13

H. B. BARLOW

Physiological Laboratory, Cambridge University

Possible Principles Underlying the Transformations of Sensory Messages

A wing would be a most mystifying structure if one did not know that birds flew. One might observe that it could be extended a considerable distance, that it had a smooth covering of feathers with conspicuous markings, that it was operated by powerful muscles, and that strength and lightness were prominent features of its construction. These are important facts, but by themselves they do not tell us that birds fly. Yet without knowing this, and without understanding something of the principles of flight, a more detailed examination of the wing itself would probably be unrewarding. I think that we may be at an analogous point in our understanding of the sensory side of the central nervous system. We have got our first batch of facts from the anatomical, neurophysiological, and psychophysical study of sensation and perception, and now we need ideas about what operations are performed by the various structures we have examined. For the bird's wing we can say that it accelerates downwards the air flowing past it and so derives an upward force which supports the weight of the bird; what would be a similar summary of the most important operation performed at a sensory relay?

It seems to me vitally important to have in mind possible answers to this question when investigating these structures, for if one does not one will get lost in a mass of irrelevant detail and fail to make the crucial observations. In this paper I shall discuss three hypotheses according to which the answers would be as follows:

1. Sensory relays are for detecting, in the incoming messages, certain "passwords" that have a particular key significance for the animal.
   2. They are filters, or recoding centers, whose "pass characteristics"

can be controlled in accordance with the requirements of other parts of the nervous system.

3. They recode sensory messages, extracting signals of high relative entropy from the highly redundant sensory input.

These hypotheses are presented in order of increasing sophistication, and in the following pages most space is given to the last one, for the simple reason that it requires more thought—and has certainly consumed more of mine recently. I have omitted the idea that sensory relays are mere accidents of embryological or evolutionary development whose sole function is to pass on information without transforming it significantly, but this uninteresting possibility should probably be borne in mind, especially when considering the earlier relays. I am using the term "sensory relays" rather loosely, and I intend it to include synapses at the highest levels.

As with the bird's wing, the summaries are in physical rather than biological language, but before discussing them in greater detail two explanations and an apology are needed. First, it is unlikely that sensory relays perform just one operation of such outstanding importance that one can say it is *the* most important function, regarding all others as subsidiary in the same way that one can legitimately regard signaling, or sheltering young, as subsidiary functions of a bird's wing. Hence the present hypotheses are not mutually exclusive, nor do they exclude other theories about the important operations of sensory relays. Second, these are really orientating ideas, not detailed hypotheses about mechanism of action. The appropriate test for them is whether they help to make sense of the facts already known about "sensory integration," and whether the further investigations they prompt one to make are fruitful. Correct or not, I feel sure that ideas of this sort are needed. A bird's ability to fly is certainly an important fact, but it might easily be missed by someone concentrating his attention too narrowly on the anatomy and physiology of wings.

The apology is for the absence of a discussion of the experimental evidence bearing on these ideas. Nevertheless they do come from puzzling over experimental facts, not from abstract speculation. The "password" idea came from the realization that a frog's retina had an organization that made it quite unsuitable for the kind of task we use our own eyes for. The recoding idea came from recognizing that the retinal organization (in the cat in this case) was not only rather complicated but could also vary with the state of adaptation of the eye. It seemed to me that one could only hope to understand the complex, variable transformations the retina was imposing on the

sensory messages if one knew what they were directed toward, or what part they played in the whole animal. It may be too ambitious to try to answer this question, but at least I want to make it clear that I do not regard these ideas as moulds into which all experimental facts must be forced. They are just attempts to make some sense out of what would otherwise be a muddle.

Password Hypothesis

In studying sensory physiology many of us start with the idea that what we discover will be simply related to the subjective sensations of which we are aware by introspection. This is of course naive: the primary effect of the sensory messages an animal receives is not to enrich its subjective experience of the world but to modify its behavior in such a way that it and its species have a greater chance of survival. Accordingly it would be one step less naive to expect that, when sensory messages are transformed at sensory relays, they are being organized in accordance with the responses that the initiating stimuli would have produced in the normal animal. The subjective sensations they would produce in ourselves may or may not be relevant. Cutaneous stimuli that elicit flexion and withdrawal in the spinal cat are probably roughly congruent to those we call "painful," and those that elicit a scratch reflex may be analogous to those we call "tickling"; thus having these categories in mind is as helpful as thinking of the responses themselves. But we have no subjective category that adequately describes the class of stimuli that elicits the snapping response in frogs, though this is obviously an important category to have in mind when investigating the frog's visual system.

