT; a compound related to LSD, usually referred to as BOL, inhibits 5HT as markedly as does LSD but has no mental effects, etc.

associated with barbiturate hypnosis and with sleep. Great claims have been made for them in the United States but they have not generally proved so effective in Britain. Their chief use is for reducing the overactivity of severely disturbed psychotic patients.

It is uncertain how much the numerous tranquillizers have in common. For example meprobamate, 'Miltown', which is now said to be the fourth most widely prescribed drug in America, differs from the others in that it is used chiefly in neurotics rather than psychotics; there is evidence that it reduces symptoms due to tension. Some of the tranquillizers probably only differ from sedatives and hypnotics in degree and not in kind. Both chlorpromazine and reserpine have numerous side-effects. In the past multiple remedies were given but this is now discredited: instead, many actions tend to be combined in one drug. As yet it is impossible to predict which of the drugs will continue to be used and what their ultimate position will be.

OTHER ASPECTS OF PSYCHOPHARMACOLOGY

Although the term 'psychopharmacology' is used chiefly in connexion with the two groups of drugs that have been described, there is no sound reason why other psychological research on the effects of drugs should not be included. Psychologists have not yet contributed much to the study of hallucinogens and tranquillizers, but they have done a great deal of work on other drugs; this is largely ignored by textbooks on both pharmacology and psychology, presumably because it concerns what was formerly regarded as a no-man's land, but now, labelled as psychopharmacology, it has become a meeting place for a variety of scientists. So far, no book covering this field has appeared.

The first systematic studies on the psychological effects of drugs were carried out by the great psychiatrist, Kraepelin, whose centenary in 1956 was overshadowed by that of Freud. Many drugs, and especially alcohol, were intensively studied by means of the somewhat rudimentary techniques provided by experimental psychology, which was then in its infancy. And

statistics, which loom so large now, had not even been born. Nevertheless many of his conclusions still appear to be valid, which is perhaps not so surprising when it is recalled that the vast majority of scientific discoveries have been made without the aid of statistics.

The psychological effects of alcohol were also investigated by many other psychologists, but a great number of them performed the experiments on themselves, so caution is needed in interpreting and generalizing their findings. On the whole, alcohol has been found to impair almost all types of performance, although it may temporarily improve muscular work on an ergograph, and it has also been shown to improve the performance of 'neurotic' cats.

More recently the amphetamines ('Benzedrine' and 'Dexedrine') have received the attention formerly devoted to alcohol. These drugs can delay the onset of deterioration of performance in prolonged monotonous tasks, but their effects are too complex and variable to be summarized here. Their dangers have only recently been realized, with the result that in Britain they are now on Schedule IV and can only be obtained by means of a doctor's prescription. Perhaps it is paradoxical that alcoholic beverages, which have proved so much more dangerous, have not been placed on Schedule IV.

The enormous number of new drugs constantly being introduced has necessitated the use of more sophisticated techniques to assess them, both clinically and psychologically; and the importance of the design of the experiments and their statistical evaluation is increasingly recognized. One effect of the emphasis on the use of adequate controls has been to draw attention to the importance of placebos.

Placebos

The term placebo (Latin for 'I shall please') was originally applied to the sometimes impressive but nonetheless pharmacologically inactive 'medicines' that doctors have been known to prescribe from time to time. It has now been extended to apply to the 'dummy' treatments, usually by inert tablets or saline

injections, that are used in controlled studies of drugs. The inert preparation should be as similar as possible to the drug under investigation. For example, the tastes should be similar; it is futile to use a tasteless placebo tablet if the drug itself is very bitter, but until this was very recently pointed out to them, two well-known pharmaceutical firms were providing tasteless inert tablets to compare with bitter drugs. This is just one of the many errors that contribute to the delay which usually elapses between the introduction of a drug and the final assessment of its value.

Another complication is provided by the remarkable frequency of reactions to placebos. In a review of fifteen drug studies, Beecher found that the incidence of relief by placebos ranged from 15 to 58 per cent, with an average of 35 per cent. Asthma, hayfever, cough, symptoms of tension, insomnia, and headaches were among the conditions that have been reported to have been relieved. Even objectively observable changes may occur. In some studies when neither doctor nor patient knew whether the placebo or the drug was being given, the reactions were so severe that the code had to be broken; sometimes the reactions were due to the placebo. Nearly forty 'toxic' effects of placebos have been described.

Placebo reactions may either add to or subtract from the drug effect. Whichever occurs will tend to make the assessment of drugs more difficult. In order to eliminate placebo reactors from such studies and also to enable this capacity to be made use of in treatment, attempts are now being made to find out what kind of personality these placebo reactors have, and how their reactions can be explained. Probably both suggestibility and learning are involved, and there is some evidence that neurotic introverts may be the most susceptible to placebo reactions. Incidentally, it has been found that an undue proportion of neurotic introverts volunteer for drug experiments.

THE MUTUAL CONTRIBUTIONS OF PSYCHOLOGY AND PHARMACOLOGY

The explanation of placebo reactions (which should be within the domain of the psychologist) is not the only contribution that

psychology can make to pharmacology. It has been increasingly realized that drugs may act differently according to the personality and mental state of the patient. For example, it is now possible to predict accurately on the basis of height, weight, age, sex, etc., the dose of a barbiturate (given intravenously) that is necessary to produce a particular effect. Recently good evidence has been produced by Shagass indicating that, among neurotic patients, less barbiturate is needed to produce a given effect (in the electro-encephalographic record) in hysterics (neurotic extraverts) than in anxiety states (neurotic introverts). These findings are, incidentally, in accordance with McDougall's theory of introversion-extraversion which has recently been extended by Eysenck. If this relationship between personality and required dosage is confirmed, 'type of personality' will eventually find its way into the formulae for predicting dosages that are to be found in works on pharmacology.

Pharmacology could also contribute to psychology. For example, a theory has been put forward that suggestibility and hypnosis are due to conditioning. As conditioned reflexes can be suppressed by certain drugs (e.g., certain tranquillizers and also arecoline) this provides a method by which the theory could be tested. If suggestibility (which is closely related to the ability to be hypnotized) were unaffected or decreased under the influence of these drugs, this would make the theory appear less likely. (It appears rather unlikely anyhow, which is the main reason why such an experiment has not yet been carried out.)

PROSPECTS

In a short account it is only possible to mention some of the many aspects of the rapidly expanding science of psychopharmacology, whose literature increases at an ever-accelerating rate. Unsystematic and poorly controlled clinical studies, culminating in results that are quite inconclusive, are not yet obsolete, but they may be expected to grow fewer as the multidisciplinary approach to the subject becomes more widespread.

The long delay before the implications of the chemical similarity between mescaline and adrenaline were explored would

probably not have occurred if there had been closer collaboration between biochemists and psychiatrists. And there may well be other, equally significant (and, once they have been pointed out, equally obvious) relationships that may be brought to light by the combined work of pharmacologists, bio-chemists, psychologists, electrophysiologists, and psychiatrists, all of whom have their contribution to make to psychopharmacology.

Although it is unnecessary to stress the importance of co-ordination in scientific work, it may be in place to draw attention to the lamentable lack of coordination that has been customary in what should really be a single subject, namely the science dealing with behaviour and its disorders. In the past, far from carrying out combined operations, there has usually been a wide gulf between psychologists and psychiatrists who are, in their turn, separated from the psychoanalysts – themselves divided into different camps. Psychopharmacology may well prove to be of the greatest value in helping to bridge these traditional gaps, as well as in leading to closer liaison with other sciences. It so often happens that advances occur on the borderlines between sciences that there is every reason to believe that progress will be rapid and of far-reaching importance. Even now, great claims are being made for psychopharmacology. Whether or not these are justified at present, there can be little doubt that, as it develops, it will help to resolve some of the baffling problems of mental disorders and contribute much to their treatment.

FURTHER READING

Most works on psychopharmacology are only to be found in specialist periodicals: a few are listed below. These contain numerous further references. Apart from some recent symposia on psychopharmacology which have been published in book form, the relevant books are mostly out of print, not in English, or literary rather than scientific (or all three).

Abramson, H. A. (ed.): *Neuropharmacology*. (Transactions of the second conference sponsored by the Josiah Macey Jr Foundation.) N.Y., 1956.

Eysenck, H. J.: *The Dynamics of Anxiety and Hysteria* (London, Routledge and Kegan Paul, 1957).

Himwich, H. E.: 'Prospects of Psychopharmacology', *Journal of nervous and mental Disease*, **122** 413–23.

Klüver, H.: *Mescal: the 'Divine' Plant* (London, Kegan Paul, 1928).

Lasagna, L.: 'Placebos', *Scientific American* **193** 68–71. (A simple account of placebos. A more detailed treatment of the same subject may be found in the *Journal of Mental Science*, 1957 **103** 344-54, by the present author.)

Leake, C. D.: 'Drugs affecting mood and behaviour', *Texas Reports on Biology and Medicine*, 1955 **13** 793–818.

McKellar, P.: *Imagination and Thinking : A Psychological Analysis* (London, Cohen and West, 1957).

Marrazzi, A. S., Hart, E. R., and Cohn, V. H.: 'Pharmacology of the Nervous System' (Chapter 35, pp. 565–94, in Spiegel, E. A. (ed.): *Progress in Neurology and Psychiatry* (N.Y. and London, Grune and Stratton, 1956).

Osmond, H., and Smythies, J.: 'Schizophrenia: a new approach', *Journal of Mental Science*, 1952 **98** 309–15. (Subsequent work may be found in later issues of this journal, e.g., 1954.)

Sourkes, T. L.: 'Some current biochemical theories of psychosis', *Canadian Psychiatric Association Journal*, 1956 **1** 73–81.

THE INTERIOR OF THE EARTH

W. H. MARSHALL

IT is probably true to say that we know as much about the deep interior of the Sun as we do about that of the Earth, although the Sun is 93 million miles away and the Earth is just under our feet. The study of each has to proceed by observation of effects produced at the surface by events inside. True, in the case of the Earth, we can bore down below the surface and examine the material brought up in the greatest possible detail. But the deepest terrestrial borings penetrate only two or three miles, and the rest is like deducing the structure of a golf ball from samples of the skin only a few thousandths of an inch in thickness. There is, nevertheless, a very considerable body of knowledge of the inside of the Earth, and some of the work done recently is of considerable importance.

All aspects of Earth science are of special interest during the International Geophysical Year, and, although studies of the interior form a relatively small part of its programme, it is hoped to provide here a background for some of the new knowledge that will come from this great piece of international co-operation.

The Earth is an oblate spheroid of equatorial diameter 12,756·8 kilometres, and polar diameter 12,713·8 kilometres; it has a volume of about $1{\cdot}1 \times 10^{27}$ cubic centimetres. Since its mass is 6×10^{27} gm, its mean density is approximately 5·5 gm per c.c. This density is substantially higher than that of surface rocks, for which the average value is 2·8 gm per c.c., the maximum being 3·3. Clearly the material inside the Earth is of considerably higher density than that at the surface.

This conclusion is confirmed by the Earth's moment of inertia, the value of which can be derived from observations of the precession of the equinoxes, and of the variation of gravity over

the Earth's surface. As is well known, the moment of inertia of any body depends on the distribution of the body's mass. A body with its mass concentrated towards the centre has a smaller moment of inertia than one with its mass uniformly distributed throughout its volume; while one (like a flywheel) with its mass concentrated towards its periphery has a larger moment of inertia. Now the value of I for the Earth is found to be 0·334 MR^2, where M and R are its mass and radius, while that of a uniform sphere is 0·4 mr^2. So the mass of the Earth is concentrated towards its centre, and this can only mean that the density of the material in the interior is greater than at the surface. The small inaccuracy involved in treating the Earth as a sphere has no effect on this conclusion.

This result is, of course, no more than might be expected from the compression that must be caused by the weight of the overlying material throughout most of the body of the Earth. It is interesting to see that information about the Earth's interior can be obtained in such a relatively simple way, but to obtain more significant knowledge we must turn to seismology, the science of earthquakes.

The outer parts of the Earth are subject to stresses which cause movements at the surface, sometimes slow and comparatively gentle, sometimes rapid and extremely violent. The local effects of an earthquake have been graphically described by F. Kingdon-Ward in *Science News 22*. In this discussion it is the more distant and less destructive effects that are of importance. Whenever the Earth's crust moves, waves are sent out from the focus of the earthquake; they travel through the body of the Earth, and can be recorded by means of seismographs at stations all over the world. Sample records (seismograms) are shown in Inset 4, which not only shows the nature of the record, but also indicates that the task of interpreting a seismogram is both difficult and complicated.

Seismic waves are of two types: longitudinal and transverse. The longitudinal waves are of the same type as sound waves, i.e., they consist of a set of compressions and rarefactions, which displace the particles of the medium along the line in

which the wave is travelling. Since the first waves to arrive at any station are of this type, they are known as primary waves, or P-waves. The transverse waves are like those set up in a rope when its end is moved up and down. The displacement in this case is perpendicular to the direction in which the wave is moving. Since they arrive after the P-waves, they are known as secondary or S-waves. The record may be further complicated, in the case of near earthquakes, by the presence of surface waves, but these do not concern us here.

As the seismogram of Inset 4 suggests, the paths of seismic waves through the Earth are not simple. Not only do they penetrate to different depths according to their initial directions, they are also reflected and refracted at any surface where the properties of the material change substantially. Nevertheless, it has been possible, by comparing records from a large number of stations, and from many earthquakes, to deduce the velocities of the waves within the Earth. At depths of less than 35 km, these are neither simple nor regular: but this relatively thin skin is only the crust and below the crust the phenomena are much more simple. As Sir Edward Bullard says: 'It seems that below a depth of about 35 km the complications of geology are left behind and the Earth is a body in which the properties are very closely a function of the radius only.'

Figure 1 shows graphically the values given by Sir Harold Jeffreys for the velocities of both the P-waves and the S-waves at all depths from the base of the crustal layers down to the Earth's centre. The curve shows four more or less marked discontinuities, where either the velocities or the rates at which they change increase more rapidly than usual. These discontinuities must be caused by changes in the properties of the material at corresponding depths in the Earth, and for this reason it is customary to speak of discontinuities in the structure of the Earth itself. The region at the base of the crust, for example, is known as the Mohorovičić discontinuity, and the one at a depth of some 410 km as the 20° discontinuity. The name is, perhaps, somewhat unfortunate, because it suggests a sharp boundary in the Earth, just as in the velocity curve. In fact, however, as K. E.

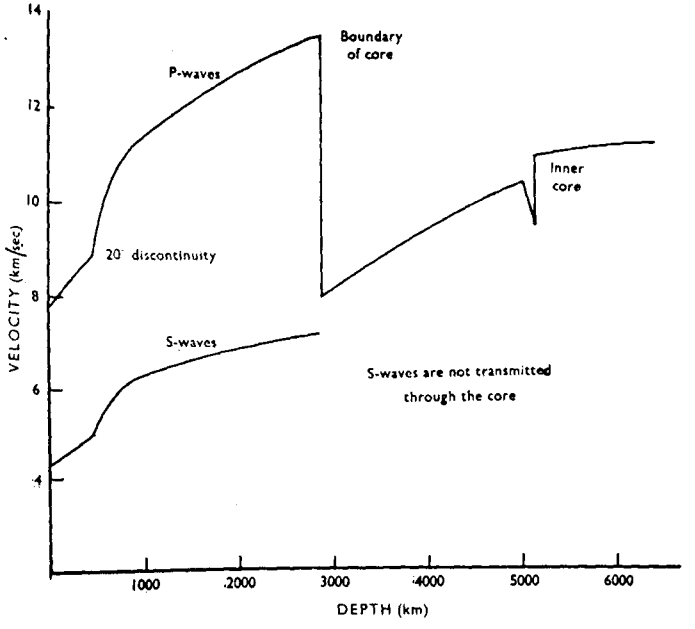


Fig. 1. Velocities of P- and S-waves at various depths in the Earth's interior.

Bullen says: '... both seismological and petrological sources of evidence indicate that the layers are probably separated not by surfaces of discontinuity in the precise mathematical sense, but by zones within which changes of properties take place relatively rapidly.'

The five shells indicated by the seismic waves are shown diagrammatically in Figure 2, together with the names given to them, and some of the information we have about them.

The most striking deduction from the curve is that the Earth has a central core which transmits P-waves at about two-thirds of their velocity in the layers above it, and S-waves not at all.

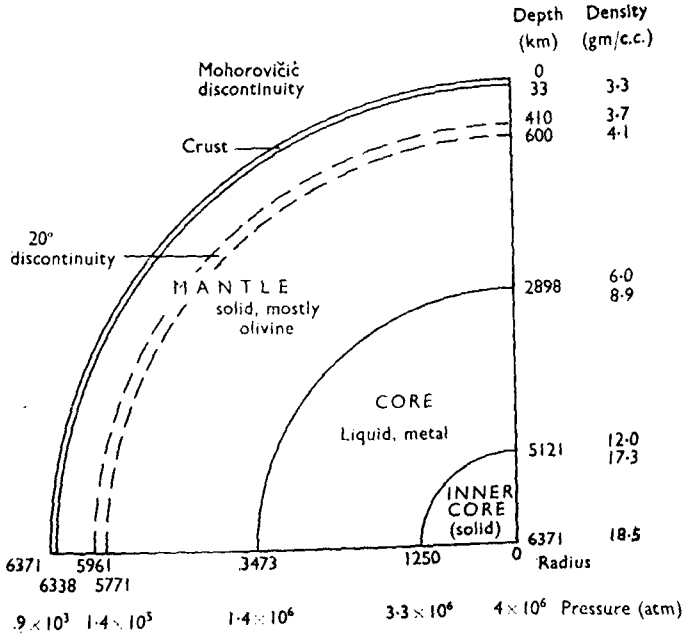


Fig. 2. Regions in the Earth's interior to correspond with the velocities of the P- and S-waves at various depths.

These properties immediately suggest that it is liquid, since only materials of very low rigidity cannot transmit shear waves. It is not inconceivable that it might be a solid with very little rigidity but with the power of absorbing and scattering shear waves; but the liquid nature of the core is suggested in other ways.

The first confirmation comes from the existence of tides in the body of the Earth. The tides raised by the Moon and the Sun are most familiar, because most noticeable, in the seas and oceans that cover more than half of the Earth's surface. The detailed study of these tides shows, however, that they cannot be fully described in terms of a fluid layer on a completely solid rigid Earth. The description must be modified to take account of a

tidal yielding of the Earth under its liquid covering, and this requires that the Earth's core be liquid rather than solid.

The fluidity of the core is also suggested by the Earth's magnetism. The surface features of the Earth's magnetic field might be produced either by a dipole at the centre, or by uniform magnetization of the whole Earth. The details of the field, however, show no relation to geographical features, so it is difficult to believe that surface layers play any considerable part in its development. The field, therefore, appears to be primarily a feature of the Earth's interior, and we have to find a likely origin for it there.

Now, one of the outstanding features of the field is that it is constantly changing, and that the time-scale for its changes must be measured in centuries rather than in millions of years. These changes are therefore unlikely to originate in the mantle, which is certainly solid, and much too rigid to show changes on any but the geological time-scale. So it seems that, though the field may not originate in the core, movements in the fluid core required by the other arguments provide the only possible explanation of the rapid changes in the Earth's magnetism.

If the earth is not uniformly magnetized, a fluid core also offers an acceptable explanation of the existence of, as well as of the variations in, a magnetic field. A magnetic field normally indicates the presence either of magnetized material or of an electric current. It is thus conceivable that the Earth's field might be due to the core being magnetized; but there is every reason to believe, as we shall see later, that the temperature of the core is of the order of some thousands of degrees. This is far above the Curie point of iron, and it seems unlikely that iron could retain any magnetism at such temperatures. There are, on the other hand, several ways in which electric currents might be produced in a core of liquid metal. Each has its advocates and there is no final agreement on any one; but there is no doubt that the Earth's magnetic field can be produced in the core, and that its variations strongly suggest a fluid core.

