

Neural Representations: Some Old Problems Revisited

Peter M. Milner

McGill University

Abstract

■ Hebb's (1949) cell assembly, originally conceived as an explanation for stimulus equivalence, also serves as a neural representation of stimuli. Association between cell assemblies was a major theme of Hebb's book, but the state of physiological knowledge at the time was such that no satisfactory basis for it could be devised. Subsequent theory has been more concerned with the recognition and "attractor" features of the cell assembly than its capacity to represent and associate concepts. This is unfortunate because while generalization is important, so is discrimination, which is not well served by an attractor model. This dilemma is avoided by postulating that stimulus representation and stimulus equivalence involve different neural circuits. Human beings can instantly form and use associations between many more concepts than there are synapses on the average cortical neuron, indicating that the associative links between engrams are sparse. The connections

within an engram, on the other hand, must be dense to ensure that a weak input can activate all its neurons. It would appear that two processes are anatomically and physiologically different, which may account for the fact that engrams remain distinct in spite of being associated with each other. The fact that a single concept may have very many associations puts a heavy demand on the process of selective attention to avert complete chaos. I propose that attention is a manifestation of motivation. Motivation facilitates responses, which in turn facilitate engrams of associated stimuli. The enhanced engram activity is fed back through centrifugal paths to intensify sensory input that has previously played a part in executing the planned responses. Attention may also contribute to a mechanism that prevents the engrams of component parts of an object from being assimilated into the engram of the whole. ■

NEURAL REPRESENTATIONS AND REALITY

Our nervous systems, which is to say we ourselves, rely on an extremely complex stream of neural impulses from receptors for all the information we have about our surroundings. The smells, songs, squirrels, streets, stars, and people that we encounter are concocted by our brains from the incoming electrochemical signals. When we see a house, or hear a dog bark, it is because the pattern of receptor stimulation has activated a neural representation of something we believe to be out there. We can never know where the stimuli *really* originate because we cannot get outside our sense organs. In fact, the concept of true reality is elusive; the neural representations are as near as we can get to it. They constitute the sum total of our inherited and acquired knowledge; clearly they are things that physiological psychologists should be interested in.

We cannot ask a newborn animal what it is experiencing, but we can observe its reactions to stimuli and it is obvious from such observations that its nervous system already has a considerable amount of organization. Some birds peck at bright objects, and follow moving ones, almost as soon as they are hatched, and most newborn mammals quickly find their mother's nipples. Stimuli that

evoke specific responses in naive animals must be recognized by their nervous system, so some representations of the outside world are innate. In adult higher vertebrates, however, the vast majority must have been learned. How this is done is still a matter for speculation.

STIMULUS EQUIVALENCE AND THE CELL ASSEMBLY

Early discussion of learned neural representations, or engrams (Schacter, 1992), was centered on their role in stimulus equivalence. A simple example of stimulus equivalence is that having learned that the shape O is called an oval, we do not have to relearn it after a small shift of visual fixation moves the image to a different part of the retina. None of the receptors stimulated during learning needs to be stimulated for the figure to be recognized later; only the stimulus pattern must be similar. More complicated examples of stimulus equivalence are objects like houses, or trees, that vary greatly in shape and size and yet access a common representation in the brain.

The first attempt to explain the phenomenon within the sphere of modern neurophysiology postulated that

visual stimuli induce waves of electrical activity that spread throughout the cortex, adding and subtracting from each other to set up interference patterns that were supposed to remain constant despite changes in the size or location of the stimulus (Lashley, 1942). Lashley was never able to develop this model beyond a vague idea, but the problem was taken up by one of his students, Hebb, who believed that stimulus equivalence must be incorporated into the engram as it is learned.

Hebb's model is described in his book, *The Organization of Behavior* (1949), and it is noteworthy not so much for its contribution to the solution of the stimulus equivalence problem as for its originality as a neural structure corresponding to percepts and concepts. According to Hebb the neurons that are fired by a particular stimulus bind to each other via strengthened synapses to form what he called a *cell assembly*. These connections ensure that if some of the neurons in an assembly fire, the rest are recruited. Hebb thought that this makes similar stimuli fire the same group of neurons so that they have the identical meaning.

