

VI Gestalt and Universals

Among other things which we have discussed in the previous chapter is the possibility of assigning a neural mechanism to Locke's theory of the association of ideas. According to Locke, this occurs according to three principles: the principle of contiguity, the principle of similarity, and the principle of cause and effect. The third of these is reduced by Locke, and even more definitively by Hume, to nothing more than constant concomitance, and so is subsumed under the first, that of contiguity. The second, that of similarity, deserves a more detailed discussion.

How do we recognize the identity of the features of a man, whether we see him in profile, in three-quarters face, or in full face? How do we recognize a circle as a circle, whether it is large or small, near or far; whether, in fact, it is in a plane perpendicular to a line from the eye meeting it in the middle, and is seen as a circle, or has some other orientation, and is seen as an ellipse? How do we see faces and animals and maps in clouds, or in the blots of a Rorschach test? All these examples refer to the eye, but similar problems extend to the other senses, and some of them have to do with intersensory relations. How do we put into words the call of a bird or the stridulations of an insect? How do we identify the roundness of a coin by touch?

For the present, let us confine ourselves to the sense of vision. One important factor in the comparison of form of different objects is certainly the interaction of the eye and the muscles, whether they are the muscles within the eyeball, the muscles moving the eyeball, the muscles moving the head, or the muscles moving the body as a whole. Indeed, some form of this visual-muscular feedback system is important as low in the animal kingdom as the flatworms. There the negative phototropism, the tendency to avoid the light, seems to be controlled by the balance of the impulses from the two eyespots. This balance is fed back to the muscles of the trunk, turning the body away from the light, and, in combination with the general impulse to move forward, carries the animal into the darkest region accessible. It is interesting to note that a combination of a pair of photocells with appropriate amplifiers, a Wheatstone bridge for balancing their outputs, and further amplifiers controlling the input into the two motors of a twinscrew mechanism would give us a very adequate negatively phototropic control for a little boat. It would be difficult or impossible for us to compress this mechanism into the dimensions that a flatworm can carry; but here we merely have another exemplification of the fact that must by now be familiar to the reader, that living mechanisms tend to have a much smaller space scale than the mechanisms best suited to the techniques of human artificers, although, on the other hand, the use of electrical techniques gives the artificial mechanism an enormous advantage in speed over the living organism.

Without going through all the intermediate stages, let us come at once to the eye-muscle feedbacks in man. Some of these are of purely homeostatic nature, as when the pupil opens in the dark and closes in the light, thus tending to confine the

flow of light into the eye between narrower bounds than would otherwise be possible. Others concern the fact that the human eye has economically confined its best form and color vision to a relatively small fovea, while its perception of motion is better on the periphery. When the peripheral vision has picked up some object conspicuous by brilliancy or light contrast or color or above all by motion, there is a reflex feedback to bring it into the fovea. This feedback is accompanied by a complicated system of interlinked subordinate feedbacks, which tend to converge the two eyes so that the object attracting attention is in the same part of the visual field of each, and to focus the lens so that its outlines are as sharp as possible. These actions are supplemented by motions of the head and body, by which we bring the object into the center of vision if this cannot be done readily by a motion of the eyes alone, or by which we bring an object outside the visual field picked up by some other sense into that field. In the case of objects with which we are more familiar in one angular orientation than another—writing, human faces, landscapes, and the like—there is also a mechanism by which we tend to pull them into the proper orientation.

All these processes can be summed up in one sentence: we tend to bring any object that attracts our attention into a standard position and orientation, so that the visual image which we form of it varies within as small a range as possible. This does not exhaust the processes which are involved in perceiving the form and meaning of the object, but it certainly facilitates all later processes tending to this end. These later processes occur in the eye and in the visual cortex. There is considerable evidence that for a considerable number of stages each step in this process diminishes the number of neuron channels involved in the

transmission of visual information, and brings this information one step nearer to the form in which it is used and is preserved in the memory.

The first step in this concentration of visual information occurs in the transition between the retina and the optic nerve. It will be noted that while in the fovea there is almost a one-one correspondence between the rods and cones and the fibers of the optic nerve, the correspondence on the periphery is such that one optic nerve fiber corresponds to ten or more end organs. This is quite understandable, in view of the fact that the chief function of the peripheral fibers is not so much vision itself as a pickup for the centering and focusing-directing mechanism of the eye.