So far, then, we have a picture of the Earth's interior divided broadly into a solid mantle, and a liquid core, with subsidiary

regions within each, where the velocities of the P- and S-waves change. We have now to trace the changes in properties which cause the changes in velocity.

It is known that in material in hydrostatic equilibrium (i.e., where the internal pressures are balanced as in a liquid at rest), the velocities of elastic waves depend on the density of the material. So, assuming the necessary equilibrium, we can deduce this equation:

$$\frac{d\rho}{dr} = -\frac{\gamma m \rho}{\alpha^2 - {}^4/_3\beta^2}$$

where ρ is the density at radius r, m the mass within that radius, γ the gravitational constant, α and β the velocities of the P- and S-waves respectively. Since the right-hand side of this equation depends on m and ρ as well as on α and β, it cannot give a unique value for the rate of change of density with depth. We know, however, that the density of the material at the base of the crust cannot be very far from 3·3 gm per c.c., and using this value, together with estimates of the jump in density at each discontinuity, we can derive a value of the density at all depths.

There is clearly an element of arbitrariness in this procedure, and the values given by different investigators differ somewhat in detail. The variations are kept in check, however, by the need for each model to give the correct values for the Earth's total mass and moment of inertia. The values shown in Figure 3 are those given by W. H. Ramsey. The biggest deviation from other models is in the value given for the inner core, which is, however, so small that even quite substantial changes here make very little difference to the overall properties of the model. It may be worth while to remark here, that in this model the inner core is represented as solid.

We come now to the problem that raised a recent and still unresolved controversy, namely, the identity of the material in the various zones. It is generally agreed that the mantle consists of olivine, which is essentially a mixture of the silicates of iron (10 per cent) and magnesium (90 per cent). At pressures attain-

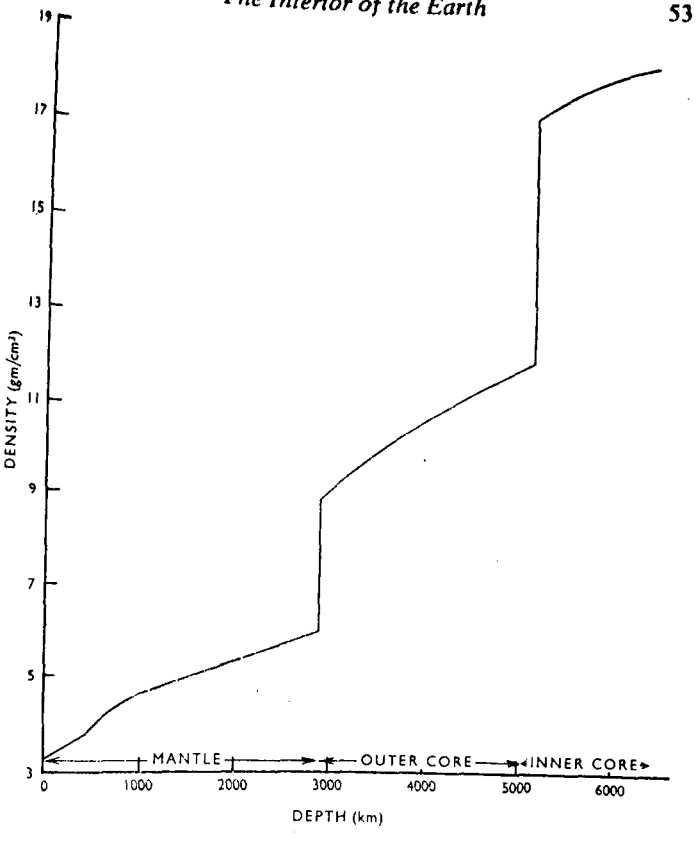


Fig. 3. Estimated density of the material of the Earth at various depths.

able in the laboratory the density of this material is the same as that of the mantle, and it also transmits seismic waves at the required velocities. At the same time the materials that come from below the crust in volcanic outflows are almost entirely silicates. The general increase in density with depth in the mantle may be attributed entirely to the effect of pressure, or to pressure

plus an increase in the proportion of iron silicate in the lower layers.

Although the constitution of the mantle is generally agreed, there is still speculation about its structure. As indicated above, it is thought by most to be a fairly uniform solid, but T. Gold has recently suggested that it may have a spongy structure with, under the influence of gravity, rivulets of molten iron percolating through it towards the core. This kind of structure would presumably lead in time to a slow decrease in the Earth's moment of inertia, as the dense iron flows inwards; but the time required for an empirical check on this hypothesis would be very very long, and it seems unlikely that the problem will be resolved in this way. A more likely method would be to look for effects on seismic waves. These presumably exist, but they are probably very small, and it may be a long time before seismologists can say with any certainty whether the mantle is a uniform solid or a 'sponge'. Meanwhile, it is probably best to regard it as a fairly uniform solid.

We have already seen that the core consists of a material whose density is considerably greater than that of the mantle in contact with it. This increase in density, from 6 to 9 gm per c.c., is so great that it is natural to assume that it is due to the appearance of a new material; and this, in fact, was for a long time the only explanation offered. It is also natural to consider only the materials that are common in the crust; to do otherwise would be to invite a situation in which 'interpretations throng and clash, and neatly equal the commentators in number'.

Of common materials the densest is iron, and for many years it was almost an article of faith that the core of the Earth consisted of liquid iron alloyed with a certain amount of nickel. The inner core was also nickel-iron, but solidified under the enormous pressure at that depth.

This theory fits quite well with the available evidence. It gives a picture of an Earth consisting predominantly of iron, magnesium, silicon, and oxygen. These also are the elements that predominate in meteorites, which until recently were the only samples of cosmic matter that could be examined on the Earth.

They are also, with the exception of hydrogen and helium, the commonest elements in the Sun and the stars, so it was reasonable to suppose that the general trend of abundance was much the same all through the universe. And this idea has received additional support in recent years from the study of cosmic rays. It is now possible to perform, as it were, a chemical analysis of these also, and here again, the general trend in the abundance of the elements is the same as in the stars and meteorites. Meteorites, too, give a strengthened argument, because while most of them are 'stony' or consist of a mixture of rock and iron there are some that are wholly iron. This is just what we might expect if meteorites were the debris of a shattered planet with an iron core and a rocky mantle. But although this theory of the origin of meteorites is widely held, it is by no means proved, and the doubt about it is too great to allow it to clinch the argument about the Earth's core.

Further evidence in favour of the iron core theory may be derived from the Earth's magnetism. We have already seen that the existence of a magnetic field and the nature of its variations point to an origin in a liquid metallic core. Iron is obviously the commonest material suitable for such a core; but since the strongest rival to iron as the material of the core is also metallic this argument now carries little weight.

The advocates of an iron core, however, have not by any means had it all their own way. In 1936 W. Kuhn and A. Rittmann criticized the theory on physico-chemical grounds. If we try to picture an Earth soon after its formation, cooling to a hot liquid mass, we may imagine an iron core forming in the same way as iron separates from slag in a blast furnace. (It may be remarked, incidentally, that this gives a picture very like Gold's spongy mantle.) Kuhn and Rittmann, however, claimed that such a process would take a time greater than the age of the Earth to form a core of the mass of the Earth's core, and they therefore concluded that the core could not consist of iron.

More recently, Ramsey has criticized the iron-core theory on several grounds. Recent measurements, by S. K. Runcorn and others, of the variations of the Earth's magnetism with depth

suggest that the Earth is uniformly magnetized, so that an iron core is not needed to explain its magnetism though a liquid core still gives the simplest explanation of its variations. Further, Ramsey quotes the opinion of Jeffreys that the mean density of the planets Mercury and Mars makes it impossible for them to have the same chemical composition as the Earth, if the latter has an iron core. And it is not easy to suggest a possible origin for the Solar System that does not entail approximately the same constitution for all the planets of our Solar System. Thirdly, he points out that the argument from meteorites is invalid because the proportion of stony to iron meteorites is about 35 to 1, which could hardly be the case if they originated in the disruption of an iron-cored planet similar to the Earth.

Finally, he makes what seems to the present writer probably the strongest objection of all. From its diameter and mean density, it can be calculated that the Earth's core contains about 30 per cent of the Earth's whole mass, and when we add to this figure the iron content of the mantle and crust, we find that about 40 per cent of the Earth is iron: the percentage of iron in meteorites is about 25, and for solar matter, ignoring hydrogen and helium, about 20. The proportion of iron may, of course, be reduced by taking it to be alloyed with nickel, as its advocates usually do. But to reduce the iron to a reasonable amount means increasing the nickel far beyond what may be considered likely.

Fundamentally, of course, the reason for Ramsey's criticism of the iron core is that he has an alternative to offer. His suggestion is based on the theory of solids, and it requires the core of the Earth to be formed by the effect of high pressure on olivine. On theoretical grounds it is to be expected that this high pressure modification will be of much higher density than normal olivine, and that it will be metallic in its properties.

In a substance like olivine, the molecules are arranged in a regular crystalline form, and the electron shells are closed so that no electrons are free to wander through the material. It is, therefore, impossible for an electric current to flow through it. When the distances between such units are decreased, very

powerful repulsive forces come into action, making it difficult to compress them. In order to compress them, in fact, their structure must be changed, and this may be done in two ways. Firstly, the shape of the crystal may be changed, the new shape being one of smaller volume and therefore stable at higher pressures. In olivine, the normal crystal is rhombic, and a change to a cubic form would reduce its volume and, of course, increase its density. This change was suggested in 1936 by J. D. Bernal and H. Jeffreys to account for the 20° discontinuity. According to Bernal's calculations it gives an increase of about 9 per cent in the density, and the pressure needed for it is of the order of that at about 400 km below the Earth's surface (140,000 atmospheres). It seems, in fact, to provide a very satisfactory explanation of the 20° discontinuity, though, as we shall see later, it is not universally accepted.

Once the crystals have been compressed as much as possible, it is the structure of the molecules themselves that must next be broken. This means that electrons must be forced out of their normal places, and the energy needed to do this is very great. In hydrogen it is of the order of 10 electron-volts per molecule. For a complex substance like olivine, it must be even higher; unfortunately the calculation of the actual energy required is difficult, and it has not yet been possible to get more than a rough estimate. It seems likely, however, that the critical pressure is of the same order as that at the surface of the Earth's core (1·4 million atmospheres).

Unlike the change of crystal form, this second modification gives an increase in density of anything from 50 per cent upwards, so it is well able to account for the marked increase at the boundary of the core. Since it also sets electrons free to wander about inside the material, it changes it into a substance with metallic properties. A high pressure modification of olivine thus seems to meet the requirements of the material of the core, provided, of course, it can exist.

Ramsey's theory can also explain quite naturally the existence of the inner core. It is unlikely that the molecular structure will be completely broken down at once. Rather, we should expect

the outermost regions to yield first, and the inner ones only on application of a still higher pressure. Like the first, this pressure will have a critical value, and so it should give a transition with a fairly sharp boundary. The inner core thus represents the third step in the transformation of olivine into the kind of material found in white-dwarf stars.

The arguments for and against Ramsey's suggestion are in some cases highly technical, but we may summarize them briefly as follows. If the Earth is of homogeneous composition (apart from a plausible concentration of iron in its lower parts) the percentage of iron in it is reduced to about 30, which is more in accord with the apparent cosmic abundance than the 40 per cent demanded by the iron-core theory. If the 20° discontinuity is due to a change in crystal structure rather than a change of material, it seems logical to suppose that the other discontinuities may be explained in the same way. Since the change to a metallic form depends on the materials being subjected to pressures above a certain critical value, the theory also explains the accepted absence of cores in the smaller planets. They are just not massive enough to exert the required pressure at any point inside them. It is difficult to see how they could fail to form iron cores unless, as mentioned above, their chemical composition has always been different from that of the Earth.

The theory also explains the observed sharpness of the boundary between the core and the mantle. An iron core, separating out as in a blast furnace, might be expected to show a broad transition zone. The modified olivine, on the contrary, must show a fairly sharp boundary, since the change takes place at a well-defined critical pressure.

As for the existence of a metallic form of olivine, it is not outside the bounds of probability. There are, for example, metallic and non-metallic forms of arsenic, and the metallic form is more than twice as dense as the other, 5·1 gm per c.c. as against 2·0 gm per c.c. Also, a metallic form of phosphorus has been prepared by subjecting it to very high pressures, and again it shows a great increase in density.

In this connexion, it is interesting to note some recent work

by R. H. Christian and B. H. Alder, who have observed metallic properties in crystals of iodine and caesium iodide subjected to pressures of about 450,000 atmospheres – they produced these pressures for very short times by means of the blast from conventional high explosives. This result does not, of course, prove that Ramsey is correct. It probably removes any doubt that a metallic form of olivine might exist at sufficiently high pressures; but the question really is whether it appears at the pressure prevailing at the surface of the Earth's core.

Finally, Ramsey points out that his account of the Earth's interior accords well with H. N. Russell's suggestion that the planets and the white-dwarf stars are one family. In these stars, as is well known, the material is so highly compressed that it attains densities of the order of a ton per cubic inch. And this incredible density arises from the complete breakdown of the normal atomic structure, all the electrons being torn from their orbits to form what the atomic physicist calls degenerate matter. Thus the planets and the white dwarfs are like a single group in which the properties of matter are determined by pressure alone, independently of either temperature or chemical composition.

TABLE I. *Variation of Compressibility within the Earth*

Depth (km)	Pressure p*	Bulk Modulus k*	dk/dp
2500	1·1	5·7	3·2
2700	1·2	6·0	2·8
2900	1·3	6·4	2·9
3000	1·4	6·6	3·0

* The unit in columns 2 and 3 is 10^{12} dynes/cm^2.

This last argument appears to gain strength from some work by Bullen, which led him to formulate a 'compressibility-pressure' hypothesis. Table I shows the variation of the compressibility within the Earth, as derived from the velocities of the seismic waves, and it is clear from it that neither the value of

k nor its rate of change alters significantly at the boundary of the core. This leads Bullen to suggest that at great pressures the compressibility depends only on the pressure and is independent of the chemical composition of the material. But Ramsey deduces from the theory of solids that compressibility at pressures of the magnitude of those inside the Earth does in fact depend on atomic weight; and so he concludes that Bullen's hypothesis is true because the materials of the mantle and core are chemically identical.

The theory sounds very plausible; but there are objections to it, which cannot be ignored. H. C. Urey, for example, does not believe that the 20° discontinuity can be accounted for by changes induced by pressure. At the moment the pressure required is beyond the reach of experiment, but he estimates that magnesium germanate, a substance chemically related to olivine, should undergo the change, if at all, at accessible pressures. Yet all attempts to prepare a cubic modification in his laboratory have failed. Furthermore, the view of the 20° discontinuity taken by Ramsey, Bernal, and Jeffreys is not universally accepted. B. Gutenberg, for example, describes it as a gradual change in density in the region between about 400 and 1,000 km below the surface. If this should prove to be correct, it could not be explained by the change suggested by Bernal, which is essentially a 'catastrophic' one taking place at a fairly well-defined critical pressure.

In recent months, however, A. E. Ringwood has reported a thermodynamic study of a mixture of nickel germanate and magnesium silicate, which 'is considered to provide strong evidence supporting the Jeffreys-Bernal hypothesis'. The spreading of the change over a range of depths and pressures is then to be explained by the presence of other substances.

More serious still is an argument by W. M. Elsasser, who deduces from the thermodynamic properties of olivine that the high-pressure modification cannot exist. Elsasser also uses quantum theory to prove that iron, alloyed with nickel, would have the right density at the required pressures, so he favours the iron-core theory. Bullen, however, has pointed out that some

modifications of Elsasser's work may be required, which would tell against the iron core.

On the astronomical side, recent work on the planets Mars and Mercury have led to new values for their density and ellipticity, and these agree much less well than the old ones with those calculated from Ramsey's theory.

The latest work on this aspect of the subject has been published recently by Bullen, in a paper on 'The Constitution of Mars'. In this, he reviews the various suggested values of the mass, radius, and oblateness of Mars, and applies the most probable results to a model of Mars, related to his own model of the Earth. This leads to conclusions about the proportion of uncombined iron in each planet. He now assumes that the Earth's *inner* core consists of iron, which has separated out from a molten mixture of iron and olivine. The remaining uncombined iron is not enough to form the outer core, which must be a modification of the material of the mantle. Further, since iron and olivine are believed to be immiscible, he concludes that the 'outer core of the Earth is the one place where there is practically no uncombined iron', and any iron outside the inner core must be in the mantle. Once more this seems to suggest Gold's spongy mantle, with its rivulets of iron.

It will be noticed, incidentally, that Bullen offers a plausible alternative to Ramsey's second pressure modification as the cause of the inner core, and this one satisfies the feeling that, somehow, iron should separate out somewhere.

What, then, are we to deduce? The arguments on each side are strong, but not conclusive, and the wisest course is probably that counselled by Bullen, who says: 'Until such time as more direct experimental evidence can be brought to bear on the behaviour of materials at pressures of the order of a million atmospheres and above, or until the quantum-mechanical calculations are carried to greater refinement, it would seem that there is a special need for discussing the available evidence in terms of probabilities rather than in terms of definite conclusions.' Definite conclusions must thus await further work in seismology, astronomy, high-pressure physics, and quantum

theory, and since the work is proceeding apace in all these fields, there is reason to hope that a decision between the two theories may be made within the next few years.

This state of suspension of belief is the rigorously scientific one, but most of us will nevertheless have a bias towards one or other of the two theories. Preference depends to some extent on imponderables, but at the present time the author feels that Ramsey's theory is the more attractive, chiefly because it pictures a more uniform Earth and planets, and because it keeps the Earth's iron content in line with those expected from other sources. Perhaps also, it attracts by its newness, by breaking with tradition and so opening up once more a region that had been largely regarded as closed. On balance, too, it seems that the latest work is tending to favour the modified olivine core rather than the iron one.

The problem of temperature-distribution inside the Earth is complex, and only a very brief summary of possible conclusions can be attempted here. First it is of interest to note that the temperature of the Earth's surface is determined chiefly by sunshine. The heat coming from the inside is several thousand times less than that coming from the Sun. Nevertheless there is a substantial amount of heat flowing out through the Earth's crust, which shows that the temperature increases from the surface inwards. The same conclusion follows from the flow from active volcanoes of molten rock at around 1,100° C.

It is known that heat is being generated in the Earth's crust by radioactive elements, chiefly radium, thorium, and potassium. The respective rates for these elements are $8{\cdot}0 \times 10^{-2}$, $6{\cdot}3 \times 10^{-9}$, and $2{\cdot}5 \times 10^{-13}$ calories per gram per second, and if all the Earth contained the same proportion of these elements as the surface rocks, the heat generated would keep the whole Earth liquid. It is found, however, that the radioactivity is much higher in granitic rocks than in the basic and ultrabasic rocks below the crust. (In this connexion it is interesting to note that the background radiation in Aberdeen, the 'Granite City', is above average.) The generation of heat is therefore largely confined to a layer at the surface less than twenty miles thick. A

small amount in the interior is required by some theories of the magnetic field.

What, then, of the temperature in the deep interior? This is determined for the most part by the requirement that the core shall be liquid. Unfortunately, it is not certain just how melting points change at the high pressures in the core and mantle, and if the core is of metallic olivine, its melting point is not known at all. An iron core, however, could hardly have a temperature of over 2,000° C and be magnetic as some theories of the magnetic field require. One thing is certain. The conductivity of the mantle is so low that very little of the interior heat can have escaped since the Earth solidified, and the mantle, for the most part, cannot be much below its melting temperature.

The only reliable value of the temperature at considerable depths is derived from the electrical conductivity of the material. It is known from studies in geomagnetism that at a depth of some 900 km this has a value equal to that of silicates at temperatures of 1,000–1,500° C. Since the material of the mantle is mostly silicate, we may reasonably conclude that this represents the Earth's temperature about 1,000 km down. To go further, we must make some assumption about the early history of the Earth, if not about its origin.