Probably the most ingenious aspect of Hebb's model is that the strengthened interconnections between the neurons in cell assemblies form closed loops around which neural impulses circulate for some time, providing the basis for short-term memories. Furthermore, after sufficient learning has taken place, the assembly can be fired in the absence of the original stimulus, thus constituting an idea, or mental image of the object.

PERILS OF THE PIONEER

Unfortunately, during the 1940s when Hebb was devising this model, neurophysiology had just begun the great leap forward that is still going on. At that time the electrical theory of neural transmission still held sway, providing no simple explanation for inhibition or the differential action of drugs. The diversity of synaptic types, in fact, the very existence of receptor proteins, was hardly suspected and the processes underlying synaptic plasticity were completely unknown. Hebb's postulation of neural structures that represented percepts and concepts was an advance that brought physiological psychology from a study of peripheral processes into the realm of cognitive psychology, but the neurophysiology on which it was based was about to undergo a revolution.

A good theory can only be improved by a more thorough understanding of the underlying processes, but progress may also reveal flaws. Superficial flaws introduced by incorrect physiological information are usually put right without too much difficulty, and such changes have been made to cell-assembly-like models of the engram. Other shortcomings are more serious. The origin of the cell assembly as an explanation for stimulus equivalence was not the best of beginnings. Generalization is certainly an important and puzzling phenomenon,

but discrimination is equally important and tends to be ignored in theories of neural representation. Any theory of engram formation must take into account the relationship between category and instance—the ability we have to distinguish our own hat, house, and dog from hats, houses, and dogs in general.

Distinctive memories can be formed very quickly. If I am introduced to a stranger and speak to her for only a few seconds I may (or may not) forget her name, but after an hour or two, perhaps even a week, I still recognize her face and voice. I have added a complicated and detailed new representation to my database far more quickly than any network model I have heard of can acquire very simple ones, and with negligible effect on the previous contents.

Hebb's cell assembly, and especially later modifications of it (e.g., Milner, 1957; Hopfield, 1982; Kaplan, Weaver, & French, 1990; Hetherington & Shapiro, 1993), is what is now called an attractor. It is a pattern of activity with a lower threshold than any similar patterns and, as a result, it is impossible for slightly different patterns to compete with it. How then is it possible to distinguish one person from others? Paradoxically, stimulus equivalence is just as effective for the unique assembly as it is for the more general ones. My new acquaintance's face does not have to stimulate the same retinal receptors that were stimulated when the introduction took place for me to recognize it, undermining Hebb's view that stimulus equivalence requires a long period of exposure to different examples of the percept.

The original cell assembly was too loosely specified to convey its exact properties. Sometimes, indeed, Hebb endows it with characteristics without indicating clearly how they could arise in the structure he describes. It is difficult, for example, to understand why the synaptic modification that links neurons to form an assembly fails to involve more and more neurons until the whole brain becomes one immense and useless cell assembly. Apparently Hebb assumed that more than one cell assembly could be simultaneously active because, according to his learning postulate (1949, p. 62), this was a necessary condition for association. Later and more precisely specified versions intended for computer simulation usually allow only one assembly to be active at a time, which raises the problem of how associations between them take place. In fact, most of these later models tend to focus on the stimulus recognition and categorization functions of engrams, neglecting the association function that I consider central to the problem of intelligent behavior.

MORE QUESTIONS

If the association problem is seriously addressed, a number of questions arise, some of which were touched on by Hebb but are largely ignored in subsequent engram models. One is the matter of immediate memory. The

electrical theory of synaptic transmission strongly implied that increased synaptic strength depended on more and larger synapses, which take time to grow. Thus it was assumed that immediate memory must employ a different mechanism. A popular choice was neural echoing, or reverberation (Müller & Pilzecker, 1900; Hilgard & Marquis, 1940), and the closed loops of Hebb's cell assemblies offer a seductive explanation for such a mechanism. The complexities involved in remembering a telephone number such as 363-4336 by this method seem not to have been adequately considered.