These preliminary remarks should have indicated what is meant by the "password" hypothesis. Specific classes of stimuli act as "releasers" and evoke specific responses; these classes of stimuli are thought of as "passwords" which have to be distinguished from all other stimuli, and it is suggested that their detection may be the important function of sensory relays. Looking at the case of flexion withdrawal, one sees that here the discrimination is mainly achieved, not at a sensory relay, but by having a class of sensory fibers that respond to potentially harmful stimuli. One knows little about the sensory discriminatory mechanism for the scratch reflex, except that it lies in the spinal cord. Probably no-one has recorded from a cell that performs the operation of distinguishing scratchworthy from unscratchworthy cutaneous stimuli, but it is worth asking whether, if one were picked

up, its function would be spotted during an ordinary physiological investigation. There is an objection that could be raised here. It might be held that the decision whether to scratch or not cannot be taken without considering the state and requirements of the rest of the animal. In this case you could not expect to find such a discriminating unit at a low level in the nervous system, but only at a level where all necessary information has been brought together. This applies to a unit that decides whether to scratch or not, but it does not apply to a unit that does the preliminary sorting of cutaneous stimuli into a class that should be scratched, and a class that should not. It is units doing this preliminary classification that one is led to expect if one bears in mind the responses ordinarily elicited by the stimuli employed.

Take the visual system of the frog as a specific example. The range of visual responses is rather limited. A small moving object elicits a sequence of reactions consisting of alerting, turning toward the object, hopping toward it if necessary, and finally hopping and snapping at it. Frogs also follow, with eye, head, and body movements, a moving object in the visual field, but a large moving object, especially if it is in the upper part of the visual field, may provoke an escape reaction in which the frog dives under a stone or into the deepest part of the pond. Yerkes (1903) was unable to get any evidence that frogs used vision to locate themselves in their habitat, nor did he find evidence of form discrimination or learned visual reactions. To some extent the neurophysiology fits in with this. The fact that it is predominantly change of retinal illumination which elicits discharges is obviously related to the fact that it is movement which is most effective in eliciting behavioral responses. One may be able to go further and identify the "on-off" units as the detectors of snapworthy objects (Barlow, 1953), for their properties are such that they respond vigorously to the type of stimulus that is particularly effective in eliciting the hunting sequence. Lettvin, et al. (1959) have recorded responses from the frog's optic tectum that seem to fit in with the behavioral requirements in a most striking manner, and it seems possible that the neurophysiology of the frog's hunting and feeding habits will become comprehensible in some detail.

If there is a moral to be drawn from the password hypothesis, it is as follows. We know that specific stimuli elicit specific responses, and it is reasonable to look out for the physiological mechanisms responsible for the preliminary classification of "releasers," even at early stages in the sensory pathways. To do this one needs some knowledge of the behavioral results of the stimuli one employs—and one must use stimuli that have specific behavioral results.

Controlled Pass-Characteristic Hypothesis

The idea that the incoming flow of sensory impulses is regulated or controlled at sensory relays is fashionable and has been experimentally fruitful, but the existence of such control raises many further points of interest. For instance, it is obvious that the control may be much more specific than is implied by the analogy of a volume or gain control. Sensitivity to one type of stimulus might be increased while another is decreased, or, combining this with the previous hypothesis, the whole characteristic of the relay might be changed, so that, in effect, the "password" is altered.

Another point is that it is not always obvious or easy to assess the significance of even a simple form of control, particularly if one fails to take into account more than the sensory pathway itself. To illustrate this, let us consider an example in the periphery. The γ efferents control the range of muscle length over which the discharge of the spindles shows finest gradation in accordance with changes in that length; since they appear to act as a zero adjustment, it was natural to think that the function of this control was to adjust the muscle spindles so that they could continue to give finely graded discharges at whatever length the muscle happened to be. The incompleteness of this picture of their function emerges when one takes into account the fact that afferent impulses from muscle spindles evoke a reflex discharge down the α efferents, causing powerful contraction in the muscle fibers lying in parallel with the spindles. Clearly activation of the γ efferents will bring about a reflex shortening, in the manner described by Eldred, Granit, and Merton (1953). In comparison with contractions produced by direct α-efferent excitation, the amount of shortening occurring in such servo-assisted contractions will be relatively independent of changes in the externally applied load and will be affected only slightly by moderate losses of muscle power resulting from fatigue. The task of controlling movement is thereby greatly simplified, and in understanding this we have gained considerable insight into the way the nervous system manages its affairs. If one is to gain comparable insight into the significance of controlled transmission at sensory relays, one must look beyond the effect of the control upon the afferent impulses themselves and consider what part these impulses play in the behavior of the intact animal.