One of the most remarkable phenomena of vision is our ability to recognize an outline drawing. Clearly, an outline drawing of, say, the face of a man, has very little resemblance to the face itself in color, or in the massing of light and shade, yet it may be a most recognizable portrait of its subject. The most plausible explanation of this is that, somewhere in the visual process, outlines are emphasized and some other aspects of an image are minimized in importance. The beginning of these processes is in the eye itself. Like all senses, the retina is subject to accommodation; that is, the constant maintenance of a stimulus reduces its ability to receive and to transmit that stimulus. This is most markedly so for the receptors which record the interior of a large block of images with constant color and illumination, for even the slight fluctuations of focus and point of fixation which are inevitable in vision do not change the character of the image received. It is quite different on the boundary of two contrasting regions. Here these fluctuations produce an alternation between one stimulus and another, and this alternation, as we see in the

phenomenon of after-images, not only does not tend to exhaust the visual mechanism by accommodation but even tends to enhance its sensitivity. This is true whether the contrast between the two adjacent regions is one of light intensity or of color. As a comment on these facts, let us note that three-quarters of the fibers in the optic nerve respond only to the flashing "on" of illumination. We thus find that the eye receives its most intense impression at boundaries, and that every visual image in fact has something of the nature of a line drawing.

Probably not all of this action is peripheral. In photography, it is known that certain treatments of a plate increase its contrasts, and such phenomena, which are of non-linearity, are certainly not beyond what the nervous system can do. They are allied to the phenomena of the telegraph-type repeater, which we have already mentioned. Like this, they use an impression which has not been blurred beyond a certain point to trigger a new impression of a standard sharpness. At any rate, they decrease the total unusable information carried by an image, and are probably correlated with a part of the reduction of the number of transmission fibers found at various stages of the visual cortex.

We have thus designated several actual or possible stages of the diagrammatization of our visual impressions. We center our images around the focus of attention and reduce them more or less to outlines. We have now to compare them with one another, or at any rate with a standard impression stored in memory, such as "circle" or "square." This may be done in several ways. We have given a rough sketch which indicates how the Lockean principle of contiguity in association may be mechanized. Let us notice that the principle of contiguity also covers much of the other Lockean principle of similarity. The different

aspects of the same object are often to be seen in those processes which bring it to the focus of attention, and of other motions which lead us to see it, now at one distance and now at another, now from one angle and now from a distinct one. This is a general principle, not confined in its application to any particular sense and doubtless of much importance in the comparison of our more complicated experiences. It is nevertheless probably not the only process which leads to the formation of our specifically visual general ideas, or, as Locke would call them, "complex ideas." The structure of our visual cortex is too highly organized, too specific, to lead us to suppose that it operates by what is after all a highly generalized mechanism. It leaves us the impression that we are here dealing with a special mechanism which is not merely a temporary assemblage of general-purpose elements with interchangeable parts, but a permanent sub-assembly like the adding and multiplying assemblies of a computing machine. Under the circumstances, it is worth considering how such a sub-assembly might possibly work and how we should go about designing it.

The possible perspective transformations of an object form what is known as a group, in the sense in which we have already defined one in Chapter II. This group defines several sub-groups of transformations: the affine group, in which we consider only those transformations which leave the region at infinity untouched; the homogeneous dilations about a given point, in which one point, the directions of the axes, and the equality of scale in all directions are preserved; the transformations preserving length; the rotations in two or three dimensions about a point; the set of all translations; and so on. Among these groups, the ones we have just mentioned are continuous; that is, the operations belonging to them are determined by the values of

a number of continuously varying parameters in an appropriate space. They thus form multidimensional configurations in n -space, and contain sub-sets of transformations which constitute regions in such a space.