Some current theories require the Earth to have been originally gaseous, subsequently liquefying, and later solidifying as it cooled. These would clearly imply very high temperatures inside – something of the order of 6,000° C. Other theories picture the Earth being formed by the accumulation of small cold bodies, though this also is usually followed by a stage of liquefaction, perhaps by the heat generated by radioactivity. If this were so, the temperature in the deep interior may be no more than 3–4,000° C. In any case, the value will depend on the amount of radioactivity in the mantle and core. As mentioned before, some theories of the Earth's magnetic field require this to be higher than the normal value for ultrabasic rock in the crust, and the temperature in the core could be raised appreciably by it.

The only conclusion possible at the moment is that the Earth's internal temperature probably rises from its surface value to a

value at the centre of several thousand degrees. But whether 'several' means 2–3 or 6–7 is a question the answer to which must await further work in several fields.

Finally, we may wonder where the liquid rock that flows out during a volcanic eruption comes from. It is not very plausible to think of it flowing all the way up from the core, and yet it is not the ordinary solid of the upper layers. The most attractive suggestion is perhaps that which links volcanoes with earthquakes. The energy released in an earthquake does not all go into seismic waves. Some of it is stored as thermal energy in material which in the course of time is heated above melting-point. The liquid rock so formed may then flow out when a weakness develops in the crustal layers.

This theory is not universally accepted, but it fits well with the fact that earthquakes and volcanic activity occur in the same clearly defined zones of the Earth. It brings us round again to the earthquakes from which so much of our knowledge of the Earth is derived. It reminds us once more that the Earth, like our knowledge of it, is constantly changing and that its changes are often as startling as they are unpredictable.

INFORMATION THEORY, PROBABILITY MODELS, AND PSYCHOLOGY

W. SLUCKIN

IN recent years experimental psychology has been stirred up by the impact of some new mathematical techniques. An impressive number of papers, monographs, and symposia are nowadays devoted to a variety of psychological uses of information theory. Fewer in number, but perhaps of no less significance, are publications concerned with probability models in psychology. What are these new fields of study? What is their significance, and where are they likely to lead?

What is information theory?

Information is a concept that we use very frequently in everyday life. It refers to items of knowledge – in the simplest terms, to what we are told. In the nineteen-twenties efforts were first made, mainly by telecommunication engineers and by statisticians, to refine this concept and make it more precise. Before long, philosophers, physicists, psychologists and others began to take an interest in and contributed to this activity. The meaning of the term information, as used technically, has been restricted in its range in certain ways.* For good measure, not just one but several precise meanings and uses have been given to the term information. Furthermore, the situation has been complicated by the inter-relatedness of these several uses.

We shall confine ourselves to the concept of information as defined and understood in one particular way; we are going to

* The same thing is commonly done with everyday words. Such words as force, power, current, resistance, resilience, reliability, and many others have been adapted to precise scientific uses.

be concerned with information which, to be distinguished from other kinds, is sometimes called selective information. We are going to concern ourselves only with selective information (*a*) because this conception of information is more widely discussed and applied than any other, and (*b*) because it is the theory of selective information that has particular relevance to psychology.

Let us begin by glancing at the historical background of this conception of information and the information theory that has developed from it.* The need to talk about information in a precise way arose in the field of telecommunications. From the start engineers wanted to know how to send a message by telegraph with the minimum of cost and a maximum of speed. To this end the choice of code had to be considered and so the frequency with which different letters occurred in the language in which messages were being sent. Some deterioration in clarity is, of course, inevitable when a message is being transmitted; what had to be considered was how much deterioration would be permissible. Thus it became necessary to know the greatest amount of information that could be reliably transmitted in a given time at a given cost.

As early as 1928, R. V. L. Hartley of the Bell Telephone Laboratories in the U.S.A. pointed out that, from the standpoint of the recipient, information is provided by the unexpected. What is fully expected is devoid of information. If the reader is quite certain what my next sentence is going to be, he will gain no information by reading it.

From the point of view of the recipient of signals, a signal to which there are no possible alternatives conveys no information. The greater the number of alternative possibilities, the greater is the amount of information conveyed by one particular signal. For example, a single digit of the decimal system of number notation will transmit more information than a single digit of a binary system, such as one that consists of dots and dashes.

Notice that the semantic aspects of information are irrelevant

* When the word information is used without a qualification it generally refers to selective information.

to this discussion. We are concerned with distinguishable signals, with the amount of code that is transmitted, and not with what the code stands for. We are not concerned with the meaning of the information conveyed. In everyday speech, however, the word information does refer, of course, to the semantical content of messages. Therefore, the phrase 'theory of information' is perhaps a somewhat misleading one. Indeed, a more reasonably descriptive name for this field of study might be 'statistical theory of signal transmission'.

A further terminological complication, and not only terminological, is that there is another theory of information which in fact parallels the one with which we are concerned. This theory *is* concerned with the transmission of meaning (or semantical information); it is Carnap's 'theory of semantical content'. In this field we might start off with the assertion that a statement that something will occur conveys a considerable amount of information only if this something is unlikely to occur. A statement that something will occur conveys little information if this something is very likely to occur. Thus, if somebody tells me that the sun will rise tomorrow, he tells me very little; his message conveys next to no information because the sun, judging by what has happened before, will in fact rise tomorrow. But if someone predicts that it will snow in England next July, then his message conveys very much information.

However this may be, our interest is confined to selective information. If signals, or unit messages, are all equi-probable, the probability of each is inversely proportional to the number of possible choices. The probability of an event is, in this sense, its relative frequency; in statistics, probability and relative frequency are synonymous. Thus, the more improbable the signal the more information it conveys; in other words, the information per signal varies inversely with the signal's probability.

The measurement of information

The unit of information is arbitrarily but conveniently defined as the amount of information that is conveyed by a signal which is one of only two possible signals. If 'yes' or 'no' are assumed

to be the only two possible answers to the question 'Will you marry me?' the actual answer will yield just one unit of information, whether, in fact, it happens to be 'yes' or 'no'. Such a unit of information is a binary unit, called for short binit, or bit. The binary unit is the accepted standard unit of expressing in quantitative terms the information content of signals or element messages.

Whenever, as far as the recipient is concerned, the number of alternative signals is reduced by half, one bit of information is passed. This is so when two choices are reduced to one, or ten reduced to five, or twenty reduced to ten, etc. If there are four choices, then to reduce them to one only, two binary steps must be made; therefore a signal that is one of four possible signals conveys two bits of information. Similarly, one of eight alternatives, requiring three 50 per cent reductions, conveys three bits of information. And one of six choices supplies less than three but more than two bits. In general, information content is measured by the logarithm to the base of two of the number of alternatives. If H is the information per signal,

$$H = \log_2 (\text{number of choices}).$$

The probability of a signal is inversely proportional to the number of choices. If $p(s)$ denotes this probability, then

$$H = \log_2 (1/p(s))$$
$$\text{or, } H = -\log_2 p(s).$$

In words, the amount of information contained in a signal is measured by the negative log (to the base of two) of the probability of the signal.

A message consists of a number of signals. We have seen that the amount of information in a particular signal depends on its probability. The probability of a signal, or its relative frequency, has so far been assumed to be constant. If the theory of information is to be useful, it must apply to signals that are not characterized by a constant probability; for all kinds of signals encountered in practice have probabilities that vary from message to message. Thus the amount of information in a signal must be expressed in terms of the signal's average probability.

If the probability of the particular signal in the i^{th} message is p_i, then the information contained by the signal in that message is equal to $-\log_2 p_i(s)$, or simply, $-\log_2 p_i$.

However, the contribution of the i^{th} message in the final reckoning of the information per signal is in proportion to the relative frequency of the signal in the message. So the contribution of each i^{th} message is given by $p_i(-\log_2 p_i)$. And the information contained in the signal, when all the messages, say k in number, are taken into account, is

$$H = \sum_{i=1}^{i=k} p_i (-\log_2 p_i)$$

where i has all the possible values from 1 to k. Notice that the value of the information in a signal is given by the weighted average information per signal, taking into account all the messages.

Suppose now that we send a series of messages; let us call this series, *series x*. If the channel along which the messages are transmitted were a perfect one, *series x* would arrive at the receiving end. But channels are often imperfect. There may be some information lost in transmission; and some spurious information may be picked up. As a result, what is received is a more or less distorted series of messages; let us call this series, *series y*.

For any series of messages the average information per signal is given by H, as defined above; its calculation is based on the frequencies of the occurrence of the signal. The quantities that may now be considered are as follows:

$H(x)$: input information per signal,

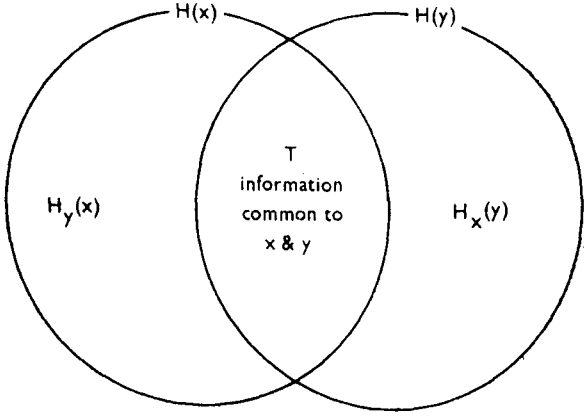
$H(y)$: output information per signal,

$H_y(x)$: information per signal lost in transmission (sometimes called equivocation),

$H_x(y)$: information per signal added in transmission (sometimes called noise),

T: information per signal actually transmitted.

The manner in which these quantities are inter-related may be represented in a diagram, as by G. A. Miller of Harvard.



The total area of this diagram, denoted by $H(xy)$, consists of the sum of information lost, transmitted and added. Thus

$H(xy) = H_y(x) + T + H_x(y)$

$= H(x) + H_x(y)$, i.e., input plus noise

$= H(y) + H_y(x)$, i.e., output plus equivocation.

Transmitted information is given by

$T = H(x) - H_y(x)$, i.e., input minus equivocation

$= H(y) - H_x(y)$, i.e., output minus noise.

It is important to note that normally, as the input information per signal, $H(x)$, increases, the transmitted information, T, approaches its limit which is prescribed by the channel capacity, C, of the given communication medium.

Although this general picture of the transmission of information derives from telecommunication engineering, it has some applicability in many other fields; and some of these fields are to a greater or lesser extent the concern of psychology.

In the broadest terms, it is obvious that information is received from the external world in perception. Information may be stored in memory. Again, when responses of any kind are made, information is generated. Any organism might be regarded as a transmitter and receiver in a communication system. What use is it to adopt such a viewpoint?

It is clear that the analogy between the simple model for the transmission of information and the human situation that is of interest to psychology has many important limitations. Man performs many functions other than those of a communicational transducer. It is very easy to point out in what ways man is not a mere mechanism for handling information. But this need not be our concern here.

It is interesting that, despite its grave limitations, the informational model may be of some use to experimental psychology; of how much use, it is perhaps rather too early to judge. It is perhaps more interesting to explore the admissible uses of information theory in psychology than to apply oneself to showing that in some situations the theory is a useless one. So let us examine some of the principal ways in which experimental psychology has made use of information theory.

Reaction times

The reaction time of a subject is the time that it takes the subject to react, or to start a response, to a given stimulus. This delay in response is taken up by the arousal of the stimulated sense organ, the passage of the resulting nervous impulse to the brain, the activity of the brain, the passage of a requisite impulse to the responding muscles, and, possibly, the preparation of the muscles for action. The reaction time of a subject is also known as his response latency; the latter phrase refers to the fact that the effect of a stimulus remains latent until the response takes place.

On and off for over a hundred years reaction times have been investigated in psychological laboratories. Wundt in Germany, followed by J. McK. Cattell in America, and more recently Piéron in France, and many other investigators have expended – and some think, wasted – much time in such studies.

The interest in reaction times had largely died down, when W. E. Hick in Cambridge made an altogether fresh approach to

this field. In a paper published in 1951 he suggested that while a subject is responding to a stimulus, his *rate* of gain of information is constant. Hence, longer reaction time (response latency) is associated with greater informational gain.

Since the eighteen-eighties it has been known that when a subject has to react appropriately to one of a number of alternative stimuli his reaction time is longer than when he reacts to only one possible stimulus. In the language of the psychological laboratory, multiple-choice (or, disjunctive) reaction times are longer than simple reaction times, the reaction times increasing as the number of alternative stimuli and responses increase. Hick suggested that the time taken to respond to a stimulus is proportional to the amount of information provided by the stimulus. The average multiple-choice reaction time is longer than the average simple reaction time because the signal for a disjunctive reaction provides more information than the signal for a simple reaction.

We know that a message to which there are no alternatives conveys no information; thus, $\log_2 1 = 0$. It might appear at first sight that the stimulus in a simple reaction-time experiment provides no information. This, however, is not the case. The alternative to the presence of stimulus is the absence of stimulus. The information gained when simple reaction occurs is one bit; $\log_2$ (two alternatives) $= 1$. If there are two alternative stimuli, the information gained by the presentation of either is $\log_2$ (choice of three). Thus, if n denotes the number of alternative stimuli, the information gained per presentation is $\log_2(n+1)$ bits. And the disjunctive reaction time is proportional to $\log_2(n+1)$.

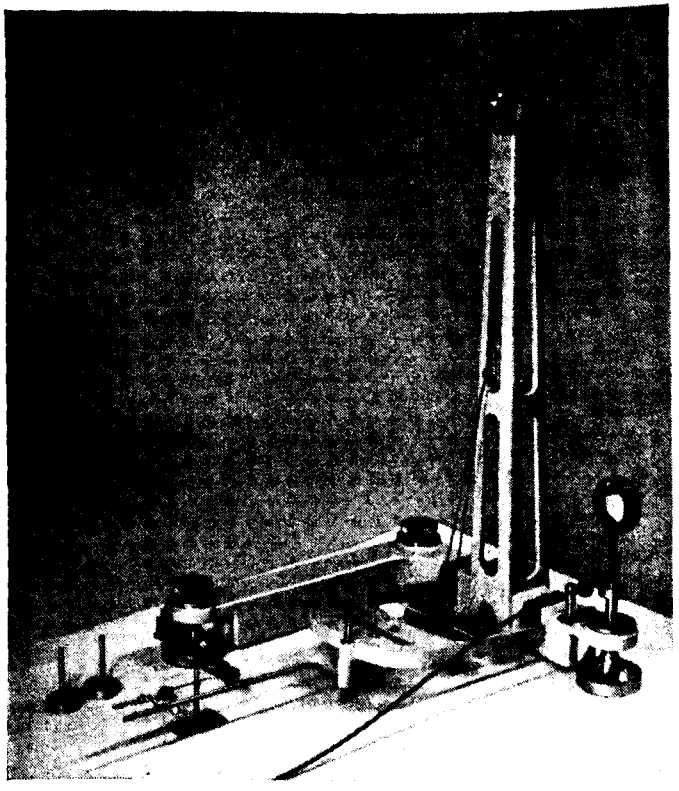
Hick used both data from old experiments and data from his own experiments, and found that disjunctive reaction times were in fact on the whole proportional to $\log_2(n+1)$. If T is the reaction time, then

$$T = k \log_2(n+1) \text{ milliseconds,}$$

where k is a constant to be determined empirically.

A number of studies along such lines were also initiated by other investigators, notably by R. Hyman in America. Later

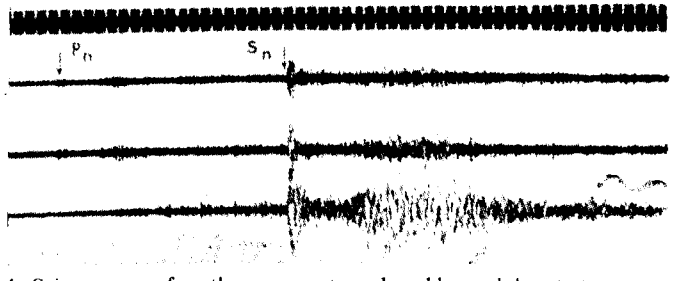
1. An experimental sign plate using an electroluminescent light source. The structure of the plate is similar to that described in Figure 1*b* of H. K. Henisch's article; zinc sulphide phosphor, shaped electrodes. Excitation at 300 volts and 400 c/s. (*The experimental unit was made available through the courtesy of Dr S. T. Henderson of Thorn Electrical Industries Ltd.*)



2. Seismograph designed by the late J. J. Shaw, C.B.E., who spent many years developing highly sensitive instruments for detecting and recording earth tremors caused by distant earthquakes.

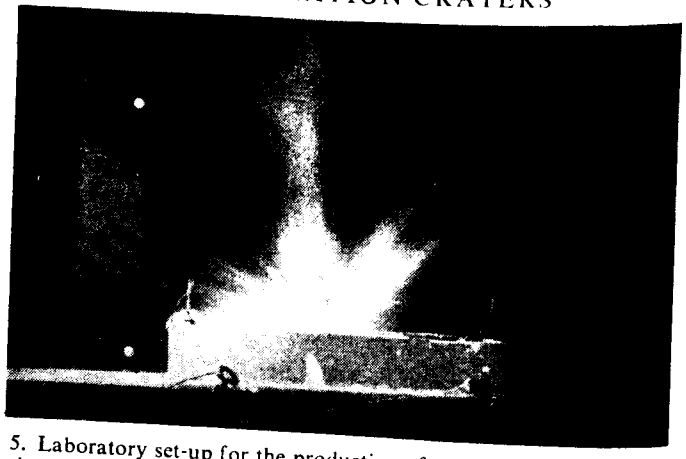
The general principle of the instrument is to multiply and record photographically the movement of a boom 19 inches long which carries a 1-lb weight, together with a damping vane that moves in a strong magnetic field and brings the boom to rest after each excursion. Ideally the instrument is mounted on a substantial masonry pier built up from a rock foundation. Any ground movement can, if necessary, be multiplied 500 times. (*Courtesy of Hilger and Watts Ltd.*)

3. For much geophysical research a field seismograph is needed. That shown here was developed by P. L. Willmore in the Department of Geodesy and Geophysics in the University of Cambridge. The seismometer unit is sealed in a metal case which can, if necessary, be buried in open ground. (*Courtesy of Hilger and Watts Ltd.*)

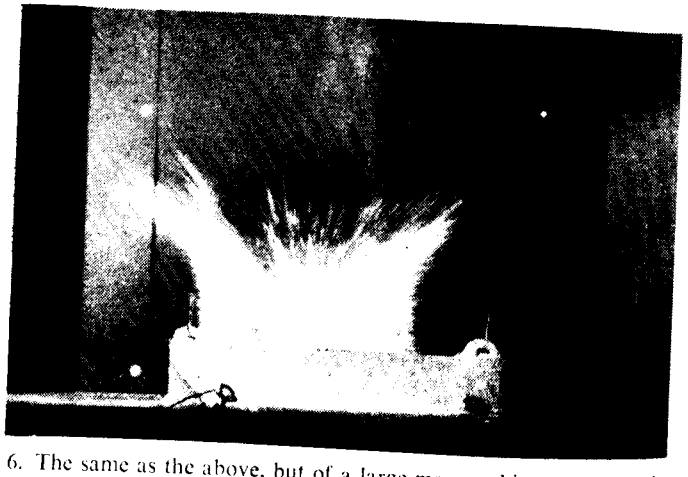


4. Seismogram of earth movement produced by rock burst at a range of 389 kilometres, recorded by a three-component short period seismometer. (*From Monthly Notices of the Royal Astronomical Society* (*Geophysical Supplement*) 1950 **6** 129, *by permission of the Author and the Society.*)

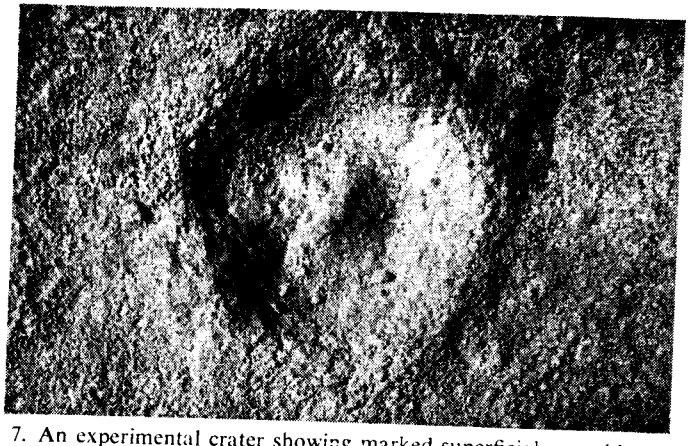
DISINTEGRATION CRATERS



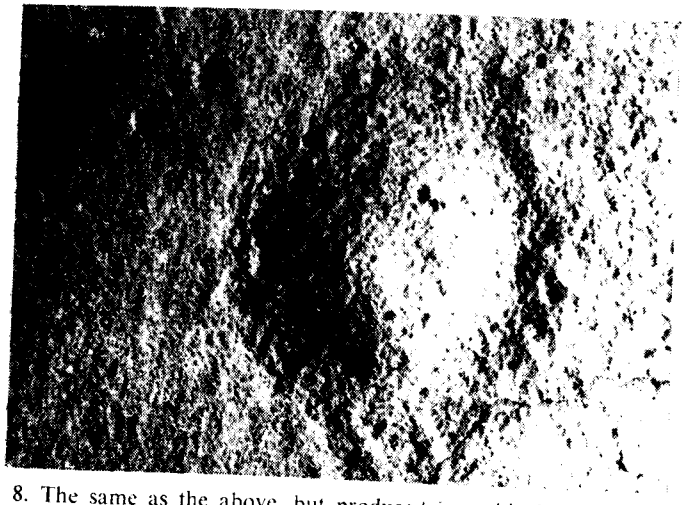
5. Laboratory set-up for the production of experimental disintegration craters. A falling powder missile brings two wires into contact and so actuates the flash-gun by which the photograph is taken. The impact illustrated here is of a small missile on shallow base powder.