Another point comes from theorists considering how to make a machine capable of learning to recognize complex patterns. In two schemes that have been offered (Lee, 1959; Selfridge, 1959), feedback is required from higher centers to points early in the pathway of

incoming information. The basic idea is to have elements early in the pathway that can change their transmission characteristics. They change in this way only when the feedback from above signals lack of success (for example, that the recognition problem has not been solved), but when success is signaled, the transmission characteristics being used are held unchanged.

Desirable or effective transmission characteristics thus survive as a result of a selective process rather analogous to natural selection acting on genes and causing evolutionary adaptation of species to their environment. It seems just possible that control fibers entering sensory relays might be exerting such a selective action, and because of its interesting implications this possibility might be worth exploring. The semipermanent change of "set" of the relays which this idea suggests needs to be looked for by experimental techniques rather different from those used to investigate continuous, moment-to-moment control of the type usually considered.

Redundancy-Reducing Hypothesis

The first hypothesis postulated preset mechanisms for detecting and passing on restricted classes of key signals, rejecting messages that did not fit into these classes. One can liken this to permanent editorial policy: for instance, one periodical only publishes information about sporting events and personalities, another rejects everything except original scientific papers. The second hypothesis suggested that the acceptance or rejection of messages might be controlled from elsewhere, either to make a temporary change in the type or amount of information passing, or to make a more permanent adjustment to the accept-reject criteria of the sensory relays. Pursuing the editorial analogy, one can liken the temporary control to rejection on the grounds of lack of space or to suit an editorial whim, the more permanent control to the long-lasting influence an editor can exert on the preliminary selection of news by his reporters. Now it is clear that there is one important editorial criterion for acceptance or rejection that is not included in either of these broad categories. Is this *news?* Has it been said before, or has it been said elsewhere? If so, it is redundant and can be rejected.

The idea that sensory relays try to ensure that what they pass on really is news is close to the basic one behind the third hypothesis. But one is liable to several misinterpretations if one thinks of the hypothesis solely in terms of the analogy. For this reason I have used

the language of information theory to state the hypothesis, together with certain simplifying assumptions. Then I have given a brief reminder of the meaning of terms such as redundancy and information, and following this an account of the sort of recoding the hypothesis leads one to expect, and the sort of predictions it leads one to make. I think the statement of the assumptions and hypothesis are precise and accurate, but they demand an accurate understanding of the meaning in information theory of the terms used; the brief reminder given here may not be sufficient to prevent misconceptions suggested by the editing analogy or by phrases like "stripping the sensory messages of their redundancy," and the only way to avoid these is to read an authoritative exposition of information theory (for example, Shannon and Weaver, 1949; Woodward, 1953).

The idea that reduction of redundancy is an important operation in the handling of sensory information is not a new one. Attneave (1954) argued that it made sense of many of the psychological facts of perception, and the points of view set out by MacKay (1956) and Craik (1943) are certainly closely related. I have written about it elsewhere (1959, in press) from a physiological point of view, and much further back in time one finds the idea, applied to much higher mental processes, clearly expressed in the writings of Ernst Mach (1886) and Karl Pearson (1892): their argument was that concepts, hypotheses, and laws of nature serve the purpose of bringing order and simplicity to our complex sensory experiences in order to achieve "economy of thought"; this seems to be the same idea as recoding to reduce the redundancy of our internal representation of the outer world.

Simplifying assumptions

1. For present purposes sensory pathways can be treated as noiseless systems using discrete signals.

2. The discrete signals are single impulses in particular nerve fibers in particular time intervals. For any one fiber and time interval, an impulse is either present or absent, so the code is binary.

3. The constraints on the capacity of a nerve pathway are the number of fibers F, the number of discrete time intervals per second R, and the average number of impulses per second per fiber I. The average number of impulses per fiber is assumed to be a variable constraint.