Now, just as a region in the ordinary two-dimensional plane is covered by the process of scanning known to the television engineer, by which a nearly uniformly distributed set of sample positions in that region is taken to represent the whole, so every region in a group-space, including the whole of such a space, can be represented by a process of *group scanning*. In such a process, which is by no means confined to a space of three dimensions, a net of positions in the space is traversed in a one-dimensional sequence, and this net of positions is so distributed that it comes near to every position in the region, in some appropriately defined sense. It will thus contain positions as near to any we wish as may be desired. If these "positions," or sets of parameters, are actually used to generate the appropriate transformations, it means that the results of transforming a given figure by these transformations will come as near as we wish to any given transformation of the figure by a transformation operator lying in the region desired. If our scanning is fine enough, and the region transformed has the maximum dimensionality of the regions transformed by the group considered, this means that the transformations actually traversed will give a resulting region overlapping *any* transform of the original region by an amount which is as large a fraction of its area as we wish.

Let us then start with a fixed comparison region and a region to be compared with it. If at any stage of the scanning of the group of transformations the image of the region to be compared under some one of the transformations scanned coincides

more perfectly with the fixed pattern than a given tolerance allows, this is recorded, and the two regions are said to be alike. If this happens at no stage of the scanning process, they are said to be unlike. This process is perfectly adapted to mechanization, and serves as a method to identify the shape of a figure independently of its size or its orientation or of whatever transformations may be included in the group-region to be scanned.

If this region is not the entire group, it may well be that region *A* seems like region *B*, and that region *B* seems like region *C*, while region *A* does not seem like region *C*. This certainly happens in reality. A figure may not show any particular resemblance to the same figure inverted, at least in so far as the immediate impression—one not involving any of the higher processes—is concerned. Nevertheless, at each stage of its inversion, there may be a considerable range of neighboring positions which appear similar. The universal “ideas” thus formed are not perfectly distinct but shade into one another.

There are other more sophisticated means of using group scanning to abstract from the transformations of a group. The groups which we here consider have a “group measure,” a probability density which depends on the transformation group itself and does not change when all the transformations of the group are altered by being preceded or followed by any specific transformation of the group. It is possible to scan the group in such a way that the density of scanning of any region of a considerable class—that is, the amount of time which the variable scanning element passes within the region in any complete scanning of the group—is closely proportional to its group measure. In the case of such a uniform scanning, if we have any quantity depending on a set *S* of elements transformed by the group, and

if this set of elements is transformed by all the transformations of the group, let us designate the quantity depending on S by $Q(S)$, and let us use TS to express the transform of the set S by the transformation T of the group. Then $Q(TS)$ will be the value of the quantity replacing $Q(S)$ when S is replaced by TS . If we average or integrate this with respect to the group measure for the group of transformations T , we shall obtain a quantity which we may write in some such form as

$$\int Q(TS)dT \quad (6.01)$$

where the integration is over the group measure. Quantity 6.01 will be identical for all sets S interchangeable with one another under the transformations of the group, that is, for all sets S which have in some sense the same form or *Gestalt*. It is possible to obtain an approximate comparability of form where the integration in Quantity 6.01 is over less than the whole group, if the integrand $Q(TS)$ is small over the region omitted. So much for group measure.

In recent years, there has been a good deal of attention to the problem of the prosthesis of one lost sense by another. The most dramatic of the attempts to accomplish this has been the design of reading devices for the blind, to work by the use of photo-electric cells. We shall suppose that these efforts are confined to printed matter, and even to a single type face or to a small number of type faces. We shall also suppose that the alignment of the page, the centering of the lines, the traverse from line to line are taken care of either manually or, as they may well be, automatically. These processes correspond, as we may see, to the part of our visual *Gestalt* determination which depends on muscular feedbacks and the use of our normal centering, orienting, focusing, and converging apparatus. There now ensues the problem

of determining the shapes of the individual letters as the scanning apparatus passes over them in sequence. It has been suggested that this be done by the use of several photoelectric cells placed in a vertical sequence, each attached to a sound-making apparatus of a different pitch. This can be done with the black of the letters registering either as silence or as sound. Let us assume the latter case, and let us assume three photocell receptors above one another. Let them record as the three notes of a chord, let us say, with the highest note on top and the lowest note below. Then the letter capital F, let us say, will record

————— Duration of upper note
——— Duration of middle note
— Duration of lower note

The letter capital Z will record

—————
—
—————

the letter capital O

—
— —
—

and so on. With the ordinary help given by our ability to interpret, it should not be too difficult to read such an auditory code, not more difficult than to read Braille, for instance.