6. The same as the above, but of a large mass making an energetic impact with deep base powder. In both photographs the vertical cloud results from a small quantity of powder left behind by the falling missile.



7. An experimental crater showing marked superficial resemblance to many lunar features. (See, for example, the illustrations in *Science News 45*). This crater was produced by a large missile in deep base powder. Light falling from left to right.

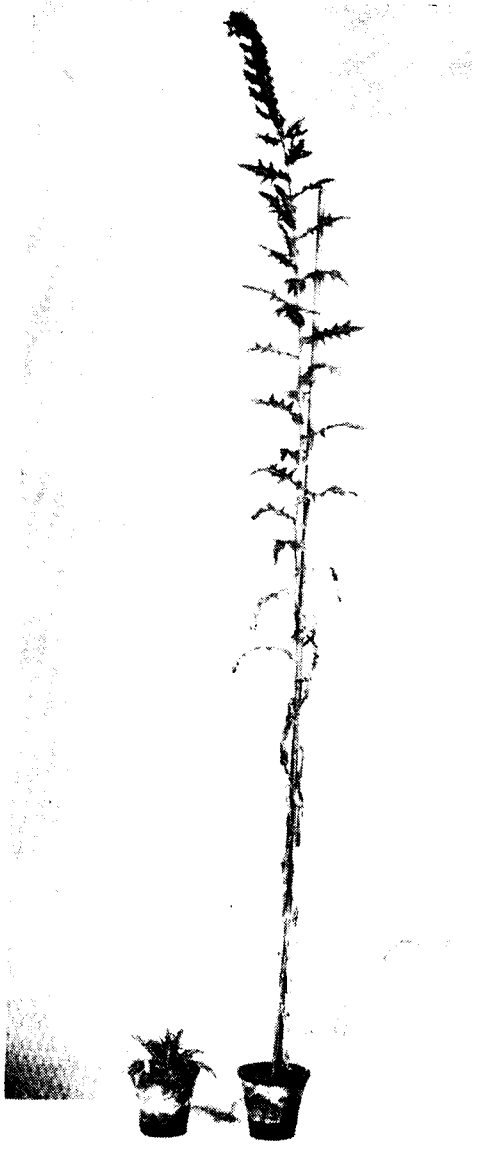


8. The same as the above, but produced in a thin layer of base powder by a medium missile. These conditions give a pronounced central mountain. The nearby secondary craterlets should also be noted. Light falling from left to right as above.

(a)

(b)

9. The effect of gibberellic acid on Common Henbane (*Hyoscyamus niger* biennial strain). This plant remains vegetative unless 'vernalized' by low temperature treatment. In each of the illustrations the untreated plant is on the left, and a plant treated with 500 μg of gibberellic acid on the right. (*a*) (*above*) 10 weeks after treatment, (*b*) (*below*) 12 weeks after treatment, and (*c*) (*opposite*) 22 weeks after treatment. (*Courtesy of Dr P. W. Brian of the Akers Research Laboratory (I.C.I. Ltd), Welwyn, Herts.*)



(c)

10. The effect of increasing doses of gibberellic acid on the growth of dwarf pea seedlings. Reading from the right: nos. 1 and 2 are untreated; no. 3 received 0·01 μg per plant; no. 4, 0·02 μg, and thence doses increasing twofold. The photograph was taken 29 days after treatment. (*Courtesy of Dr P. W. Brian of the Akers Research Laboratory (I.C.I. Ltd), Welwyn, Herts.*)

investigators interested themselves in situations where signals to which the subject has to react are not all equi-probable, some signals occurring, in fact, more frequently than others. In these circumstances the subject's response time to a signal tends to be proportional to the *average* information content in this signal.

PSYCHOPHYSICS

Originally psychophysics was conceived as a study of the relationship between physical stimuli and the corresponding sensations. Sensations are always private; they cannot be publicly investigated or measured. But sensations mediate between stimuli and responses. And responses can be observed and measured. So nowadays psychophysics is in fact concerned with stimulus-response relations. Interpreted broadly, this would perhaps include the whole field of perception and learning, and not merely studies of reaction times and the like. In point of fact, however, only a limited area of psychology goes traditionally by the name of psychophysics. It includes primarily investigations of the absolute and differential sensory thresholds, and the construction of subjective sensory scales. It also includes experiments in which the subject is called upon to identify stimuli, or to make absolute judgements. It is these experiments that are of especial interest to us.

All judgements are, of course, essentially relative. It is traditional, however, in psychological laboratories to refer to judgements in which particular stimuli are identified, rather than compared with other stimuli, as absolute judgements. Consider an experimental situation where there are k stimuli varying in magnitude along a single dimension, e.g., a series of light circles on a screen, the circles varying in size only. The circles are presented to the subject one by one, and the subject is required to assign each to one of m categories, where m may or may not be equal to k. Of course, the simplest situation arises where $m=k$, that is when the subject is given a genuine opportunity to identify correctly the magnitude of each stimulus.

Assume that stimuli are presented to the subject in a random

order again and again. For the sake of numerical example, imagine that $k=m=6$, that is, that there are six stimulus categories and six possible responses; and suppose further that each stimulus is presented twenty times altogether. Then, the responses may be set out in a scatter diagram or contingency table, as shown below.

		Stimuli (k)						
		1	2	3	4	5	6	N_m
Responses (m)	1							
	2							
	3							
	4							
	5							
	6							
N_k		20	20	20	20	20	20	N_m120

N_k-values in the bottom row are numbers of presentations of each stimulus category, in our case twenty, and N_m-values in the extreme right-hand-side column are numbers of responses that fell into each response category.

Let us now refer again to the two overlapping circles that represent the flow of information (p. 70) $H(x)$, input information, is now stimulus information, and $H(y)$, output information, is response information. Information may be lost or added in transmission, so it is not out of place to talk in this connexion about equivocation, $H_y(x)$, and noise, $H_x(y)$.

The transmitted information, T, indicates the degree of dependence of responses upon stimuli; it therefore indicates the subject's ability to discriminate between the different stimulus categories. Thus T may serve as a measure of the subject's power of discrimination. If the transmitted information nearly equals the input and output information (overlap of circles being great), then the stimuli are well discriminated. If transmitted information is relatively small (the overlap of circles being slight), then the stimuli are poorly discriminated.

The channel capacity, C, is the capacity of the subject to discriminate among stimuli of the given kind. Where stimuli vary along a single dimension (e.g., circles differ in size only, or musical notes are alike in every respect except loudness), the subject will generally be able to discriminate perfectly only between a certain limited number of stimuli, perhaps seven or eight or nine, and no more. This number is a function of, or depends upon, the channel capacity.

Consider what happens as the number of stimuli that are presented to the subject increases. Input information increases with each additional stimulus category. So at first does the transmitted information. After a certain value of input information is reached, transmitted information will increase more and more slowly, until it reaches its upper limit C. Continued increase in the number of stimuli will now lead to a rapid increase in the number of response errors, thus preventing T from rising above some value C, which corresponds to the maximal discrimination of which the subject is capable.

Garner and Hake, of Johns Hopkins University in the U.S.A., have shown how to calculate from the frequencies of occurrence of various responses, as entered in a contingency table such as that shown on p. 70, the amounts of stimulus and response information and of 'noise'. From these the amount of transmitted information, or discrimination, can be determined. This quantity is approximately equal to the logarithm of the number of perfectly discriminable stimulus categories. In other words, the nearest integer to 2^T is the number of perfectly discriminable stimuli. The number of such perfectly discriminable stimuli is called by Miller the span of absolute judgement. The span of absolute judgement has been investigated for different kinds of stimulation in all the sensory modalities.

A point of considerable interest is that the maximal transmission of information is attained when stimuli differ from one another in a manner that makes them equally discriminable.

Scales for equal discriminability have been constructed by several investigators for several kinds of stimulation.

We cannot leave this topic without referring to the obvious fact that outside the laboratory people are constantly called upon to identify objects that differ from other objects, not just in one way, but rather in a number of ways. There is no reason why precise investigations of this kind of judgement should not be carried out under controlled laboratory conditions. Several studies have, in fact, been conducted where stimuli vary along two or more dimensions. For example, discs of light may vary both in size and in brightness, or sounds may vary in pitch and loudness at the same time. In these circumstances the channel capacity for transmitting information is expanded; that is, the number of perfectly discriminable categories increases with the number of stimulus dimensions, although the function does not in general appear to be a linear one.

Verbal communication

A less obvious and, therefore, all the more interesting application of information theory in psychology is in the field of certain formal aspects of verbal communication. Spoken language has various constraints inherent in it. Generally the words we use at any given time determine to some extent, sometimes slight and sometimes considerable, the words that we shall use next. Words tend to follow certain sequences; they do not occur at random. Some contexts limit quite severely the number of possibles from which the next word can be chosen; other contexts impose a less severe limitation upon the choice of subsequent word or words.

The context tells us something about the word to follow by reducing our freedom of choice as to the next word. We can determine this reduction of freedom from the knowledge of the relative frequency with which word sequences of various kinds occur in the language. This enables us to calculate the information that is provided by the context about the next item. And when this item occurs, some of the information that it provides is the same information as that which has already been supplied

by the preceding context. The part of the information conveyed by the word that had been available before the word occurred is, in a sense, superfluous or redundant. However, such redundant information is by no means necessarily worthless, as we shall see later.

Let us now establish the analogy between our picture of the flow of information and the sequential information of verbal communication. Referring again to the diagram on p. 70, $H(x)$, input information, is the information provided by the context. $H(y)$ does not now represent output information; it is in this case another input information. $H(y)$, in this sequential situation, is the information from the word that follows the context. Thus we have here two sources of input information that may or may not be related. If they are related, the two circles representing the two inputs overlap. The overlap of information from the two sources, T, represents the redundant information.

To pursue the interpretation of the diagram still further, $H_y(x)$ is the information in the context that has no bearing on the word that follows. $H_x(y)$ is the information that is not contained in the context. $H(xy)$ is the total amount of information from both the context and the word that follows.

The quantity $H_x(y)$ represents the fresh information that is supplied by a word. This information may be measured for each new word in any text. The average amount of information of this kind expressed in bits indicates the rate, in bits per word, at which the text generates information.

In this manner Miller and his co-workers have made extensive studies of sequences of letters in written English. If letters followed one another at random, each would generate about 4·7 bits of information (equal to $\log_2 26$, there being 26 letters in the alphabet). In fact, owing to constraints inherent in written English, each new letter generates on the average in standard English about 1·4 bits of information. But this is an average. Some letters generate a good deal more than this. Some letters generate hardly any new information. The letter u when preceded by q provides information which is wholly redundant; it adds no fresh information; to drop it would make written English no less intelligible.

A moderate amount of redundancy in verbal communication is a useful thing. It ensures that an occasional error in transmission or reception has a good chance of being corrected. There would be no need for verbal redundancy as an aid to comprehension if errors did not occur. But they do occur inadvertently, for example, through lapses of attention. However, to say that some degree of redundancy is useful is not to favour excessive redundancy.

OTHER USES OF INFORMATION THEORY

It might be argued that the study of verbal communication in the manner described in the last section is peripheral to experimental psychology. But we might have chosen for discussion several other fields of psychology which have made, or are attempting to make, use of information theory. Let us glance at some of them.

Memory has long been studied in the psychological laboratory. Informational analysis has been used both in experiments designed to investigate forgetting over appreciable periods of time, and in experiments concerned with the span of immediate memory. More recently Brown, of Birkbeck College, in the University of London, attempted to consider immediate and delayed memory as part of the same process. As he puts it, a memory trace may adequately specify an item even after some decay of the trace, i.e., after some initial features of the trace have been lost. This is so because of some redundancy in the initial trace. Immediate memory span is a function of the time that it takes some of the early items to decay below the threshold at which they can be adequately specified.

In the study of motor skills, Fitts, in America, concerned himself with testing the hypothesis that motor systems have a fixed informational capacity. He and others have also applied the information theory approach to problems concerned with the display of signals and with motor controls; these problems are part of what has come to be known as human engineering.

In another field of applied psychology Bendig and others have investigated the worth of various lengths of rating scales, using

information measures. Information transmitted by each scale length was computed. This turned out to be correlated with the logarithm of the number of the scale categories. One study indicated that the information transmitted was a constant proportion of what might have been transmitted under optimal (noiseless) conditions.

Bendig has also applied information theory to class-concept formation, where the class concept is arrived at by the process of elimination as in the game of 'twenty questions'. Hovland and Weiss considered the transmission of information in the formation of concepts through positive and negative instances.

One might add that uses other than those briefly mentioned here have been made of information theory in psychology, including attempts to apply it to test situations, and to the recoding procedures that are the basis of all mnemonic devices. It is quite remarkable how much work of this kind has been done since about 1950; there is little doubt that more will be done yet.

STOCHASTIC PROCESSES AND BEHAVIOUR

Consider a sequence of events each of which happens in accordance with some particular probability; it may be a sequence of mathematical symbols, or of letters in a book, or of weather on consecutive days, or of responses of an animal, or any such sequence. A time series of this kind is known as a stochastic process. In recent years many attempts have been made to regard behaviour sequences as stochastic processes, particularly sequences of events in learning behaviour.

Psychologists have always been concerned with showing that seemingly patternless acts of behaviour do, in fact, conform to certain definite patterns. Assumptions are often made about the character of such patterns or sequences, and explanations are suggested as to why they are such as they are. However, the initial difficulty of discerning a pattern at all, and of being able to describe it unambiguously, is greater than might be supposed. Furthermore, any acceptable explanation that may be offered with regard to a sequence of acts of behaviour is *ipso facto* rooted in the specification of the sequence in question.

Many of the time series of behaviour events that one might wish to examine are not, of course, of the very simple kind in which the recurring events have a constant, unchanging probability. In the more complex series the probability of a given act of behaviour might depend on the particular position of the act in the time series. Fortunately, mathematical descriptions of some such processes are available. A relatively simple stochastic process of this kind is one in which the probability of transition from one state to another depends solely on the first of the two states. This type of sequence of events, where each event is directly related only to the event immediately before it, is known as the Markov Chain, after the Russian mathematician A. A. Markov who, in the first decade of this century, made a special study of such series.

Behaviour sequences that do not exhibit any simple regularity sometimes turn out to be susceptible to description in terms of Markov processes. It might be thought that the characteristic nature of the Markov Chain would severely limit its applicability to real behaviour sequences, but the limitation is a little less severe than might appear at first sight. This is because whenever the transition probabilities depend in a *constant* way on a number of earlier events, the time series can be restated (or recoded) so as to form a Markov Chain. Even so, in real life situations, the probabilities of events about to occur may depend in some inconstant and complex way on very many previous events.

The application of stochastic descriptions of behaviour seems most promising when behaviour sequences may be suspected of statistical regularities that are relatively simple or easily specifiable. Learning processes would seem, on the face of it, to meet these conditions. It is not surprising, therefore, that in recent years many attempts have been made to fit stochastic models to learning situations.

PROBABILITY MODELS IN LEARNING

The process of learning in an animal or a human being under controlled conditions may often be analysed into a series of

right and wrong moves. Whenever this is so, a statistical analysis of the sequence of moves or trials may be attempted. Broadly speaking, as trials proceed the probability of success in each new trial increases, this probability depending in some way on some or all the previous moves: successes, or failures, or both.

Early attempts at analyses of this kind were made in the thirties. But it appears that their significance was not very widely recognized. Further steps in this direction were made in the early nineteen-fifties. And very recently considerable progress was made by Bush and Mosteller in America, and by Audley in Britain.

The basis of the Bush-Mosteller approach to providing a stochastic model for learning may perhaps be best understood by a reference to their own simple example. A hungry rat is placed in a T-maze; it runs along the lane to the junction, and then has to turn right or left. Food is found at the end of one of the two lanes, always the same one. If the rat takes the wrong turning, it is removed from the maze and is placed at the entrance again. As the rat learns, the errors become less and less frequent. As trials go on, the probability of a correct move becomes higher and higher. For simplicity it may be supposed in the first place that this probability depends only on the number of trials and increases by some constant fraction. This fraction is a parameter that may be estimated from actual experimental data, if indeed the data do fit the model (some acceptable criterion of goodness of fit is of course necessary). To get a better agreement between theory and fact, it may well be necessary to complicate the model and introduce another parameter; thus, the fraction by which the probability of a correct move is determined may be supposed to depend on two constants, one associated with previous successes, and one with previous failures. And to get a fully acceptable fit between theoretical expectations and experimental data certain further complications may have to be introduced into the model – so much so, that Bush and Mosteller have written a very sizeable book about problems of this kind.

While Bush and Mosteller derive their stochastic models of

learning with the help of what are called operators which transform one function into another, Audley and Jonckheere in a recent paper start off by considering an imaginary situation where coloured balls are successively drawn from an urn. They consider what happens to the probabilities of drawing a ball of one of two colours, when after each drawing the ball is replaced and the contents of the urn are changed in a way which depends on whether one kind of ball or the other had been drawn.

It will by now be clear that drawing one kind of ball may represent success in trial-and-error learning, and drawing another kind may represent failure. One alters the probability of the next success (or failure) by altering each time in some suitably prescribed manner the contents of the urn. It is possible to introduce various constraints into such a general scheme simplifying the model as far as it is compatible with experimental data. It is desirable, for example, to be able to express the process as a Markov chain, the characteristics of which are mathematically known and manageable. It is noteworthy that the Audley-Jonckheere approach and the Bush-Mosteller approach are ultimately reducible to one another.

The learning curves traditionally studied in experimental psychology are in general shape much like smooth growth curves. Such smooth learning curves are obtained by the addition of a large number of markedly jagged graphs, each of which relates a single individual's performance to a time function. The parameters of the traditional learning curves ignore both the differences between individuals, and the possible (and, indeed, very probable) interdependence of the observed responses of each individual subject.