These are quite specific assumptions which cannot be justified by experimental results, nor are they really essential parts of the redundancy-reducing hypothesis. This would gain in generality if one dispensed with them entirely, but it is very difficult to discuss coding in the nervous system without making some assumptions about what variables in a nerve message are used to convey information, and it seemed wise to make those assumptions both explicit and as simple as possible. The assumptions adopted here are certainly oversimple in some respects. For instance, the first one side-steps the question of intrinsic neural noise, such as might be caused by random perturbations in transit time of impulses or by chance interruptions in synaptic transmission. This is not because I want to deny the importance or existence of these effects, but because the present hypothesis has something interesting to say about how the nervous system handles certain extrinsic properties of nerve messages—properties that are inherent in the physical stimuli impinging on the sense organs themselves. From this point of view, intrinsic noise, which is added to the messages at or after the sense organs, is a complicating factor that might obscure the issue, and so it seems best to neglect it at this stage.

Another point on which the assumptions might be criticized is that they fail to state some additional restraints that one feels pretty sure nerve fibers and synapses work under. For instance, FitzHugh (1957) has produced evidence that it is not the presence or absence of a single impulse in a particular short time interval that matters in a nerve message, but the aggregate number of impulses in a longer time interval. This additional restraint greatly decreases the capacity of a nerve fiber; consequently, if it holds in the higher parts of the nervous system, as well as in the simpler situation investigated by FitzHugh, my assumptions allow too much information to be passed down a nerve fiber. This has been done deliberately, because the safe course here is to assume that the nervous system is efficient. If it is clearly demonstrated that the nervous system is inefficient in some particular well-defined way, this can quite easily be incorporated into the hypothesis and its implications correspondingly modified, whereas our whole frame of thought might be undermined if it turned out that the nervous system was more efficient than we had supposed.

In fact the assumptions are simple; they suggest what we should look out for if the nervous system is smarter than we are inclined to think; and they define a communication system that will be helpful in discussing the hypothesis; but physiologically they are certainly oversimple and unproved, and they may be quite wide of the mark.

Hypothesis

The hypothesis is that sensory relays recode sensory messages so that their redundancy is reduced but comparatively little information is lost. To clarify this, what is meant by "information," "redundancy," "message," and so on, must first be explained.

A "message" is a set of "signals"; for example, it might be the particular pattern of impulses that arrives along a set of 10 fibers during an interval of $1/10$ seconds. These signals are carried into the relay along a set of fibers that constitute the "input" channel, and they generate impulses in other neurons that are the "output" signals in the output channel. If one writes down all the different input messages that occur, and for each input the output message that results, this will constitute the "code" relating input to output.

"Information" is a quantitative attribute of a message if the prior probability of receiving it is known. This usually means that it belongs to an ensemble or population of mutually exclusive and statistically independent messages whose frequency distribution is known. If P_m is the probability of the message m in such an ensemble, then the information attributed to m is $H_m = -\log P_m$; the average information of all messages is

$$H_{av} = -\sum_m P_m \log P_m$$

summed for all members of the ensemble. The rate of flow of information is H_{av}/T, where T is the average duration of messages from the ensemble, weighted for frequency of occurrence, that is,

$$T = \sum_m P_m T_m$$

The "capacity" C of a channel is equal to the greatest rate of flow of information that can be passed down it. This is calculated from its physical limitations and the constraints on the way it is used. For instance, with the constraints assumed under Simplifying Assumptions 3 above, the capacity of a nerve pathway is

$$C = -FR\left[\frac{I}{R}\log\frac{I}{R} + \left(1 - \frac{I}{R}\right)\log\left(1 - \frac{I}{R}\right)\right]$$

If messages of average information content H and duration T are passing down a channel of capacity C, the "relative entropy" of the messages is the ratio of rate of flow of information to capacity H/CT. The "redundancy" is 1 minus this ratio $[1 - (H/CT)]$; it can be thought of as the fraction of the channel capacity that is not occupied by the message it is being used to transmit.

Returning to the code that relates input to output, if there is a one-to-one relation, it is clear that there will be no loss of information, because the probability of each output is the same as that of its corresponding input. At first it might be suspected that the redundancy would also be unchanged, but this is not so; redundancy depends upon the capacity of the channel as well as the amount of information it is passing; so if the output has a lower capacity, and carries the same information, the redundancy must be less. The task achieved by a redundancy-reducing code is, in fact, to pack the messages more neatly, so that they can be passed down a smaller channel, with less unused space.