However, all this depends on one thing: the proper relation of the photocells to the vertical height of the letters. Even with standardized type faces, there still are great variations in the size

of the type. Thus it is desirable for us to be able to pull the vertical scale of the scanning up or down, in order to reduce the impression of a given letter to a standard. We must at least have at our disposal, manually or automatically, some of the transformations of the vertical dilation group.

There are several ways we might do this. We might allow for a mechanical vertical adjustment of our photocells. On the other hand, we might use a rather large vertical array of photocells and change the pitch assignment with the size of type, leaving those above and below the type silent. This may be done, for example, with the aid of a schema of two sets of connectors, the inputs coming up from the photocells, and leading to a series of switches of wider and wider divergence, and the outputs a series of vertical lines, as in Fig. 8. Here the single lines represent the leads from the photocells, the double lines the leads to the oscillators, the circles on the dotted lines the points of connections between incoming and outgoing leads, and the dotted lines themselves the leads whereby one or another of a bank of oscillators is put into action. This was the device, to which we have referred in the introduction, designed by McCulloch for the purpose of adjusting to the height of the type face. In the first design, the selection between dotted line and dotted line was manual.

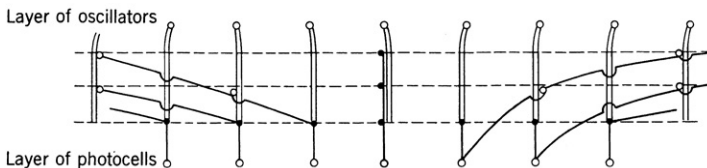


Fig. 8

This was the figure which, when shown to Dr. von Bonin, suggested the fourth layer of the visual cortex. It was the connecting circles which suggested the neuron cell bodies of this layer, arranged in sub-layers of uniformly changing horizontal density, and size changing in the opposite direction to the density. The horizontal leads are probably fired in some cyclical order. The whole apparatus seems quite suited to the process of group scanning. There must of course be some process of recombination in time of the upper outputs.

This then was the device suggested by McCulloch as that actually used in the brain in the detection of visual *Gestalt*. It represents a type of device usable for any sort of group scanning. Something similar occurs in other senses as well. In the ear, the transposition of music from one fundamental pitch to another is nothing but a translation of the logarithm of the frequency, and may consequently be performed by a group-scanning apparatus.

A group-scanning assembly thus has well-defined, appropriate anatomical structure. The necessary switching may be performed by independent horizontal leads which furnish enough stimulation to shift the thresholds in each level to just the proper amount to make them fire when the lead comes on. While we do not know all the details of the performance of the machinery, it is not at all difficult to conjecture a possible machine conforming to the anatomy. In short, the group-scanning assembly is well adapted to form the sort of permanent sub-assembly of the brain corresponding to the adders or multipliers of the numerical computing machine.

Lastly, the scanning apparatus should have a certain intrinsic period of operation which should be identifiable in the performance of the brain. The order of magnitude of this period

should show in the minimum time required for making direct comparison of the shapes of objects different in size. This can be done only when the comparison is between two objects not too different in size; otherwise, it is a long-time process, suggestive of the action of a non-specific assembly. When direct comparison seems to be possible, it appears to take a time of the order of magnitude of a tenth of a second. This also seems to accord with the order of magnitude of the time needed by excitation to stimulate all the layers of transverse connectors in cyclical sequence.

While this cyclical process then might be a locally determined one, there is evidence that there is a widespread synchronism in different parts of the cortex, suggesting that it is driven from some clocking center. In fact, it has the order of frequency appropriate for the alpha rhythm of the brain, as shown in electroencephalograms. We may suspect that this alpha rhythm is associated with form perception, and that it partakes of the nature of a sweep rhythm, like the rhythm shown in the scanning process of a television apparatus. It disappears in deep sleep, and seems to be obscured and overlaid with other rhythms, precisely as we might expect, when we are actually looking at something and the sweep rhythm is acting as something like a carrier for other rhythms and activities. It is most marked when the eyes are closed in waking, or when we are staring into space at nothing in particular, as in the condition of abstraction of a yogi,¹ when it shows an almost perfect periodicity.