Bush and Mosteller do, of course, recognize the interdependence of responses; indeed, this is the very basis of their work. They assume, however, the same parameters for all subjects. Audley considers his subjects individually because they may differ from one another in that each is characterized by some *particular* probability of responding correctly on the first trial and of responding subsequently in accordance with his earlier successes or failures.

The new approach to the study of learning consists essentially in assuming that the learning process under investigation conforms to some stochastic process. Parameters are then empirically established so as to increase as much as possible the likelihood of the observed data fitting the theoretical schema. When this is achieved the learning process has been accurately described; and dependable predictions about such a learning process may be made. What kind or what level of explanation of learning processes this approach provides is for the philosophers of science to attempt to say. That it tells us something about learning that we have not known before is beyond any reasonable doubt.

Information theory and probability models have found a place for themselves in psychology. Theorizing and experimentation in both go hand in hand, and both proceed at present at a fairly fast rate. There should be much to hold our attention in this sphere in the coming years.

FURTHER READING

Bush, R. R., and Mosteller, F.: *Stochastic Models for Learning* (New York, Wiley, 1955).

Cherry, C. (Ed): *Information Theory* (London, Butterworth, 1956).

Quastler, H.: *Information Theory in Psychology* (Glencoe, Ill., The Free Press, 1955).

Shannon, C. E., and Weaver, W.: *The Mathematical Theory of Communication* (Urbana, University of Illinois Press, 1949).

Further information may be found in a publication *Current Trends in Information Theory* (University of Pittsburgh Press, 1953), and in papers in recent volumes of the following journals: *British Journal of Statistical Psychology, Quarterly Journal of Experimental Psychology, Journal of Experimental Psychology* (American), and the *Psychological Review*.

THE OESTROGENS AND THEIR METABOLISM

P. H. JELLINCK

THE striking changes that occur in man and the more highly developed animals when either the testes or ovaries are removed have long been recognized, but less than a century has elapsed since it was realized that these reproductive organs were also endocrine glands, which passed their physiologically active products, the sex hormones, directly into the blood, to be carried to distant parts of the body.

Testes produce mainly masculinizing compounds, collectively known as androgens, while the corresponding feminizing products of the ovaries are called oestrogens.

The physiological and psychological effects of these hormones are best illustrated by considering the results of excision of the sex glands (castration), but first the characters of sex should be defined. The organs, other than the sex glands (gonads), that are essential for procreation, such as the external genitalia, as well as the uterus, Fallopian tubes, and vagina of the female, and the seminal vesicles and prostate of the male, are known as the sex accessories, while the other characters of sex that make their appearance at puberty, and are chiefly concerned with attracting the opposite sex, and with birth and nutrition, are called the secondary sex characters. These include the development of the mammary glands in women, differences in hair distribution and in the size and shape of the pelvis and larynx in the two sexes, as well as the psychological manifestations of sex. Other examples are the antlers of stags, the very distinctive plumage of birds, and the reddening and swelling of the sexual skin of female monkeys during a certain stage of their menstrual cycle.

Effects of castration

Removal of the gonads from young animals prevents the mature development of the sex accessories, and the appearance of the secondary sex characters.

Castration of the Leghorn cockerel produces atrophy of the comb and wattles, the development of more body fat than in the normal bird, and it abolishes the psychological responses characteristic of the male. Courtship behaviour is lost and pugnacity diminished. Removal of the ovary (ovariectomy) from a young hen results in the growth of cock-like spurs, a large comb, and male plumage. The antlers of young stags do not develop after castration, and ovariectomized female deer may produce horns.

The effects of castration on cattle, horses, and other domestic animals are well known, and in the past this operation has even been practised on human beings. We need but mention the eunuchs of Oriental countries who could be entrusted to take charge of the women in the harem, and the *castrati*, or choir boys, who were emasculated to keep their soprano voices. Castration of boys before puberty results in a feminine distribution of fat, a high-pitched voice, very little hair on the face and body, and suppression of male sexual feeling. It can also retard ossification of the epiphyses of the long bones, with consequent enlargement of stature. Ovariectomy is followed by corresponding effects; if performed before puberty characteristic feminine attributes do not appear; the girl tends to become mannish in type, accessory organs fail to develop fully, and menstruation does not occur. In many female animals that will only mate during certain periods when they are in oestrus (sexual heat), the rhythmic oestrus cycles do not ensue, or are completely abolished. If castration is carried out after puberty the effects are not so marked and sexual desire is often retained.

Effects of transplantation

It has been found that the majority of these striking changes can be prevented if the excised testes or ovaries are transplanted

to a different body site. This suggests that the gonads produce an internal secretion, and are in fact endocrine glands. Thus, in 1849, Berthold showed that transplantation of the testes of a castrated cock annuls the usual effects on the comb, a discovery that was ignored until, forty years later, Brown Séquard made exaggerated claims of increased vigour and bodily well-being after injecting macerated testes into his own body. Present information indicates that his testicular extracts did not in fact contain sufficient male sex hormones to produce any physiological effect whatsoever. Attempts at rejuvenation by operations in which monkey testes ('monkey glands') were grafted, were unsuccessful because of differences in species.

Transplantation has also been performed in the human female, after complete excision of the ovaries, with successful results, and back in 1905 Marshall and Jolly found that oestrus could be induced in ovariectomized dogs by implanting ovaries from another dog into the peritoneum, or else by injection of ovarian extracts.

These and other experiments have firmly established the existence of sex hormones and have led to the eventual isolation of pure crystalline androgens and oestrogens. Administration of these compounds has the same effect on the sex organs and secondary sex characters as gonadal transplants or extracts, and in addition they have been shown to be capable of modifying the developing gonads of birds and lower vertebrates. It should be stressed, however, that whether an animal or a man develops ovaries or testes is determined genetically when the parental ovum and spermatozoon unite.

Oestrogens also play an important role in pregnancy, especially in promoting growth of the uterus.

Oestrogen production by the ovary is controlled by the gonadotrophic hormones of the pituitary.

THE OESTROGENS

Definition

An oestrogen may best be defined as any compound that produces oestrus, with the accompanying morphological changes

of the uterus and cellular content of the vaginal wall, when injected into ovariectomized rats or mice. This includes compounds such as oestradiol and oestrone that have been isolated from the ovary and are referred to as natural oestrogens.

The term ovarian or gonadal hormone, as applied to these substances, is not strictly correct because oestrone is also produced by the placenta and the adrenal glands. Similarly, there may be objection to the term male or female sex hormone since relatively enormous amounts of oestrogens are present in the urine of stallions, and horse testis contains the highest concentration of oestrogenic substances that has been found to occur in any tissue – about 300 times as much as in mares' ovaries. On the other hand, ovaries are known to be able to produce androgens. To complicate matters still further, oestrogens and androgens are to some extent bisexual in their action since they each produce certain changes in the sex accessories of both male and female, and sometimes they will act synergistically, i.e. enhance each other's activity. It may, therefore, not sound quite so surprising that recent experiments have conclusively demonstrated the direct conversion in the living organism of the most potent androgen, namely testosterone, into oestrone.

The natural oestrogens

Oestrone was the first naturally occurring oestrogen to be isolated in pure crystalline form by two subsequent Nobel Prize winners, working independently, Doisy in the U.S.A. and Butenandt in Germany. Both obtained the hormone from human pregnancy urine, a source from which Marrian, a year later, in 1930, isolated a closely related compound, oestriol. (See Fig. 1 for formulae of compounds being discussed.)

The most active oestrogen, oestradiol, was synthesized in the laboratory before it was obtained in 1935 from sows' ovaries. The yield was only about 10 mg. from 4 tons of slaughter house ovaries.

As already mentioned, these oestrogens have been found in a number of other tissues as well, though usually in smaller amounts; oestrone has also been obtained from palm kernel

extracts, and oestriol from female willow flower. Their function in plants remains a complete mystery.

Two other closely related oestrogens, produced by the equine species, are equilin and equilenin, a rich source of which is the urine of pregnant mares. Credit for their isolation goes to the French chemist Girard.

Chemically they are all steroids, i.e., they all have the perhydro-cyclopenteno-phenanthrene nucleus, modified to some extent, and they are closely related to cholesterol, the bile acids, progesterone, the androgens, and the adrenocortical hormones. The natural oestrogens differ from these other steroids by having one or more aromatic rings in their molecules.

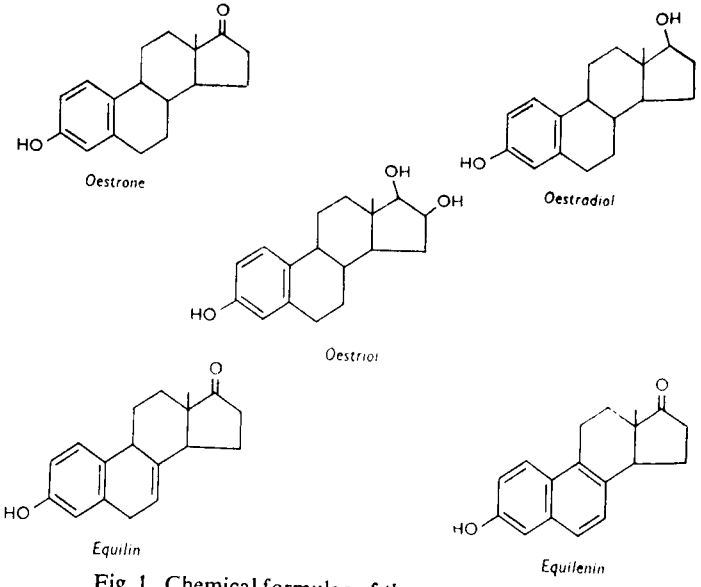
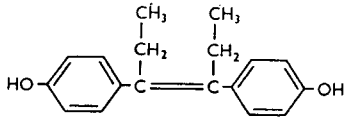


Fig. 1. Chemical formulae of the natural oestrogens

The synthetic oestrogens

The discovery that oestrogenic activity was exhibited by substances not occurring in the body, or belonging to the steroid

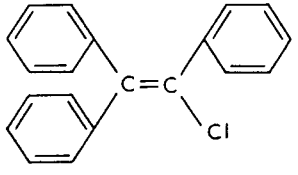
group, is due to Sir Charles Dodds, Professor of Biochemistry at the Middlesex Hospital Medical School, London, and his collaborators. Between 1933 and 1938 they synthesized and tested for activity a very large number of organic compounds, their work culminating in the synthesis of stilboestrol, and related series of highly active oestrogens. Sir Robert Robinson and his colleagues at Oxford were also involved in this research.



Stilboestrol

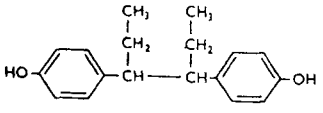
The great advantage of stilboestrol over the natural oestrogens lies in the ease of its manufacture, with resulting lower cost, and also in its high activity when given by mouth. It has been used clinically in the treatment of delayed puberty, the suppression of lactation on weaning, and the treatment of the menopause, as well as in cases of cancer of the prostate. Other uses include its administration in some countries to prisoners with criminal sexual drive, and the implantation of pellets of stilboestrol in poultry in order to produce physiological capons.

Other highly active synthetic oestrogens are also known: these include some that are chemically modified natural oestrogens, e.g., ethinyl oestradiol, doisynolic and marrianolic acid, named after Doisy and Marrian respectively, and some that are totally unrelated to natural oestrogens such as triphenyl chloroethylene. In this compound even the phenolic group, present in all oestrogens mentioned so far, is absent.

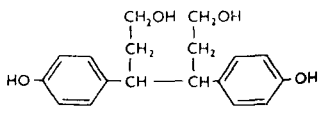


Triphenyl chloroethylene

Minute modifications in chemical structure can completely alter the biological activity of these substances. Thus dihydroxyhexoestrol is inactive, while hexoestrol is one of the most potent oestrogens known. Another interesting oestrogenic compound,

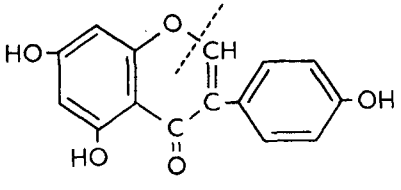


Hexoestrol
(active)



Dihydroxyhexoestrol
(inactive)

which occurs naturally in certain types of plants, is the isoflavone, genistein. Its presence in a strain of subterranean clover, extensively cultivated for sheep pasturage, was found to be responsible for reproductive disorders of ewes feeding on it. Examination of the chemical structure of genistein shows that on breaking the oxygen containing ring a stilboestrol-like compound would be formed. This suggests that in nature there may possibly be precursors to oestrogens that have so far been obtained only in the laboratory.



Genistein

It is certainly interesting to consider that an extremely complicated biological reaction, involving the whole of the female genital tract and all the secondary sexual characteristics, should be carried out by two groups of substances; first those produced in the body of a complex steroid type and secondly by relatively simple derivatives of hexane and other compounds. So far it has not been possible to correlate chemical structure with oestrogenic activity.

Before leaving this topic, it should be mentioned that since

the total synthesis of oestrone by Miescher and his colleagues in 1948, the term synthetic oestrogen could apply to both the naturally occurring hormones and stilboestrol-like compounds. The term artificial oestrogen has therefore been suggested for the latter group.

Oestrogen assay

In order to study the metabolism of oestrogens under physiological conditions, only minute amounts may be used, on account of their high biological activity, with the result that progress in the field has been limited by the sensitivity of the methods available for their detection and quantitative estimation. Some idea of the activity of the oestrogens may be obtained from the value of the international unit which is 0·0001 mg. (0·1 μg) of oestrone. This amount is sufficient to induce oestrus, and the accompanying morphological changes, in an ovariectomized mouse, and although a proportionally larger amount is needed to produce any effect in man, it is still a very small quantity.

A large fraction of any administered natural oestrogen is rapidly metabolized, so that to detect perhaps a few micrograms of a likely metabolite, even in urine which is free of interfering protein and lipoid material, is a task requiring considerable skill and experience.

Apart from the classical methods of chemical isolation, initially colorimetric analysis and bioassay were adopted. In the last of these, oestrogenic activity could be determined either by vaginal smear technique in which the changes in cellular content in the vaginal wall of ovariectomized rats or mice are observed (the cells are cornified when the animals are in oestrus), or by measuring increase in weight of uteri of immature mice.

More recently, fluorimetry, which is highly sensitive for the oestrogens though not very specific, chromatography, and isotopic tracer techniques have come into the foreground. The labelling of oestrogens with radioactive atoms especially shows great promise for solving many outstanding problems of their metabolism.

METABOLISM OF NATURAL OESTROGENS

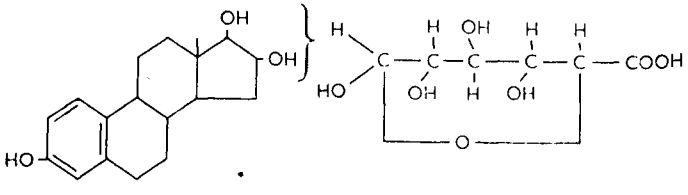
This section is concerned with the biogenesis and the fate of the oestrogens in the body. It is convenient to discuss these two topics separately.

It may be stated at the start, that the manner in which these hormones exert their profound physiological and psychological effects still remains a mystery.

The first problem is to consider what happens to oestrogens when they are administered, and which organs are mainly involved in their metabolism. Data have been obtained either clinically or by the use of experimental animals. The approach includes *in vivo* work, in which the oestrogen is administered to the whole intact animal, and *in vitro* experiments, in which an isolated organ is perfused or a slice or homogenate of a tissue incubated with hormone. Both methods have yielded interesting results, and the reactions that have been found to occur are: conjugation, interconversion, and destruction.

Conjugation

The existence of a relatively inactive form of oestrogen, which could be activated by boiling with acid, was first recognized by Marrian in 1933. Three years later he managed to isolate the sodium salt of oestriol glucuronide from human pregnancy urine. Oestriol must therefore undergo conjugation with glucur-



Oestriol glucuronide

onic acid, an oxidized form of glucose. This process, which involves the addition of another molecule, is well known in biochemistry. Amino acids are often involved, as for example in hippuric acid formation, but alcohols and phenols generally

form either glucuronic or sulphuric acid conjugates. It is, therefore, not surprising that oestrone sulphate has also been isolated – this time from pregnant mare's urine.

The activity of the conjugated oestrogens is considerably lower than that of the free compounds, so it has been suggested that conjugation may be a protective mechanism whereby excess hormone can be inactivated prior to excretion.

There is ample evidence that the liver plays an important role in the conjugation of many compounds including the oestrogens, and damage to this organ (induced experimentally by carbon tetrachloride or other toxic agent) leads to enhanced response to endogenous and injected oestrogenic hormones. In men liver cirrhosis may be associated with atrophy of the testes and the development of breasts.

β-glucuronidase, the biocatalyst or enzyme responsible for the formation as well as the hydrolysis of glucuronides, is abundant in liver. It has also been detected in many other tissues, and this suggests that the conjugation of oestrogens constitutes a factor, not only in their inactivation and excretion, but also in their transport and utilization in the body. Certainly the conjugates are more soluble than the free oestrogens.

Szegö and Roberts, who examined the blood of several species (including women), found that 75 per cent of the total oestrogens are associated with protein and that on dialysis (a process for separating small from large molecules) they pass through the membrane as acid hydrolysable conjugates. The remainder, which is not associated with protein, is also in the conjugated form, and other evidence indicates that the main oestrogen present in human blood is oestriol, though the other two have also been detected.

As far as the mechanism of glucuronide conjugation is concerned, it probably involves uridine diphosphate glucuronic acid, a compound that has recently been shown to act as glucuronic acid donor. It is essential for the formation of corticosteroid glucuronide for instance.

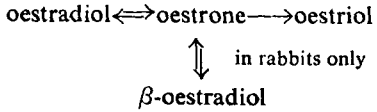
In oestrogen determinations it is essential to hydrolyse the conjugates by boiling with mineral acid or by the enzymatic action

of β-glucuronidase preparations. Acid hydrolysis usually results in some oestrogen loss while β-glucuronidase has no action on any sulphate or other conjugate that may be present.

Stilboestrol, incidentally, is also excreted as a glucuronide.

Interconversion

Many experiments have been carried out to establish the interconversion scheme of the natural oestrogens, usually by the administration of the pure crystalline hormone to animals or man followed by attempts to identify metabolic products in the urine, faeces, or bile. The results may be summarized by the following metabolic reactions:



Two points are noteworthy. One is that the formation of β-oestradiol, which is 40 times less active than the other form, has so far only been observed in rabbits. The other point is that oestriol does not appear to be converted into oestradiol or oestrone. Oestriol has been isolated only from the urine of pregnant women, and its presence in various animals after administration of oestradiol and oestrone has been inferred but not conclusively proved until recently.

There is evidence that oestrogens are excreted in the bile, pass into the intestines, from which they are absorbed, and are conveyed to the liver to be re-excreted in the bile. This enterohepatic circulation is similar to that of the bile acids.

Destruction

The greater part of any administered oestrogen is rapidly destroyed in the body. Often less than 5 per cent of the hormone can be accounted for in the urine, and maximum excretion of active material occurs within the first 12 hours.

In 1934 Zondek was able to recover from the whole carcasses, urine, and faeces of rats not more than traces of the activity of large doses of oestrone, even after only 3 hours. The yield could

be increased however by hydrolysis, showing that inactivation due to conjugation as well as destruction was taking place.

As in conjugation, it is the liver that plays a major role in the breakdown of these sex hormones. Incubation with liver slices *in vitro*, perfusion through liver, and, to a lesser extent, perfusion through kidney, all result in destruction of oestrone and oestradiol. Inactivation by liver is complete within a few hours, but the nature of the inactivation products is entirely obscure. Conjugation is ruled out because hydrolysis does not restore the activity of the incubated material, and in liver slices as opposed to the intact animal very little conjugation takes place.