Storing repetitions of the same stimulus is not the only problem introduced by a purely dynamic immediate memory; preventing excessive spread or premature decay would require a delicate balance of excitation, inhibition, and fatigue. Recovering and using the stored information would also be very difficult. In 1957 I suggested that immediate memory involved chemical traces that rendered cell-assembly neurons sensitive to nonspecific arousal input (Milner, 1957). This suggestion appears even more reasonable today in view of our present knowledge about synapses. It also has the advantage that only one retrieval mechanism is needed for both immediate and longer-term memories.

Another question has already been raised: How do associations between cell assemblies differ from the internal associations that are responsible for the assemblies' properties? It does not seem likely that both these processes can be the result of similar synaptic changes as is usually assumed. If they were, the interassembly associations would soon become intraassembly loops. A related puzzle is that parts are not submerged in the whole. Doors and windows are integral parts of my concept of a house, but they are also robust, stand-alone concepts. If the cell assembly is really an attractor, how can this be? The more comprehensive percept should have a lower threshold than the components and swallow them up.

Yet another problem that was touched on earlier is how several engrams can be simultaneously active without interfering with each other. The fact that I recognize my house does not prevent me from perceiving it as a member of the same category as millions of other houses. I may also see it as a familiar, small, red brick house. Do the engrams for "familiar," "small," and "red brick" that are associated with my house alternate with the engram for house, or are they all simultaneously active?

Perhaps the toughest problem of all concerns the fact that we have many associations with almost every engram. It is harder to think of something that is *not* associated in some way with a concept like "house" than to think of associated concepts. The brain must be a veritable rat's nest of tangled associations, yet for the most part we navigate through it with ease. When we are engaged in a task the appropriate associations automatically appear in the right sequence. Change the task

and a new set of associations takes their place. We attribute this to attention, which Hebb interpreted as prefacilitation of the appropriate cell assemblies. But where does the facilitation come from and how does it know which are the appropriate assemblies? None of these questions is answered by the commonly encountered computer simulations of neural networks.

INSTANT ASSOCIATIONS

The literature on cell assemblies, or engrams in general, leads one to suppose that they are formed slowly, over a period of years perhaps, and provide a basis for semantic memory. That is to say, they respond to a whole category of stimuli. This leaves out a vast number of individual neural representations corresponding to episodic memories, most of them relatively short-lived. Ablation of the medial temporal lobe reduces the duration of such memories to a few minutes at most, suggesting that several neural representations of stimuli, specialized for different functions, are created in different parts of the brain. Those in the medial temporal region are established almost instantly, and may last for a period of days. They appear to play an important role in habituation (Squire, 1987; Nadel, 1968; Teitelbaum & Milner, 1963).

Episodic memories may be established very quickly. The immigration officer who compares passport photographs with faces for hours a day must be able to store complex patterns in a matter of seconds and discover any discrepancies between them. Associations between engrams can also be established almost instantaneously. You probably do not associate Rembrandt with Thelonus Monk, but now that you have seen the names together you have formed an association that may last for hours, possibly days. It would even be available for about a minute in a patient with no hippocampus or medial temporal lobe (Milner, 1972).

The implications of this rather trite observation are far reaching. In the first place it means that a short burst of impulses can markedly change the effectiveness of some synapses for quite a long time. More significantly, it means that there must be potential synapses already in place between every one of our engrams, waiting to be modified should the occasion arise. The growth of a new synapse would take far longer than the time it takes to establish, and make use of, a new association between any arbitrary pair of concepts. Even though such episodic memories require an intact hippocampal region if they are to last for more than a minute or two (Milner, Corkin, & Teuber, 1968), it seems that there is at least a brief change in other engram synapses, probably located in the neocortex.

Cortical pyramidal cells have 5000 to 10,000 synapses, which means that if each neuron sends only one synapse to each of the other cells it contacts [a condition that is apparently close to being attained in the mouse cortex (Braitenberg, 1989)], and even if each of the receiving

neurons is part of a different engram (a much less likely proposition), the neuron may communicate directly with no more than 5000 engrams. In other words, individual cortical neurons are not equipped for immediate association with more than about 5000 engrams.