Now according to the assumptions, the only restraint on the output that can be varied is the average frequency of impulses. The capacity C is maximum for $I = R/2$; I is normally below this, so that if the sensory relay decreases the redundancy, it must do so by decreasing still further the average frequency of impulses being used to convey the input messages. We may suppose that the relay has a range of possible codes relating input to output: the hypothesis says that, for a given class of input message, it will choose the code that requires the smallest average expenditure of impulses in the output. Or putting it briefly, it economizes impulses; but it is important to realize that it can only do this on the average; the commonly occurring inputs are allotted outputs with few impulses, but there may be infrequent inputs that require more impulses in the output than in the input.

There is an important difference between this and the two previous hypotheses. They considered possible principles for selecting some sensory messages while rejecting others; and it was taken for granted that those selected would be rather infrequent, whereas those rejected would, in effect, all be classified alike. Consequently, both these hypotheses involve discarding a large fraction of the incoming information. In contrast, the emphasis in the present hypothesis is on the preservation of information: it is the redundancy that is discarded, and although an incidental loss of information may result this is not an essential feature of a redundancy-reducing code. If the morning paper fails to state that the sun set last night, one does not conclude that it did not happen, because one knows that this is the kind of event, which, though important, is omitted. In the same way, impulses can be economized without misleading the more central parts of the nervous system.

Recoding two binary inputs

The effect of coding to reduce redundancy is not just the elimination of wasteful neural activity. It constitutes a way of organizing the sensory information so that, on the one hand, an internal model of the environment causing the past sensory inputs is built up, while on the other hand the current sensory situation is represented in a concise way which simplifies the task of the parts of the nervous system responsible for learning and conditioning. In order to illustrate this, consider the very simple optimal code finder which has been described by the author and Mr. P. E. K. Donaldson (Barlow, 1959). Suppose there are two fibers A and B entering a relay, and two fibers X and Y leaving it. Take a time interval such that only one impulse can arrive in a fiber; now the possible input messages are impulse in A alone (Ab), impulse in B alone (aB), impulses in both (AB), or impulses in neither (ab). (Capitals symbolize impulses present, small letters impulses absent.) There are also four possible output messages, and these can be related one-to-one to the inputs in factorial $4 = 24$ ways, each of which is a reversible code in which no information is lost. To choose the appropriate code according to the hypothesis, it is necessary to measure the relative frequencies of the input messages. As soon as the commonest has been determined, it should be allotted to the output xy, since this is cheapest in terms of impulses. As soon as the rarest is known, this should be allotted to XY, the most expensive output. This narrows down the choice of code to 2 out of the 24. A choice between these two could be made if impulses in X and Y were valued differently; otherwise either would do.

The first claim for the code was in relation to the kind of neural model of the environment which Craik (1943) talked about. If one tries to think what is meant by such a neural model, it is clear that this must consist of a store of the frequencies of occurrence of a myriad of combinations and sequences of sensory stimuli. Excluding genetic factors, there is nothing else causally connected with the environment from which our internal representation of it could be constructed. It is the relative frequencies of the input states that are required for choosing the code, and so, conversely, the code chosen reflects these relative frequencies. Thus it acts as a store or model, but it is, of course, defective to the extent that it has only the rank order of input events, not their actual frequencies.

The fact that a redundancy-reducing code orders the input messages in accordance with their frequency of occurrences is also the basis

of the second claim, that such coding is a useful preliminary to the learning and conditioning tasks performed by the nervous system. There are, of course, some operations that it will not assist but will hinder. Take, for instance, the control of pupil diameter; this seems to require information about the average amount of light entering the pupil, and it is precisely such average properties that the code tends to subtract from the messages. On the other hand, in learning and conditioning, the animal does not act upon a predetermined feature of the sensory input but has to find sensory correlates of the rewards, punishments, and unconditional stimuli it receives before it can act on them. This is no simple task, for with input fibers numbered in millions the number of possible states of the input is more than just astronomical, it is meaninglessly large. Yet it would seem to be necessary to separate and inspect individually a rather large fraction of these possible states in order to have a reasonable chance of finding the required sensory correlate. This would be a formidable task.

After they have been coded, the messages are arranged according to their prior probabilities. Those containing a small number of impulses are the commonly occurring ones and lie at one extreme. Those containing many impulses occur infrequently and lie at the other extreme. In deciding which of the possible states one should inspect one would be greatly helped by this arrangement, for one could avoid allotting neural machinery to the task of discriminating between the vast numbers of possible states that contain many impulses and, therefore, occur infrequently or not at all. One could start the search with the possible states that contain few impulses and therefore include the states that occur most often, and by this means a vast curtailment of search effort would seem to be possible.