In 1940 Heller found that cyanide inhibits the hepatic oestrogen destroying system, which suggested that an enzyme of the oxidase type was involved, but later workers could not confirm this. In a recent experiment carried out by the author in which radioactive carbon (^{14}C) labelled oestrone was incubated with rat liver slices, with and without added cyanide, there was however a marked effect. Thus after incubation for 4 hours in the absence of added cyanide, 31 per cent of the total radioactivity was present in the incubation medium – even after repeated extraction with ether in which the oestrogens and related steroids are soluble. Acid hydrolysis did not alter this value to any great extent, thus showing that something more than mere conjugation had taken place. In the presence of cyanide less than 5 per cent of the radioactivity was in the water soluble/ether insoluble fraction. The liver tissue was found to contain 4 per cent of the total radioactivity in the first experiment, but only 0·1 per cent when cyanide was added. In neither case was radioactive carbon dioxide formed so that it seems unlikely that complete oxidative destruction had taken place.

In vivo experiments by various research workers using radioactive oestrone have shown that after administration of hormone both bile and urine are radioactive, and that most of this consisted of water soluble/ether insoluble unconjugated compounds.

The kidneys, too, have been shown to have some oestrogen destroying activity.

Very little is known about the nature of the decomposition

products of natural oestrogens, but a possible mode of attack is by oxidation of the phenolic ring, similar to the way in which tyrosine and other monohydric phenols are dealt with by the enzyme tyrosinase. The reaction in these cases involves the introduction of a second hydroxyl group, ortho to the first, oxidation to an orthoquinone and then further destruction of this compound in an as yet unknown manner. Tyrosinase has certainly been shown to inactivate oestrone and oestradiol.

Although it is impossible to mention all the work carried out in the field, this summary gives a general idea of the present state of knowledge of the fate of oestrogens in the body. There is obviously a wealth of information still to be discovered.

Biogenesis of the oestrogens

Had this article been written four years ago, the section on biogenesis could have been summarized in the words of Lieberman, who in an excellent review on steroid metabolism stated: 'Virtually nothing is known concerning the mechanism of biosynthesis of the oestrogens.' The fact that more is known nowadays about the origin of these hormones is due to the successful chemical synthesis of isotope labelled steroids, and the work of American and Canadian biochemists, of whom that of Heard and his colleagues at McGill University has been outstanding.

By 1953, the main metabolic inter-relationship between the different groups of steroids and steroid hormones had been elucidated (see Figure 2), and it had been shown that cholesterol, which is abundant is all animal cells, could give rise to bile acids (e.g., cholic acid), adrenocortical hormones (e.g., corticosterone), progesterone, and the androgens (e.g., androsterone), although, of course, there are other routes in the body to each of these substances. Cholesterol, however, did not give rise to the oestrogens as demonstrated in the pregnant mare in a more recent experiment.

Furthermore it had been amply demonstrated that the whole of the complex molecule of cholesterol could be synthesized from acetate by the body. By chemical degradation of this sterol, biosynthesized from isotopically labelled acetate, it was even

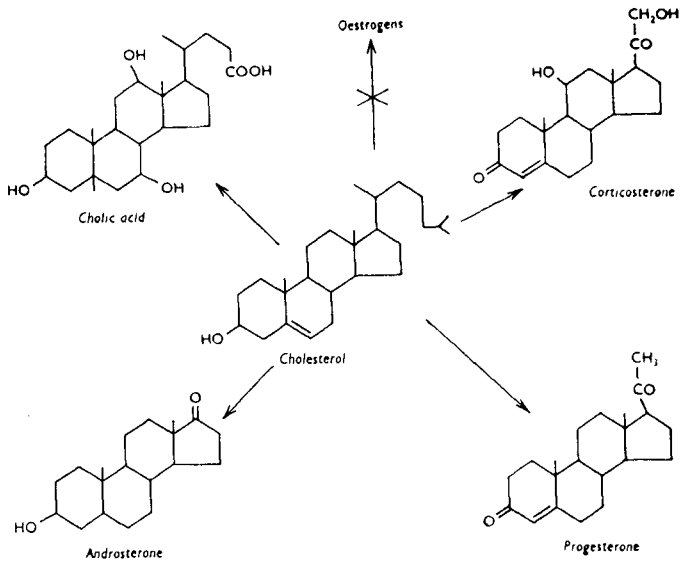


Fig. 2. Diagram showing the role of cholesterol in the formation of steroids in the body.

possible to show whether a given carbon atom had come from the methyl or carboxyl group of the acetate molecule.

Formation from acetate

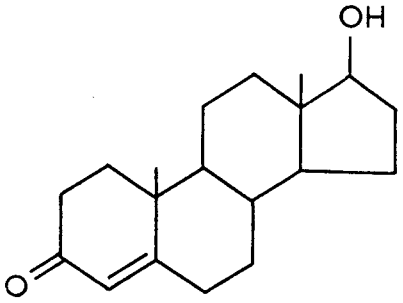
This work led Heard and his group to investigate whether acetate would give rise to oestrogens in the body. A pregnant mare was injected with radioactive (^{14}C) sodium acetate: oestrone, equilin, and equilenin were isolated from the urine. All three steroids were found to be radioactive, with the oestrone twice as active as the other two. This suggests that perhaps acetate is only utilized in the synthesis of the non-aromatic part of the oestrogen molecule and that oestrone does not give rise to equilin and equilenin. The latter point was confirmed by showing that after administration of ^{14}C labelled oestrone, the more

highly unsaturated equilin and equilenin contained no radioactivity.

The formation of oestrone and oestradiol from acetate has also been shown to occur *in vitro* in tissue slices and cell-free homogenates of dog ovaries, and in perfused human placenta.

Formation from testosterone

It was found that oestrone could also arise from testosterone, the most potent of the androgenic hormones. This surprising



Testosterone

though unequivocal reaction was shown to take place both *in vitro* in surviving human ovarian slices, and *in vivo*, in the pregnant mare. In the latter experiment, the isolated oestrone was radioactive but not the equilin or equilenin. The conversion was direct and not via acetate to which testosterone could conceivably be broken down.

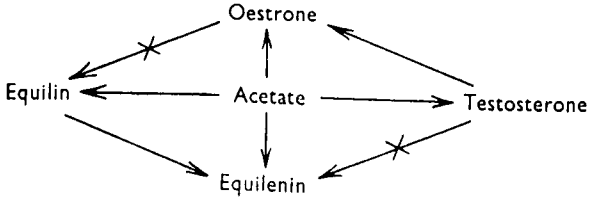


Fig. 3. Diagram showing the biogenesis of the oestrogens.

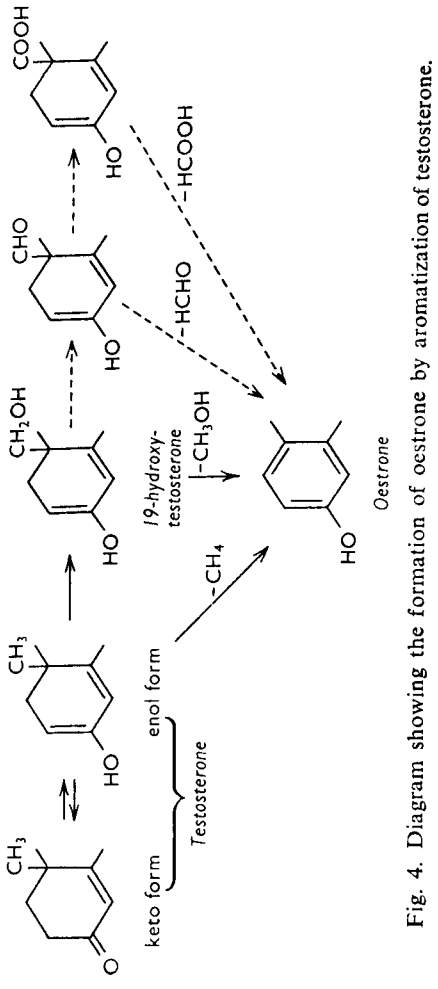


Fig. 4. Diagram showing the formation of oestrone by aromatization of testosterone.

Another reaction that has recently been demonstrated with placenta and other tissues is the conversion of androstenedione and 19-hydroxyandrostenedione, both closely related to testosterone, into oestrone. The yield was greater with the 19-hydroxy compound which strongly suggests that 19-hydroxylation is an intermediate step in the aromatization of the androgen to yield oestrone. In the following diagram, reactions established either *in vivo*, or with enzyme systems *in vitro*, are indicated by a solid arrow and hypothetical reactions by a broken arrow.

Cleavage of methane, methyl alcohol, or formic acid is not commonly observed in biological systems, but scission of formaldehyde is much more probable and, indeed, has been observed in the ouabain series of steroids.

* * *

In conclusion it may be said that work in the oestrogen field is still in a pioneering stage, and that important discoveries should be forthcoming in the next few years.

Some of the problems still unsolved include the determination of the nature of the degradation products of the oestrogens, other precursors in their biogenesis, and also the discovery of the way in which they exert their varied physiological effects. It may be that, as in the case of many of the vitamins, the characteristic action of these hormones is to be found in their participation, whether as coenzyme or substrate, in essential, specific, enzyme systems, but there is no evidence yet that steroids can function in this manner.

FURTHER READING

Dorfman, R. I., and Ungar, F.: *Metabolism of Steroid Hormones* (Minneapolis, Burgess Publication Company, 1953).

Heard, R. D. H., *et al.*: 'Biogenesis of the Sterols and Steroid Hormones', *Recent Progress in Hormone Research*, 1956 **12** 45.

'Hormones in Reproduction', *British Medical Bulletin*, 1955 **11** 83.

Lieberman, S., and Teich, S.: 'Recent Trends in the Biochemistry of the Steroid Hormones', *Pharmacological Reviews*, 1953 **5** 285.

Röbson, J. M.: *Recent Advances in Sex and Reproductive Physiology* (London, Churchill, 3rd edition, 1947).

DISINTEGRATION CRATERS

F. N. JOHNSON

IT may happen that when a body in motion makes contact with a solid, or semi-solid, surface, a crater is formed – a process that may, or may not, be accompanied by the destruction of the impinging body. The mechanism by which the crater is produced depends on the nature and energy of the missile, as well as on the composition of the surface undergoing the impact.

In *The Face of the Moon*, Dr Ralph B. Baldwin states[1] that experiments conducted by the United States Army Ordnance Department indicate that inert missiles possessing velocities between four and five miles per second, will explode spontaneously on making contact with a solid surface, and Dr H. C. Urey has pointed out[2] that, if the inert missiles contained volatiles to any great extent, explosions might possibly be induced at lower velocities than those indicated by Baldwin.

A similar effect, and one which lends itself more readily to laboratory study, is obtained when both missile and receiving-surface possess low mechanical strength. Such conditions may be achieved by using in each case a loose conglomeration of small particles – in other words, a powder. The crater-forming mechanism is then not so much an explosion as a 'pushing-aside' of the base powder, accompanied by the complete disintegration of the missile and of the base powder immediately below the point of contact.

Little has been attempted in the way of detailed study of such impact craters, although in the years immediately following the last war several experimenters touched on the subject in connexion with the suggested meteoritic origin of the lunar craters. The first of these workers was Ley.[3] His work was followed up by Crooke[4] in 1946, and in 1949 by Gale and Megson,[5] while the problem has been revised very recently by Gilbert Fielder,

the University of Manchester. The work described by the author was, however, of an independent nature. It was intended primarily as a preliminary study, and as a basis for further, more comprehensive, research on craters in general.

Quantitative Study; Technique

The prime object of the quantitative study of the powder craters was to correlate certain aspects of crater structure with each other, and with various characteristics of the productive missiles. The aspects of the craters that were important were diameter, depth, and volume. The first of these presented no undue difficulty from the point of view of measurement, nor did the second, but it was obvious that to measure crater volumes some new technique would have to be introduced. In the early experiments carried out by Fielder, cement powder was used, thus enabling the subsequent solidification of the crater, and greatly facilitating the volume measurement.

In all the experiments described here, however, the medium used was plaster of Paris – it being lighter, volume for volume, and generally superior in its disintegration qualities. After solidification, the crater could be filled with fine, dry sand from a measuring cylinder, and its volume found between limits. There was, of course, no justification for carrying out such measurements to any great accuracy. Measurements thus obtained were to be correlated with characteristics of the missiles; these latter included mass, energy, velocity, and momentum. A little thought will make it clear that, for a fixed mass of powder in the missile, energy, velocity, and momentum are all functions of the dropping height. All that was necessary, therefore, was to measure the dropping height for missiles of the same mass, and the sizes of the craters produced, and plot to obtain the relationships.

A word might be said here concerning the dropping of the missiles. Small, slightly conical cylinders, with open ends, were used. A card was held against the wider of the ends and the cylinder then filled, not packed, with the dry plaster of Paris powder. The usual order of mass was round about ten grams for a normal missile, although experiments were from time to

time conducted with missiles ranging in mass from five to twenty grams. On the removal of the card from the cylinder, the powder fell freely whilst retaining the cylinder's shape until contact with the base powder. Energies of only some two million ergs were needed in the missiles in order to produce very appreciable craters. In both explosion and powder crater formation, in addition to the formation of the main crater, secondary bodies are ejected which fall to produce smaller craters, or 'craterlets', at distances from the primary.

Other material, originating either from the missile or the base powder, is flung in streaks, or rays as they are commonly termed, having lengths varying between a few centimetres and one and a half metres. Such rays, though pointing towards the general region of the crater, are not necessarily radial to it.

In the course of the experiments several hundred impact craters were produced and, with the assistance of Mr L. Pointon, measurements taken and tabulated prior to the construction of graphs. Individual results in tabular form, being of immediate interest, were communicated[7] to the British Astronomical Association late in 1956, but are not reproduced here.

Quantitative Study; Results

The symbols employed are:

V = Crater volume.
D = Crater diameter.
d = Crater depth.
m = Mass of missile.
h = Dropping height of missile.
e = Energy of missile.
v = Impact velocity of missile.

A, B, C, E, F, are constants; *c.g.s.* system throughout.

In the analysis of the results obtained from the experiments, many interesting relationships appeared. It became possible to verify Fielder's result[8] that crater volume was directly proportional to the square-root of the missile's dropping height, provided that the mass of powder in the missile remained constant.

$$V = Ah^{\frac{1}{2}} \qquad i$$

From this result it may readily be deduced, by simple considerations, that two other relationships also hold;

$$V = Be^{\frac{1}{2}} \qquad ii$$

$$V = Cv. \qquad iii$$

In the above equations, constants of proportionality were calculated, but are of no practical importance, because, as will be appreciated, their values depend on the nature of the medium used – density, and degree of pulverization being only two of the many factors involved.

For experimental convenience, the expression $mh^{\frac{1}{2}}$ was used by the writer to represent missile momenta, and an important result was derived, of the form;

$$V = Emh^{\frac{1}{2}} \qquad iv$$

and from the graph (see Figure 1), it was further determined that, under the conditions of the experiment,

$$V = 0{\cdot}78mh^{\frac{1}{2}} + 0{\cdot}68. \qquad v$$

It must be emphasized that the relationships indicated in equations *i* to *v* hold, for the most part, over a limited range only. There is, for example, a limit to the size of crater that may be produced by a missile of fixed mass, no matter how great its velocity, so that at high velocities mass becomes critical, while at lower velocities it is the actual velocity that is the more important. Results *i* and *iii*, then, are valid for a fixed mass of missile powder over a certain range only, but *ii* is valid at all times, provided that the missiles become more massive as their velocities increase. This also holds for equations *iv* and *v*.

Relations involving crater depth and diameter were also established, but are of theoretical interest only, being less distinct in character than the more important volume considerations just quoted.

So far, however, emphasis has been laid on the relation between missile and crater characteristics, and little has been said about the connexions of crater measurements as such. In point

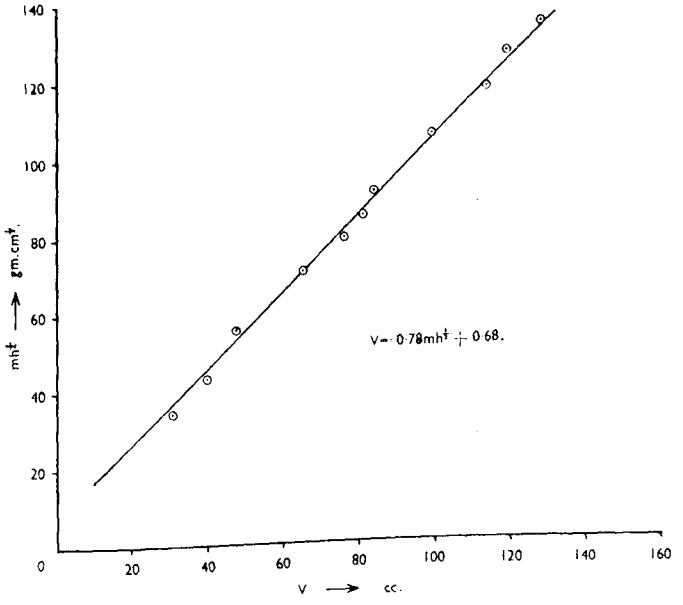


Fig. 1. Graph illustrating equation *iv*.

of fact, the experimental determination of a depth/diameter relation was of great interest, because Dr Baldwin had already published[9] an empirical relation between depth and diameter measurements taken from lunar craters, bomb craters, and shell and explosion pits. This he gave as:

$$\text{Log } D = 0{\cdot}1083 \log^2 d + 0{\cdot}6917 \log d + 0{\cdot}75 \qquad vi$$

but, as he later pointed out,[10] the points defining the lower portion of his curve were not completely satisfactory, and, following further work, he has arrived at the revised equation;

$$\text{Log } D = 0{\cdot}0462 \log^2 d + d + 0{\cdot}48. \qquad vii$$

The powder impact craters yielded an equation of the form;

$$D = Fd \qquad viii$$

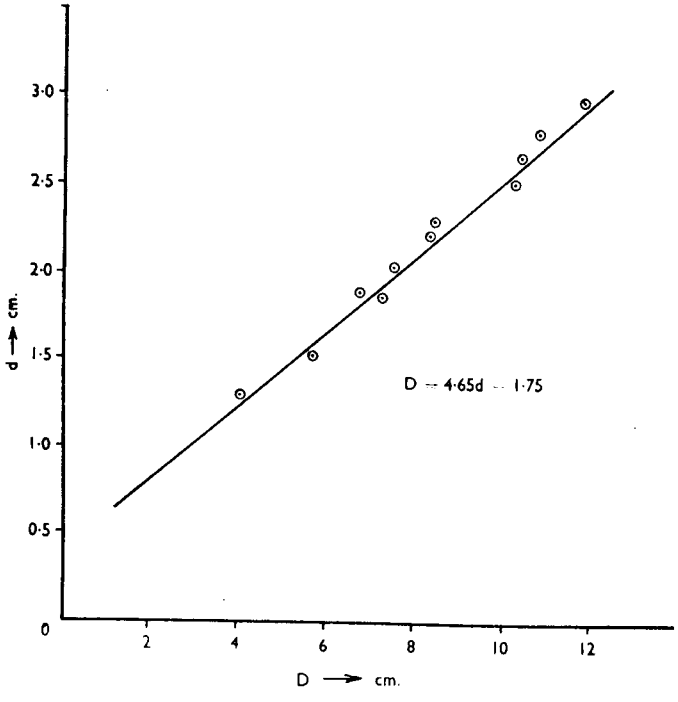


Fig. 2. Graph of equation *viii*.

and from the graph (see Figure 2) it was determined that;

$$D = 4{\cdot}65d - 1{\cdot}75. \qquad ix$$

The exact inter-connexion of this result and Baldwin's revised equation will be discussed later. Finally, it was found to be generally true that very approximately

$$V = 0{\cdot}32\, D^2 d \qquad x$$

but this result was never used to calculate crater volumes instead of measuring them experimentally.