We have just seen that the problem of sensory prosthesis—the problem of replacing the information normally conveyed through a lost sense by information through another sense still available—is important and not necessarily insoluble. What makes it more hopeful is the fact that the memory and

association areas, normally approached through one sense, are not locks with a single key but are available to store impressions gathered from other senses than the one to which they normally belong. A blinded man, as distinguished perhaps from one congenitally blind, not only retains visual memories earlier in date than his accident but is even able to store tactile and auditory impressions in a visual form. He may feel his way around a room, and yet have an image of how it ought to look.

Thus, part of his normal visual mechanism is accessible to him. On the other hand, he has lost more than his eyes: he has also lost the use of that part of his visual cortex which may be regarded as a fixed assembly for organizing the impressions of sight. It is necessary to equip him not only with artificial visual receptors but with an artificial visual cortex, which will translate the light impressions on his new receptors into a form so related to the normal output of his visual cortex that objects which ordinarily look alike will now sound alike.

Thus the criterion of the possibility of such a replacement of sight by hearing is at least in part a comparison between the number of recognizably different visual patterns and recognizably different auditory patterns *at the cortical level*. This is a comparison of amounts of information. In view of the somewhat similar organization of the different parts of the sensory cortex, it will probably not differ very much from a comparison between the areas of the two parts of the cortex. This is about 100:1 as between sight and sound. If all the auditory cortex *were* used for vision, we might expect to get a quantity of reception of information about 1 per cent of that coming in through the eye. On the other hand, our usual scale for the estimation of vision is in terms of the relative distance at which a certain degree of resolution of pattern is obtained, and thus a 10/100 vision means

an amount of flow of information about 1 per cent of normal. This is very poor vision; it is, however, definitely not blindness, nor do people with this amount of vision necessarily consider themselves as blind.

In the other direction, the picture is even more favorable. The eye can detect all of the nuances of the ear with the use of only 1 per cent of its facilities, and still leave a vision of about 95/100, which is substantially perfect. Thus the problem of sensory prosthesis is an extremely hopeful field of work.

This is a portion of the eBook [doi:10.7551/mitpress/11810.001.0001](https://doi.org/10.7551/mitpress/11810.001.0001)
at

This is a section of [doi:10.7551/mitpress/11810.001.0001](https://doi.org/10.7551/mitpress/11810.001.0001)

Cybernetics or Control and Communication in the Animal and the Machine

By: Norbert Wiener

Citation:

Cybernetics or Control and Communication in the Animal and the Machine

By: Norbert Wiener

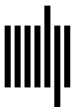
DOI: [10.7551/mitpress/11810.001.0001](https://doi.org/10.7551/mitpress/11810.001.0001)

ISBN (electronic): 9780262355902

Publisher: The MIT Press

Published: 2019

Funding for the open access edition was provided by the MIT Libraries Open Monograph Fund.



The MIT Press

© 2019, 1961, 1948 Massachusetts Institute of Technology

First MIT Press paperback edition, February 1965

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

This book was set in ITC Stone Serif Std and ITC Stone Sans Std by Toppan Best-set Premedia Limited. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Names: Wiener, Norbert, 1894-1964, author.

Title: Cybernetics ; or, Control and communication in the animal and the machine / Norbert Wiener ; forewords by Doug Hill and Sanjoy Mitter.

Other titles: Control and communication in the animal and the machine

Description: [Second edition, 2019 reissue]. | Cambridge, MA : The MIT Press, [2019] | "Reissue of the 1961 second edition." | Includes bibliographical references and index.

Identifiers: LCCN 2019005612 | ISBN 9780262537841 (pbk. : alk. paper)

Subjects: LCSH: Cybernetics. | Control theory. | System theory.

Classification: LCC Q310 .W47 2019 | DDC 003/.5--dc23 LC record available at <https://lccn.loc.gov/2019005612>

10 9 8 7 6 5 4 3 2 1