One can go a stage further along these lines; the requirement is to find which of the incoming messages are correlated with a particular event, such as a punishment, a reward, or the receipt of an unconditional stimulus. Now the frequency of this event can itself be determined, and then one could avoid wasteful searching among the possible messages that occur too frequently as well as too rarely; one could confine the search to those possible states that contain the appropriate number of impulses. If a rat runs a maze once a day, it should search for the key to the correct turning among patterns of sensory stimulation that also occur, very roughly, once a day. Obviously the frequency matching must not be too accurate, or the rat will be foiled by any nonregularity in the environment or in the experimental

design, but a degree of frequency matching does seem to make the correlate-finding problem much more feasible.

Returning now to the two-binary-input recoder, we can illustrate the following property of redundancy-reducing codes: an elementary signal in the output may correspond to an unexpected and less simple feature of the input. Suppose that the inputs A and B are both fairly infrequent, and that very nearly all instances of B are accompanied by A. This means that the input state aB is the rarest, and it must be allotted the output XY: that is, on the rare occasions it does occur, it causes both outputs to fire. Now since both A and B are fairly infrequent, the commonest input state is ab, which is accordingly allotted to xy. It then turns out that, if the code is reversible, one of the output fibers must correspond to the situation in which A and B are different from each other, so that it is active when A fires without B, or B without A. In other words, this output fiber signals when A and B hold this relation to each other, and it could not be understood or described adequately in terms of the responses to either input alone. This example, incidentally, illustrates the way in which the code is typically incomplete as a model of the environment, for it tells us that the input "B without A" has occurred less often than any other input, but it does not tell us whether it has never occurred, or occurred only very rarely. This does not mean that the code is unlike the nervous system's model, for that too is incomplete, and this type of omission might perhaps be characteristic here also.

Recoding more complex inputs

When one tries to consider a recoder for a more complicated input, an interesting situation arises. There are $(2^2)! = 24$ possible codes for two binary inputs; for n inputs, each with m discriminable levels of activity, there are $(m^n)!$ possible codes, a number that obviously gets impossibly large when m and n increase. It would be unreasonable to assume that the nervous system was able to choose any one of this number, so the range is presumably restricted by genetic and purely chance factors and possibly also by the "engineering difficulties" of arranging certain codes. The magnitude of $(m^n)!$ emphasizes that there is plenty of scope for such factors in limiting the choice of code; all the present hypothesis requires is that there should be a considerable range left to be selected from on the basis of frequencies of past sensory messages.

The way in which a code can act as a model of the environment

can be brought out more directly in a more complicated example. The principle of recoding is to find what messages are expected on the basis of past experience and then to allot outputs with few impulses to these expected inputs, reserving the outputs with many impulses for the unusual or unexpected inputs. Imagine that the incoming pattern of impulses forms a sensory "image" which can be likened to the pattern of light intensity in an optical image. To determine the expectations, take a time exposure, and develop the negative; the blackening then indicates the expected intensity in each part of the image. Now look at the sensory image at present being received through the negative. Any regions in the image that have not changed since the time exposure was started are reduced to a uniform gray, but regions that have changed stand out by being lighter or darker than their background. The procedure thus emphasizes the unusual at the expense of the usual.* In the same way, a redundancy-reducing code in the nervous system cuts down the impulse traffic from expected messages, whereas any sequence or combination of inputs that is unexpected on the basis of previous experience requires more impulses, and so stands out from the background. The code must be superior to the photographic negative in taking account of ordered sequences (that is, movement), but like the negative it is a representation of the environment—the fact that it is a negative model is not important.

This picture of the operation of a redundancy-reducing code also brings out its close relation to the "matching response" described by MacKay (1956). His conception was that a nervous center produces an outgoing signal that is an attempt to match the incoming signal. The "error" between incoming signal and matching response indicates how successful the attempt has been, and a second-stage matching response could be made to this error signal, and so on. Since the matching response must correspond to a redundant feature of the original signal, the effect of the operation is to recode the signal without this redundant element.