Qualitative Study

Those who are acquainted with astronomy, or who have seen photographs or drawings of parts of the Moon's surface,* will be familiar with the general form of a lunar crater: a very shallow bowl bounded by a relatively low wall, known as the 'ring-wall', and at times possessing a small central elevation called the 'central-mountain'.

All these features may be readily observed in photographs (see Insets 7 and 8). These photographs were taken from typical experimental impact craters and show considerable resemblance to some smaller undamaged lunar structures.

In all the impact experiments, the formation of central mountains could be avoided by using a sufficient depth of base powder and missiles of high velocity; such conditions favoured the production of craterial depressions tending towards a hemispherical form. Central mountains could be induced at any time by making the base powder somewhat denser than the missile.

The diameters of the craters ranged from eight to twelve centimetres, while the depth was always near to two and a half centimetres. In no case did the height of the central mountain exceed that of the exterior base powder.

The superficial resemblance of powder craters to those observed on the Moon really does not of necessity tell us much about the internal structure of impact craters. Following a technique devised by Fielder, stratified base powder was used. Each stratum was about five millimetres in depth and coloured by mixing with the plaster of Paris some equally dense coloured powder, as for example, litharge. It will be realized that in this, care had to be exercised so as not to have one stratum denser than another. The normal experiment was then carried out with a white missile, the resulting crater solidified and then sectioned vertically. Figure 3 shows a drawing made from one such experiment. It must be pointed out, however, that in this drawing the arching beneath the central mountain is exaggerated.

* See, for example, the photographs that accompany D. W. Allan's 'History of the Moon' in *Science News 45*.

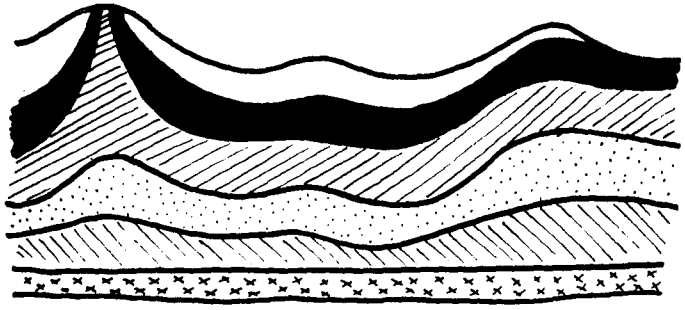


Fig. 3. Section through a stratified crater. (Vertical scale slightly exaggerated.)

An interesting point, and one worthy of note, is the disappearance of some of the lower strata through the ring-wall. The material forced up through the ring-wall in this way went to produce very short rays – about twenty centimetres long – which were tangential to the crater perimeter. A further point is the existence of a slight arching of the strata directly beneath the central mountain – exaggerated in the drawing but none the less evident in the original stratification. Dr Gerard P. Kuiper has stated[11] that a surface sustaining an impact will not rebound in such a way as to form a central mountain. Experiment, on the other hand, indicates this to be hardly the case – although it must be admitted that a great deal of the material constituting the central mountain must be attributed to the missile.

As might be expected, craters were lined with material from the missile, but there occurred several instances when the ring-wall seemed surprisingly irregular, and material from the lower strata was scattered over the interior of the depression.

Ray Systems and the Nature of the Impact

As has already been said, the material that is ejected during the crater-forming process is roughly divisible into two categories. The first embraces all the material that is ejected at large angles to the horizontal and which falls to produce small,

secondary craters in the base powder. About three or four millimetres in diameter, the craterlets were formed in chains of two kinds. The first kind was in the form of a straight line approximately radial to the parent crater, while the second was shaped like an arc, the focus of which lay inside, or near to, the primary formation. Figure 4 illustrates this point, although it may be

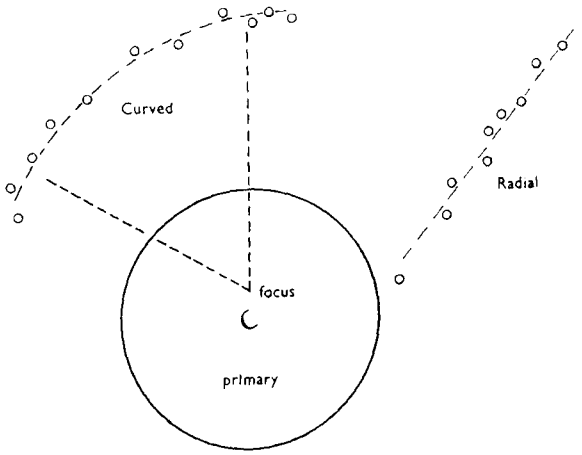


Fig. 4. Diagram of radial and arched chains of craterlets.

added that in no experiment were the chains quite so clearly delineated as in the figure. The individual craterlets formed by secondary impacts in the powder never encroached upon each other, a state of affairs which is predominant in most of the lunar craterlet chains. (See Figure 5.)

The second category of ejected matter goes to form the wonderful ray systems that are produced by even the least energetic impacts. Some three or four centimetres in width, the largest of the powder streaks belonging to each crater was often no less than one and a half metres long, and several exceeded the two-metre mark during the course of the experiments. Such lengths would seem to indicate projective angles of about forty-five degrees for the material at the tip. It has already been stated that

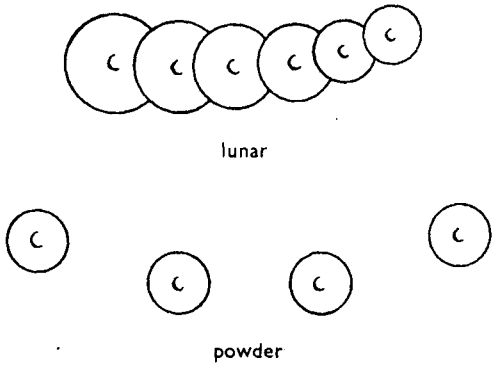


Fig. 5. Comparison of powder and lunar craterlets.

the smaller, tangential rays are the product of material from the lower regions of the ring-wall.

Powder rays, of course, have their parallel in the light coloured rays emanating from lunar craters – Copernicus illustrating the effect at its most beautiful, and those from Proclus exhibiting an odd 'split' effect, as though the passage of the material on its way out from the centre of the crater had been interrupted by some form of barrier.

Explosives and Astronomy

The powder impacts that have been dealt with in this paper certainly approximate in many respects to explosions, but not in every aspect is the mechanism *strictly* comparable to the explosion. This must be borne in mind when one attempts to correlate this form of experimental crater production with the empirical equations developed by Baldwin. The points that define the upper portions of Baldwin's curves were derived from lunar craters. If such structures were indeed formed as the result of an explosive impact by some form of meteoritic body, the explosion would have been of a uniform character, having its centre some distance below the initial point of impact. The same considerations apply to bomb explosions, but not, however, to those produced by artillery shells which, as a rule, explode immediately upon making contact with the ground. If such conditions are prevalent, the craters due to shell explosions will be correspondingly less deep for unit diameter, and it was this that caused Baldwin to replace the lower points on his original curve, which had been drawn from data of shell pits, with results obtained from controlled explosion experiments.

When measurements of the powder impact craters were plotted on an extension of Baldwin's revised curve, they yielded a curve with an identical slope, but displaced slightly upwards. Baldwin comments that this 'indicates that they are shallower relative to the diameter than are the equivalent diameter craters from sub-surface bursts. This suggests that the pellets shattered on impact and that what were actually observed were craters more like those from surface than from sub-surface bursts.'[12]

Insets 5 and 6 are of interest in that they show what is happening during the actual impact. Experimental conditions were arranged so that the missile, forcing crossed wires into contact, triggered off the flash of a camera in the darkened room. The camera shutter was open all the time.

An examination of these photographs would seem to indicate that the centre of the explosion is a small distance *above* the surface of the base powder, but whether this impression is illusory is a matter for further study. A slow-motion cine film of the complete process is obviously desirable. On the two photographs, the vertical stream of attenuated powder marks the path of the falling missile.

For many years there has been constant discussion among astronomers concerning the history of the lunar surface features, with particular reference to the craters.* There are two main points of view, the first and older of the two being that the craters are the result of one of many alternative types of magmatic action – the so-called volcanic or plutonic hypothesis,[13] while the second, which is based on the conception of random meteoritic impacts, has been extensively revised by Baldwin.

* See D. W. Allan *loc. cit.*

There are, of course, theories[14] which attempt to reconcile the impact and magmatic hypotheses, but in any case, on the basis of all available evidence, most of which is purely observational, there is no degree of certainty. The only reasonable manner in which to approach the problem is to attempt a comparison between what we already know to hold for terrestrial structures, both natural – like the Hawaiian volcanic craters – and, like the powder craters, artificial, with what we are able to observe on the Moon. No one piece of evidence, however, may be regarded as conclusive, but, by an intensive study of all facets of the problem it may be further resolved.

The powder missiles have been compared[15] with the low-velocity meteoritic bodies required by Urey's theory,[16] and, both qualitatively and quantitatively, the results obtained from the craters they produce compare favourably with those taken from lunar craters. But again, it cannot be too strongly emphasized that the relations quoted in this paper, in so far as they have been obtained from plaster craters, are applicable between restricted limits, and that more detailed, comprehensive results remain to be found.

In any event, whether seeking to verify this theory or that, or merely from an academic interest, the further study of crater-forms will be rewarded.

REFERENCES

1 Baldwin, R. B.: *The Face of the Moon* (University of Chicago Press, 1948) p. 62.
2 Urey, H. C., private conversation, University of Manchester, Oct., 1956.
3 Ley, W.: 'Experimental Moon Craters', *Popular Astronomy*, 1944 **52** 278.
4 Crooke, R. D.: 'More Experimental Moon Craters', *Popular Astronomy*, 1946 **53** 465.
5 Gale, S. I., and Megson, S. H.: 'Lunar Craters in the Laboratory', *Sky and Telescope*, 1949.
6 Fielder, G.: private communication, University of Manchester, 16 Jan. 1956.
7 Johnson, F. N.: 'Craters formed in Plaster of Paris Powder: A Note on Some Experimental Results', *J. British Astronomical Association*, 1956 **66** 319–22.
8 Fielder, G.: private communication, University of Manchester, 26 Feb. 1957.
9 Baldwin, R. B.: *op. cit.*, p. 131.
10 Baldwin, R. B.: private communication, Grand Rapids, U.S.A., 29 Apr. 1957.
11 Kuiper, G. P.: 'On the Origin of the Lunar Surface Features', *Proceedings National Academy of Sciences of U.S.A.*, 1954 **40** 1107.
12 Baldwin, R.B.: *ibidem.*
13 Moore, P.: 'Formación de los Crateres Lunares', *Urania*, 1956, No. 242.
14 Kuiper, G. P.: *loc. cit.*
15 Urey, H. C.: private communication, University of Chicago, 5 July 1956.
16 Urey, H. C.: *The Planets: Their Origin and Development* (Oxford University Press, 1952).

RESEARCH REPORT FROM FRANCE

A. W. HASLETT

A PARTY of scientific journalists was lately given an opportunity to visit a number of research and scientific establishments in France as guests of the French Government. Although disturbed and reduced in numbers by the launching of Sputnik I, the party completed its round of some ten visits, no one of which was a waste of time. Most of the establishments seen were governmental – notably some of those maintained by the Centre National de la Recherche Scientifique, the Commissariat à l'Énergie Atomique, and Électricité de France – and there was a bias in favour of new equipment and laboratories; we were visitors to be shown the best. It was also intended that we should enjoy ourselves, and that too was arranged.

Although we saw little of universities, it seemed that in France the most productive course for the head of a department in need of equipment is to persuade the Centre National to put up a new laboratory on his doorstep – and to direct both that and his own department.* An internationally known example is Professor L. Néel, of the University of Grenoble, who directs the well-equipped Laboratoire d'Électrostatique et de Physique du Métal of the Centre National and will be responsible also for the Centre d'Études Nucléaires, now under construction. Not every-

*Universities are administered directly by the Ministère de l'Éducation Nationale. Apart from equipment, well-informed guesses suggest that French University staffs are able to spend less of their time on research (one-third) than in Britain (one-half); the guesses (1957 in each case) are those made respectively by the Conseil Supérieur de la Recherche Scientifique et du Progrès Technique (France) and the Advisory Council on Scientific Policy (Britain).

one has contributed, as he has, to the opening up of a new area of his subject (the magnetic properties of materials) or can be rewarded, as he has been, by being given liquid helium in piped supply. There must be many departments that could do with more straightforward support, not necessarily on a big scale. That said, I shall concentrate on things seen and discussed.

SOLAR FURNACE

The Laboratoire de Recherches sur l'Énergie Solaire was established in 1950 in the fortress of Mont Louis, which was built by Vauban in the late seventeenth century to guard the only easy route through the Pyrénées. Its main interest is the use of solar heat for research on heat-resistant materials, such as are used in the lining of furnaces, and the biggest optical mirror in the world is being built for work on these lines. It will have an area of about 1,200 square ft, and will be a fixed installation. Subsidiary collecting mirrors, so mounted as to follow the movement of the sun, will give an effective collecting area of more than 20,000 square feet.

The advantages claimed for the solar furnace are freedom from chemical contamination and – associated with this – localization of the region of high temperatures. This may be as high as 3,500° C, although 3,000° C is looked on as a more realistic working maximum, and most present work at Mont Louis is in the range from 2,400° to 2,600° C. On a smaller scale, the idea has been taken up lately by the Stanford Institute, California, with acknowledgement to Mont Louis, and the point made that the method is 'particularly useful in the preparation of "super-pure" materials, such as single silicon crystals for research on transistors'. In this use, requirements in the way of temperature are comparatively modest; but, where even 1 part in 100 million may be an unacceptably high level for some impurities, avoidance of contamination is supremely important. Another use, mentioned at Mont Louis, is in the alloying of high-melting-point metals.

The main gain from using bigger mirrors is to increase the quantities of materials that can be treated. Temperatures of up

to 3,000° C were obtained experimentally by the director, M. Felix Trombe, with a 6-ft mirror at the Meudon Observatory near Paris before the Laboratory at Fort Mont Louis was established. With the big mirror system, it will be possible to treat two or three tons of material a day, and the intention is to do work on contract for industry, as is being done already to some extent. In the interval facilities are being increased by building four extra mirror systems.

The paraboloidal mirrors used at Mont Louis are built up from many small squares, some six inches or so across. The shape of the small mirrors can be distorted at will by an arrangement of screws and bars, and each is adjusted separately by hand to the focal length required for its eventual position in the composite mirror. This is a tiresomely repetitive operation, but is clearly a cheap way of building up very large mirrors, in which figuring to telescope standards is unnecessary.

The paraboloidal mirror of the big system now under construction will be mounted in the centre of a valley near Mont Louis. Along either side of the valley there will be rows of plane collecting mirrors, each with an area of 385 square feet. It is planned to build 60 of these mirrors – giving the effective collecting area of more than 20,000 square feet already mentioned. Purely as a spectacle, this silent army of mirrors, slowly following the sun, will, in its Pyrénéen setting, be a sight not readily to be forgotten.

THE PHYTOTRON

At Gif-sur-Yvette, near Paris, we saw a new laboratory for research on the effects of climate on plants, which will be the biggest and most elaborately equipped in any country. It has been built by the Centre National de la Recherche Scientifique at the instigation of Professor Pierre Chouard, who holds the Chair of Botany in the Sorbonne and will be also its director. He hopes that much of its work will be in association with visiting scientists – botanists, plant physiologists, geneticists, and biochemists – from other laboratories and countries. Since the laboratory will have cost the equivalent of £2,000,000 to build

and fit out, and will be expensive to run, there seems to be a good case for making it international in organization, as well as in fact, if that can be arranged. It will be working in part during 1958, and fully the year after. On the same site there are laboratories of genetics and hydrobiology.

The object of the new laboratory is to reproduce on the grand scale all possible variations of climate that affect the growth and development of plants, including combinations of conditions that are not found in nature. The first such laboratory was built in 1949 for Professor F. Went at Pasadena, California, and was named a 'phytotron' by Professor R. A. Millikan, from analogy with the cyclotron, which had marked the beginning of a new era in physics. Many university departments and research institutions have now smaller climatic chambers, in which much useful work can be done. The Gif-sur-Yvette phytotron will be unique in the extent of control which it is planned to provide, and in the number and size of its different chambers.

Thirteen of these chambers, each some 12 ft by 26 ft, will be phytotrons proper, with entirely artificial illumination, and temperature, humidity, and other factors all fully controlled. Air will be replaced at a rate ten times as great as in normal air-conditioning, and will be chemically controlled. There will be provision for the use of radio-isotopes in tracer experiments, and for special forms of illumination, giving irradiation within selected bands of the spectrum. In some, but not all of the chambers, artificial rainfall will be provided. The idea is to complete only a few of the chambers to begin with, and in fitting out the rest to be guided by experience with those first operated. It may also be convenient to include smaller compartments or cabins in some of the chambers, so that a greater variety of conditions can be studied at once. Associated with the phytotron, there will be eight chambers illuminated by the sun but with arrangements for the control of temperature; and, below these, there will be the same number of dark chambers. Finally, there will be biochemical and other laboratories.

For many experiments, it will be unnecessary to make use of the full range of control provided. Examples are research on the

relation between day-length, temperature, and flowering in plants. Studies on the effects of growth substances, on the other hand, are a case in which the most precise specification of conditions should help greatly in working out the relations between growth-promoting substances and inhibitors, and of both with natural plant hormones. Similarly, in research on plant nutrition and the balance of trace elements necessary for different species, a great saving in time should be possible. In pure botany, there is a wide field for research into the conditions that limit the range of different species.

As a more immediately practical type of research, it is envisaged that the phytotron could be used to determine the climatic limitations of new varieties of crop plants, and also their optimal ranges. In the usual procedure of plant-breeding, it is often desired to combine some desirable quality, such as high crop yield, with another such as resistance to drought or suitability for a short growing season. In such a case, the phytotron could be used to make a preliminary selection of strains meeting the climatic condition, which would then become raw material for plant breeding in some overseas research station. With such wide facilities, it should be possible to combine the more practical uses of the phytotron with its availability, preferably on an international basis, for use in more fundamental kinds of research.

ATOMIC ENERGY

We visited both the centre at Marcoule, near Avignon, for the production of plutonium, and Saclay, near Paris, which is an education centre as well as one for research. It is at Marcoule that the Commissariat à l'Énergie Atomique is building its first two reactors in which carbon dioxide under pressure will be used to extract heat from the reactors. They differ from the Calder Hall type, and from later reactors developed from it, in the fact that pre-stressed concrete, made gas-tight on the inside by steel, replaces the pressure-resistant steel shells characteristic of British designs. The effect is to relieve the pressure-resistant shell of the further function of supporting the weight of the

reactor which in British designs it is required to do. The reason for this is the greater expansion of steel, when heated, than of concrete. Because of this difference in expansion, it was thought impracticable, when designing Calder Hall, to carry the weight of the reactor directly on the concrete raft that underlies it; to do so would have implied treating the pressure-resistant shell as essentially a cover which could be bonded into the concrete beneath, and the join made gas-tight. The Marcoule reactors, on the other hand, are designed on the basis that their pre-stressed concrete shells – containing steel as well as concrete – will withstand the changes in temperature to which they will be exposed, and that the joins between shell and base will not suffer from movement. The latter is the main point on which fingers will be crossed. At the least, it is an experiment that deserved to be done; at the best, it might be the means by which gas-cooled reactors of higher power output could be built at less cost.

It was therefore of interest to a British visitor to discover not only that Électricité de France was going its own way – which substantially is the British way – in designing its first stations primarily for electricity generation, but also the reasons for this divergence. One of them is that France, with greater hydroelectric resources, has not the need that has been felt in Britain to arrive quickly at an accepted type of reactor. Another is that it is planned to rely mainly on home-produced uranium; and that not until some time in the 1960s will output of uranium reach the point at which more than a development programme can be supported. There is time, therefore, to experiment and, whatever type had been built by the Commissariat à l'Énergie Atomique, Électricité de France – although advised by the Commissariat on design – would have developed a different type.

With this background, it is no longer surprising to find that for their first nuclear power station (EDF 1, under construction near Chinon on the Loire) Électricité de France will use a cylindrical steel shell; for their second (EDF 2) probably a spherical steel shell; and that for their third they may try a variant of the Marcoule design using pre-stressed concrete. Also of interest is the thickness of the welded steel plate that they plan to use.

Whereas so far British designers have been content with pressure vessels of 3-inch thickness, Électricité de France is proceeding at once to just over 4·2 inches, and for EDF 2 is prepared to consider 5 inches. On this basis, 'improved Calders' with outputs as high as 400 MW per reactor – compared with 150 MW in the first stations, and 250 MW at Hinkley Point – look a reasonable prospect. Finally, there is a point of safety on which France and Britain have parted company. The French have not liked the idea of passing carbon dioxide from the reactor through heat exchangers in which the steam pressure in the boiler tubes was higher than the gas pressure outside the tubes. The potential consequence of a leak in that situation might be that water might be carried back in the gas circuit and react with the magnesium alloy used to protect the fuel elements. British designers have relied on a succession of safeguards: the use of a magnesium alloy as resistant to corrosion as possible; the taking of extreme precautions to ensure the soundness of the heat exchanger tubing; continuous monitoring of the returning gas stream from each heat exchanger for the presence of moisture; and provision for the prompt cutting out of any heat exchanger should a leak in it be detected. The French have begun, certainly, by rejecting the British corrosion-resistant alloy as too difficult to fabricate. In EDF 2 they may change their mind. They are not at present entirely convinced.

Aside from her own programme, France is building up a position in western Europe in education for atomic energy. Saclay alone is a centre for research and design in atomic energy which provides also a year's full-time course for engineers working in industry, and wishing to specialize in atomic energy. For comparison, the longest course at Harwell lasts 16 weeks – although there are more opportunities in Britain to learn in industry. Other courses at Saclay – which in its training capacity is the Institut National des Sciences et Techniques Nucléaires – include one or two years in radiobiology for research workers, and several more that can be taken in conjunction with doctorates at the Sorbonne. In distinction from Saclay, the Centre d'Études Nucléaires at Grenoble will provide a year's training for engineers not yet experienced in industry. All these courses are open to nationals of other countries, subject to the possession of equivalent qualifications.

HIGH-ENERGY ACCELERATORS

As well as three research reactors, the most recent of which is comparable with the newer reactors at Harwell, Saclay has five machines of different kinds for the acceleration of nuclear particles to high energies, and a sixth which, when completed in 1958, will be the most powerful machine of its kind in western Europe. This is a 3,000 MeV proton synchrotron. Pending completion of the big CERN machine at Meyrin near Geneva – due in 1960 – it will be surpassed only in the United States and Russia.

The equipment at Grenoble will be comparatively modest, but includes a new type of particle accelerator – or, more accurately, a high-speed electrostatic generator of cylindrical shape, devised some years ago by Professor N. J. Felici, of the University of Grenoble, and now developed to enter the range of voltages useful to physicists.* Like the Van de Graaff generator, it offers accurate control of voltage, and is therefore suited to measurements of the lowest energies at which particular nuclear reactions will just proceed. But it gives bigger currents, and is therefore suited also for use – at second hand – as a source of high-energy neutrons. Up to 1954, the highest potential for which a machine of Professor Felici's type had been built was 200 kV. A 600-kV machine, giving a current of 5 milliamperes and slightly less than a foot in diameter, is working in the Laboratoire d'Électrostatique. Harwell is to have a bigger one as

*The essential components are an insulating core, with ionizing electrodes; an insulating rotor, concentric with the core about which it rotates; hydrogen gas under pressure as the dielectric fluid; and a slightly conductive stator, again cylindrical, with contact metallic segments. English-language references are: Felici, N. J. *Direct Current*, 1953, **1**, 122; *Research*, 1954, **7**, 265. The makers are the Société Anonyme de Machines Électrostatiques, a specially formed company with which Neyrpic (see following paragraph) is associated.

a source of high-energy neutrons, and a 1·4 MV installation is planned for the Centre d'Études Nucléaires.

LABORATOIRE DAUPHINOIS D'HYDRAULIQUE

To a visitor familiar with British research, a hydraulics research station working with models on the problems of rivers, estuaries, and sea harbours suggests primarily the one built with Government funds at Wallingford for the Department of Scientific and Industrial Research, and needing more funds for expansion. In Grenoble, it means a laboratory that has become the largest of its kind in Europe, and in volume of work probably the biggest in any country, but which has grown up as an off-shoot from a company, Neyrpic, whose primary concern is the making of turbines. It has been a self-expanding development. In 1906 the parent company built a small experimental station for tests on turbines. In 1923 it was asked to investigate the arrangements for the supply of water to some turbines that it was building for a hydroelectric installation, and in 1927 it did a similar investigation on an already completed plant. In 1928 it imported a young engineer, Pierre Danel, who built up a reputation for the laboratory, at first in its original field of work. In 1935 it added harbours and breakwaters to its repertoire. Later, it took to the control of irrigation canals and land use. In 1955 it became a separate organization and under the name Société Grenobloise d'Études et d'Applications Hydrauliques, commonly shortened to Sogréah. Its staff of qualified engineers numbers 150–roughly the same as that of the parent company, and its laboratories extend over some 15 very fully used acres. Eight-tenths of its work is done for countries other than France. It describes itself as acting as consultants to consultants. Directly or otherwise, it has contributed to work on the port of Heligoland for Germany, Guam for the United States, the port of Wellington for New Zealand, and Valetta Harbour, Malta, for the British Admiralty. At the present time, it has a finger in the Kariba Dam in Rhodesia, Zonguldak Harbour in Turkey, and the siting of a causeway in Brazil – three reasonably well-scattered examples.

It has also done general research on such subjects as ocean swell and its interaction with sea-beds and currents, the flow of stratified fluids, and coastal erosion. It has looked at the comparative merits of unshaped natural rock and rectangular blocks in the construction of breakwaters, and concluded that concrete blocks made with cylindrical legs sticking out symmetrically in four directions would use less concrete, and would interlock readily and settle down stably under the action of gravity and wave action. It has looked at the operation of drilling for oil and decided – as have also the Russians – that, as depths were increased, it would be an advantage in some circumstances to use turbines, driven by mud under pressure, at the bottom of the drill hole. All this, it may be repeated, has been without cost to the French taxpayer, and to the benefit of the parent company, 60 per cent of whose business is for export. Remembering that the use of hydraulic models originated in the work of Osborne Reynolds in Manchester, it may reasonably be asked: why has no British firm done the same thing?

MOVING AROUND

Other visits during our tour can be rendered only as disconnected notes. With several days instead of an afternoon in Professor Néel's laboratory at Grenoble, justice might have been done to his programme of research on magnetic materials and properties, approached for their interest to physics. Experiments under his direction led recently to the discovery of a new class of rare-earth ferrites, on which work is proceeding; and there is curiosity about what really happens during repeated cycles of magnetization and demagnetization (as opposed to the simple loop diagrams that look so convincing in text-books).

The Palais de la Découverte in Paris brought a reminder that there is in Britain no national museum whose purpose is to explain science and to aid in its teaching, as opposed to the preservation of historic apparatus and the illustration of history. The Station d'Essais de Fontenay, near Paris, of Électricité de France claims the distinction that it is the only high-voltage testing station in the world that is connected directly to a high-voltage national grid; it is now testing cables for the 100-kV

direct-current link between Boulogne and Dungeness, due to come into service in the winter of 1960–61. A near neighbour, also at Fontenay – the Laboratoire Central des Industries Électriques – is unusual in being an industry-owned laboratory responsible for the maintenance of national electrical standards. Finally, although we saw nothing of the research done by French Railways, we discovered (in the wake of Sir Brian Robertson) how much has still to be learnt by British Railways about mechanized marshalling yards; saw the entry into Paris of the Nord–Paris electrification which, by the middle of 1958, will be directly connected with some 900 miles of electrified main lines, completed since 1953 on a newly developed system; and were told almost incidentally that, apart from a supposed, probably real, prejudice on the part of passengers, the 35 miles of line from Dôle to Vallorbe could be operated at any time with driverless, remotely controlled trains.

Acknowledgements are too numerous to be made in detail here. In the planning and organization of our tour as a whole, at the ten establishments that we visited, and in the meeting of individual requests when made, we could not have received more consideration from our hosts.

SOME BOOKS RECEIVED

THE ORIGIN OF LIFE ON THE EARTH by A. I. Oparin, translated from the Russian by Ann Synge (Edinburgh, Oliver and Boyd, 1957), pp. 495, 35s.

Although the problem of the origin of life has fascinated thinkers since the beginning of time, it is strange to ponder on the fact that only a hundred years ago the metaphysical world was startled and reft by the publication of a theory that dared to suggest that man was not the work of a divine agency but, as far as our planet is concerned, just the end product in a process that had been going on for thousands of millions of years – or that Pasteur's researches had not yet scotched the idea of spontaneous generation.

In this, the centenary year of the Darwin Russell-Wallace communication to the Linnean Society, the fact of evolution will, in the main, be accepted, although some, contemplating the fearful complexity of Nature, will feel that there is still much to be explained about the mechanism that has brought it about. Be that as it may, man's probing has gone one stage deeper and many branches of experimental science are converging to produce an explanation, not of evolution and its mechanism, but of the first step, the emergence of the first living things on our planet.

Professor A. I. Oparin, who is a member of the Academy of Sciences of the U.S.S.R., first published a book on the origin of life in 1936: an English translation appeared in 1938. It created a great deal of interest and in the intervening years much new scientific information has come to light which can be brought to bear on this, the profoundest and grandest of problems. Now, with the title 'The Origin of Life on the Earth', we have a completely rewritten thesis translated from the Russian by Mrs Ann Synge. The new factual material is used to create a 'more definite picture of the successive stages in the development of matter on the way to the origin of life'. Its publication was planned to coincide with the First International Symposium in the Origin of Life which was organized by the Academy of Sciences of the Soviet Union under the auspices of the International Union of Biochemistry.

Throughout the ages man has *speculated* on his origin and the emergence of the life from which he evolved. But today his ability to manipulate matter and understand its various manifestations is

becoming so certain that in all scientific modesty Professor Oparin can say: 'We must check our knowledge by experiment. We must reproduce experimentally the separate stages in the historical development of matter and finally create life again, synthetically, not by the long and devious route based on a thorough understanding of those forms of organization which we find already in a finished state in existing living things.'

It is the history of this understanding that he sets out in this book whose fascination matches the fascination of the subject.

A HISTORY OF TECHNOLOGY edited by Charles Singer *et al.*: Volume III, From the Renaissance to the Industrial Revolution (Oxford, at the Clarendon Press, 1957), pp. 766, 168s.

It is a pleasure to be able to notice the appearance of the third volume of the great history of technology which is being produced under the editorship of Dr Charles Singer, and sponsored by Imperial Chemical Industries, Ltd. As the editors point out in their preface, their time-scale is closing in rapidly. The period covered by volume I was measured in geological time, volume II some two thousand years – which brought the history up to the end of the Middle Ages – while the present volume covers a mere two hundred. And even to cover these two centuries it has been necessary, as the editors point out, to be highly selective.

This is a threshold volume: science, in the minds of Bacon and Galileo, is beginning to emerge, but we shall have to wait for volume IV to see it having any marked effect on technological advance. Its complete ascendancy, which is a phenomenon of our own century, will see the end of the present *History*.

Realizing that, as we approach modern times, the history of technology becomes a facet of the wider history of economics, the editors have rightly 'sought to emphasize those changing aspects of technology which were of the greater social and economic importance, while at the same time illustrating the gradual penetration of technological innovation by the results of scientific inquiry'. They shrank, however, from producing a complete history of modern industrial civilization – at least they say that is not their object.

The volume is subdivided into five parts to which is added an epilogue on 'The Rise of the West'. The parts are Primary Production (power from water and coal, food, metals): Manufactures including tools, harness, textiles, and glass: Material Civilization (building, drainage, military technology, and printing): Communications including cartography and navigation.

The fifth, and final, part brings us to the threshold of the Industrial Revolution with five chapters dealing with the approach to science: the calendar, the clock, other philosophical instruments, and, foretaste of the dominant role it plays in civilization today, 'Invention in the Chemical Industry'.

In all there are some four hundred line drawings in the text which, despite the greater diversity of the sources from which they are drawn in this, later, volume, maintain the high standards of their predecessors. The half-tone supplement of some sixty illustrations is less happy: some of them are more than adequate, but many lose greatly by reproduction in black-and-white. This is a minor point: we look forward to the appearance of the next volume which, on the vast scale on which this work has been conceived, will almost seem to deal with our own era, if not with our own time.

THROUGH ALCHEMY TO CHEMISTRY: A Procession of Ideas and Personalities by John Read (London, G. Bell and Sons, 1957), pp. 206, 18s 6d.

Professor Read of St Andrews has a long-established reputation for his broadly based studies of alchemy and chemistry. Indeed, in the former field few are so knowledgeable. In the thirties and forties he produced an alchemical trilogy of which, however, two parts are now out of print. Happily the present volume is, in some ways, a condensation of the longer works, so that those who have not been fortunate enough to acquire all three, or otherwise have not enjoyed the leisure to browse through them, may now become familiar with Professor Read's scholarship in condensed form, as it were.

Alchemy is not everyone's interest, but it cannot be denied that alchemical practice had a profound effect on human thought for the best part of a thousand years. Mystical and irrational though it may appear to modern minds, it has often struck the reviewer that the philosopher's stone part of it may have been less esoteric than we are inclined to imagine. Is there, for example, a correlation between the degree of alchemical activity and shortage of currency? After all, the invention of paper currency beat the alchemists at their own game and would have undermined the rationale of their mysterious experiments if the logic of expanding chemistry had not done so instead.

Professor Read's title is apt and accurate, for one is indeed led through alchemistic activity to the development of modern chemistry. The former journey takes three-quarters of the book, so the quarter that is devoted to chemistry, as we know it, is, at the very least, sketchy.

If one can muster any enthusiasm for alchemy, none better than Professor Read as expositer, but in an age that has been variously called 'The Plastics Age', 'The Age of Light Metals', and 'The Atomic Age' there is much to be said for taking alchemy for granted.

THE PREHISTORY OF AFRICA by H. Alimen, translated by A. H. Brodrick (London, Hutchinson Scientific and Technical, 1957), pp. 438, 63s.

This is not an easy book. It is a review of a subject that has advanced by leaps and bounds during the past two decades, and in doing so has altered many long-established ideas about the prehistory of man. No less an authority than the Abbé Breuil has described it as 'an important synthesis which will long remain of great use to those who explore Africa and its immense prehistoric riches'.

Few will be able to explore Africa in the literal sense, but armed with Professor Alimen's book one can engage in the most rewarding armchair exploration both in space and in time. She has sifted the vast accumulation of recent scientific research results and made it accessible to the reader who is prepared to meet her half-way. Throughout, the emphasis is on the latest knowledge of the geology, ancient geography, climatic variation, palaeontology, archaeology, and ethnography of the great continent. Nothing less will suffice because Africa in the past twenty years has proved as great a mine of intellectual wealth as it did in earlier years to those who sought but material gain. It is not far short of the mark to say that, today, Africa is the most important area in the world in which to explore the evolution of man.

Professor Alimen, who works at the Institute of Ethnography in the University of Paris, has herself carried out extensive excavations in Africa, and against this practical background she has produced the first comprehensive account of our knowledge of Africa and its significance in the story of man's past and evolution.

This is a translation of the 1955 French edition with additions, and though published in 1957 the author's preface to the English edition is also dated November 1955.

ABOUT OUR CONTRIBUTORS

H. K. HENISCH is a lecturer in physics at the University of Reading. He is also a graduate of Reading and has for many years been concerned with research in semiconductor physics, both in this country and in the United States. He is associated with the international journal on the Physics and Chemistry of Solids (Pergamon) and his latest book, *Rectifying Semiconductor Contacts* (Oxford), appeared earlier this year. His other interests include music, photography, and mountaineering.

P. H. JELLINCK was educated at St Bees School, Trinity College, Cambridge, and London University. He obtained his Ph.D. in 1954 for research on the metabolism of the oestrogens under the direction of Professor Sir Charles Dodds and continued this work in Canada at McGill University while on a Postdoctorate National Research Council Fellowship. Since his return to Britain he has taught chemistry at Norwood Technical College and was recently appointed Lecturer in Biochemistry at St Bartholomew's Hospital Medical College. He is married and takes a keen interest in rugby football. He is also the author of a book on *Biochemistry* in the *Teach Yourself* series.

F. N. JOHNSON is at present studying at Sandbach School, Cheshire, where he is specializing in science subjects. A member of the British Astronomical Association, and several other bodies, he became interested in problems concerning the origin of the lunar surface features several years ago, and intends to pursue the subject in future years.

A. M. MAYER is a graduate of University College London who turned from chemistry to plant physiology (Ph.D.). In 1950 he went to Israel where he is a member of the staff of the Botany Department of the Hebrew University, Jerusalem. His special research interests are algal physiology and the mass culture of algae, and the physiology and biochemistry of germination. He has previously contributed to *Science News* on subjects of plant physiological interest.

W. H. MARSHALL, a graduate of Glasgow University, is a teacher of science, though he spent the war years in the explosives industry. He is a Fellow of the Royal Astronomical Society, and a member of the Council of the Astronomical Society of Glasgow, and has

published four articles on astronomical topics in previous issues of *Science News*. At present he is working on a small book, *Our Planet Earth*, intended as a reader for secondary schools. Is married and has two sons to keep his feet on earth. Other interests include collecting books and trying to reduce his golf handicap to a mentionable figure.

W. SLUCKIN was born and received his early education in Warsaw. He graduated in 1942 in engineering and in 1950, after evening study, in psychology, both at the University of London; in 1955 he was awarded a Ph.D. Since 1951 he has been on the staff of the University of Durham, where he is now Lecturer in Psychology. He has published some text-books, a number of papers, and a Pelican, *Minds and Machines*.

D. S. TROUTON was born in 1924 and educated at Marlborough College; Trinity Hall, Cambridge; University College London; and University College Hospital, London. Having become interested in psychology and taken a degree in it, he qualified in medicine in 1952, after which he was mostly concerned with research in psychology and psychiatry at the Institute of Psychiatry, Maudsley Hospital, London. It is with great regret that we have to record his death in the earlier part of 1957.

CONTENTS OF RECENT ISSUES

44

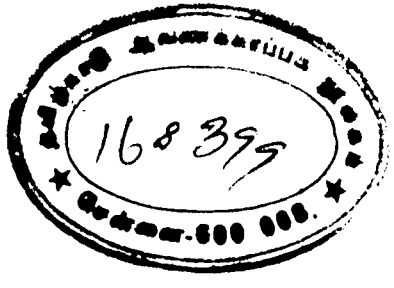
45

46

SUBJECT INDEX

Numbers 44–47

References are in all cases to number and page, e.g., **44**, 65. Articles are indexed under what is judged to be a key-word in their titles (printed in small capitals). Cross-references are suggested under a number of general subject names. Paragraphs under the heading *Research Report* are indicated by (*R.R.*). Continuing references are, in general, not given.



INDEX OF AUTHORS

Numbers 44–47

Further information about any author is to be found in the column *About our Contributors*, which appears in each issue. The numbers in heavy type indicate the number of the volume.

தவணைத்தநாள்.

Cr. எண் : 168399 ப. எண் : [illegible]

வழங்கிய நாள். (1)	திருப்பிய நாள். (2)	வழங்கிய நாள். (1)	திரும்பிய நாள். (2)

S.F. 259A-6—1,60,000—30-1-84.