If one thinks of reducing redundancy as "economy of impulses" and "emphasizing the unusual," it will be seen that such recoding tends to impart a dual character to a nerve impulse. On the one hand, it signals the occurrence of a specific, but not necessarily simple, feature of the input. On the other hand, it also contributes to the sum total of impulses required to convey the information and thus helps to indicate how improbable the current sensory input is. Impulses are reserved for the unusual so that they carry more information indi-

* See illustration of a similar redundancy-reducing photographic process in Comment on Lateral Inhibition, p. 783 in this volume.

vidually, and the total number required in the output increases with the improbability of the input.

Predictions and speculations

Three predictions following from the hypothesis are, first, that impulse frequencies in response to "usual" stimuli should be decreased; second, that the type of transformations should change according to the probabilities of the stimuli being passed through the sensory relay; and, third, that the outputs may correspond to rather complex features of the inputs, not to properties that are simple in physical or anatomical terms. I think it is probably worth looking for evidence on these points with present techniques, though it should be realized that, even if the principle is correct, there are a great many more specific details to be given before one really has a working hypothesis on organization of the sensory input. For instance, one does not know how rapidly to expect the changes in code to follow a change in input. Reducing redundancy defines a strategy, but the tactics by which this objective is attacked are all-important, and nothing has been said about this. Now the object of this paper was to set out possible strategies for sensory integration, so I am going to confine my speculation about tactics of redundancy reduction to two points.

The first is that the coding out of redundancy is an operation that lends itself to subdivision. One can suppose that small parts of the input, selected by spatial and temporal contiguity, are dealt with in isolation before being brought together for coding out the more complex forms of redundancy. If the coding was successful in the first stage, the capacity of the channel acting as input to the next stage would be reduced. This serial reduction seems to be advantageous, because the difficulty in selecting a code is related to the number of possible codes, which in turn is dependent upon channel capacity, not the amount of information the channel is carrying. Of course, the recognition of a complex pattern, to which a complex recoding operation is comparable in difficulty, can also be conceived as a subdivided series of acts, but in this case it is not at all clear how the earlier acts take one toward the accomplishment of the complete recognition. In contrast, coding out redundancy must almost inevitably lead to the situation in which a single impulse in the output corresponds to a complex feature of the input, and this complex feature will be one that enables a concise, nonredundant description of the sensory situation to be given. One thus sees a way of breaking down pattern recognition into a succession of autonomous stages that do not need controlling from above.

The second point about tactics concerns the form of the output, which has not been specified either by the assumptions or by the hypothesis itself. These would allow the number of fibers conveying the information either to increase or to decrease, or to stay the same, but on this point anatomical clues are available. The situation is that the number of fibers may increase as one goes toward the cortex (see, for example, Galambos, 1954, for auditory pathway), or it may decrease (see Walls, 1953, for visual pathway). But whatever may happen in the subcortical relays, in the sensory areas of the cortex itself, there is a vastly greater number of cells than in the incoming fibers. Now it follows from the formula for the capacity of a nerve channel (given on p. 223 according to the admittedly oversimplified assumptions used here), that the aggregate of impulses required to carry the same information at the same redundancy is lower in a large channel than in a channel with fewer fibers; not only is I, the mean impulse frequency per fiber, less, but also IF, the total number of impulses in the whole pathway. Combined with reduction of redundancy, an enormous decrease in the number of impulses required seems to be possible without the loss of any information. One consequence is that impulses in some units may occur so rarely that it is possible to conceive that a response, such as salivation or the raising of a forepaw, would become linked in direct fashion to the occurrence of an impulse in that unit alone, rather than to impulses occurring in a particular combination of units. Expansion of channels works in the same direction as redundancy reduction; they both increase the informational value of single impulses at the higher levels.

It is amusing to speculate on the possibility that the whole of the complex sensory input we experience is represented, at the highest level, by activity in a very few, and perhaps only a single, neural unit at any one instant. At first this seems a monstrous suggestion, but consider how complex a sensory situation a skilled writer can evoke with a very small number of words. These words are taken at, say, 4 per second, and are chosen from a vocabulary of the order of 10^4; with an impulse chosen from a neural vocabulary of 10^6 cells, and occurring at an average rate of, say, 1 per 1/10 second (an average of about 1 per day per fiber) a representation of the current sensory situation should be possible which would be as complete as what we actually experience.

Now the process of reducing the redundancy must be stopped at some point, and instead the nervous system must disseminate its representation of the sensory input to all parts of the nervous system that require it: having edited its newspaper, it must print it and dis-

tribute it, and this is of course a redundancy-increasing process. The present speculation is that the sensory image that is thus disseminated consists of very few impulses and perhaps only a solitary one, in a very large array of nerve fibers. But whether this particular suggestion is right or not, it offends one's intuition, and one's experience of the efficiency and economy of naturally evolved mechanisms, to suppose that sensory messages are widely disseminated through the nervous system before they have been organized in a fairly non-redundant form.

Summary

This paper is not a discussion of the physiological mechanisms of sensory pathways, but an attempt to formulate ideas about the operations these mechanisms perform. "What are sensory relays for?" is the question posed, and three hypotheses are put forward as answers.

The first—the "password" hypothesis—really says that, since animals respond specifically to specific stimuli, their sensory pathways must possess mechanisms for detecting such stimuli and discriminating between them: one might therefore look for such mechanisms in neurophysiological preparations.

The second hypothesis is the fashionable one that relays act as control points at which the flow of information is modulated according to the requirements of other parts of the nervous system. It is pointed out that such control might have more interesting consequences than are suggested by the analogy of a simple gain or sensitivity control.

Most space is given to discussion of the third hypothesis, that reduction of redundancy is an important principle guiding the organization of sensory messages and is carried out at relays in the sensory pathways. Some simplifying assumptions about the information-carrying variables of nerve messages are made, followed by a statement of the hypothesis and an explanation of the terms used. Examples of recoding are described to illustrate its consequences, and predictions (which might be experimentally testable) and speculations (for entertainment only) are made.

To strip the redundancy from the preceding pages, what I have said is this: it is foolish to investigate sensory mechanisms blindly—one must also look at the ways in which animals make use of their senses. It would be surprising if the use to which they are put was not reflected in the design of the sense organs and their nervous pathways—

as surprising as it would be for a bird's wing to be like a horse's hoof.

References

Attneave, F.   Informational aspects of visual perception.   *Psychol. Rev.,* 1954, **61,** 183–193.

Barlow, H. B.   Summation and inhibition in the frog's retina.   *J. Physiol.,* 1953, **119,** 69–88.

Barlow, H. B.   Sensory mechanisms, the reduction of redundancy, and intelligence.   In National Physical Laboratory Symposium, *Mechanisation of Thought Processes.*   London: H.M. Stationery Office, 1959.

Barlow, H. B.   The coding of sensory messages.   In W. H. Thorpe and O. L. Zangwill (Editors), *Current Problems in Animal Behaviour.*   Cambridge: University Press (in press).

Craik, K. J. W.   *The Nature of Explanation.*   Cambridge: University Press, 1943.

Eldred, E., R. Granit, and P. A. Merton.   Supraspinal control of the muscle spindles and its significance.   *J. Physiol.,* 1953, **122,** 498–523.

FitzHugh, R.   The statistical detection of threshold signals in the retina.   *J. gen. Physiol.,* 1957, **40,** 925–948.

Galambos, R.   Neural mechanisms of audition.   *Physiol. Rev.,* 1954, **34,** 497–528.

Lee, R. J.   *Self-Programming Information and Control Equipment (SPICE).*   Falls Church, Virginia: Melpar Inc., 1959.

Lettvin, J. Y., H. R. Maturana, W. S. McCulloch, and W. H. Pitts.   What the frog's eye tells the frog's brain.   *Proc. Inst. rad. Engrs.,* 1959, **47,** 1940–1951.

Mach, E.   *The Analysis of Sensations.*   Trans. by C. M. Williams from 1st German ed., 1886.   Chicago and London: Open Court, 1914.

MacKay, D. M.   Towards an information-flow model of human behaviour.   *Brit. J. Psychol.,* 1956, **47,** 30–43.

Pearson, K.   *The Grammar of Science.*   London: Walter Scott, 1892.

Selfridge, O. G.   Pandemonium: a paradigm for learning.   In National Physical Laboratory Symposium, *Mechanisation of Thought Processes.*   London: H.M. Stationery Office, 1959.

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

Walls, G. L.   The lateral geniculate nucleus and visual histophysiology.   *Univ. Calif. Publ. Physiol.,* 1953, **9,** 1–100.

Woodward, P. M.   *Probability and Information Theory with Applications to Radar.*   London: Pergamon Press, 1953.

Yerkes, R. M.   The instincts, habits, and reactions of the frog.   I.   The associative processes of the green frog.   *Psychol. Monogr.,* 1903, **4,** 579–597.