The question that immediately arises is how many engrams does the average person have? Presumably the number bears some rough relationship to the number of concepts, including response concepts, they have acquired. I believe, for reasons to be mentioned later, that most concepts are represented by a number of engrams in different modalities linked by strong associations (that sometimes fail, as when we forget a name). Thus, cats are represented by engrams for their shape, their smell, the feel of their fur, the sounds they make (which are not only separate concepts but have names of their own, purr, meow, caterwaul, etc.), and the read, heard, written, and pronounced versions of their name. I think there is little doubt that a person with only 5000 engrams would be regarded as severely retarded. The normal adult probably has at least a hundred times as many engrams as there are synapses on a cortical neuron.

If this is the case, and if any engram may be associated with any other, the absolute minimum number of neurons in an engram would be about 100, though in view of overlap and the impossibility of reliably firing a cortical neuron with input to only one synapse, a more practical size would be more like 1000 neurons. An assembly of 1000 neurons might have as many as 5 million synaptic sites available for receiving inputs from potentially associated engrams, or about 10 sites per engram. Such a sparse input could directly fire only a few of the engram's neurons, but they must all fire if the association is to be passed on to other engrams.

In the traditional cell assembly, spread of activity was postulated to occur via dense intraassembly connections established by growth of synapses over a considerable period of time, but there are difficulties in applying this solution. If the connections depend on the growth of synapses they cannot account for engrams that are immediately available; almost any stimulus, however unfamiliar, can be retained for long enough to compare it with other stimuli. In such cases the links between the neurons of an engram must be established via preexisting learning synapses at the time when the cells are fired together by the stimulus they represent.

Another serious problem is that an assembly of neurons linked by excitatory connections would be inherently unstable and liable to fire out of control at the slightest disturbance. How the brain avoids these pitfalls is not at all clear at present, but it has an endless fund of tools for achieving its ends. The spread could take place, for example, via two-way dendrodendritic learning synapses in the cortical neuropile that partly depolarize, but do not fire, the coupled neurons. Firing of the depolarized neurons could be brought about by pulses from a subcortical pacemaker, perhaps thalamic.

This sort of linking can occur only between neurons that are located fairly close together within one modality. Thus related activities in different modalities such as the smell, sight, and name of a rose, cannot merge into one engram, however often they fire together, though the engrams can become strongly associated. By seeking to understand instantaneous associations between arbitrary pairs of stimuli, we may have found a difference between the learning that creates engrams and the learning that associates one engram with another, a difference that explains why engrams never lose their individual identities even if they frequently fire together.

But what of concepts that are within the same modality and genuinely overlap, such as a face and its components—eyes, nose, mouth, and possible beard? The concept of a face cannot exist without these parts, yet the parts are not permanently merged in the whole; they maintain their independent existence and have their own associations, including their names, contrary to what the hebbian attractor tradition would predict. The answer to this question, if it can be found, must rest on an understanding of attention. Attention determines what sensory input does. If we are counting noses, those are what we see; if we are looking for a friend, the same visual stimuli arouse face engrams.

ATTENTION

Most psychologists take attention for granted. It is assumed that a rat in a maze is concentrating on finding food and that when human subjects have been asked to fixate a point on a screen they attend to it. Most attention research involves the visual system, and is often based on the assumption that attention sensitizes part of the visual field in some way. Broadbent (1956) studied attention switching in the auditory system, demonstrating that subjects can attend to a string of words going to one ear and then attend to the traces of different words that had been simultaneously presented to the other ear. This also implies a spatial allocation of attention.

Speculation about mechanisms rarely goes beyond postulating facilitation or inhibition (or both) of selected parts of the sensory pathways. Where the attention originates or even where it exerts its influence are questions not often addressed. This is no doubt partly because of a tendency to attribute attention to the "self." As Johnston and Dark (1986, p. 43) say in reference to selective attention: "It is difficult to conceptualize a process that is not well defined, and it is difficult to falsify empirically a vague conceptualization, especially one that relies on a homunculus." Hebb (1949) thought attention was the facilitation of cell assemblies, but he too was vague about its source.

When I look around my room the images on my retinas contain information sufficient to identify hundreds, perhaps thousands, of objects. It could take me

hours to list them all. If all my associations with each of them were to be added to the list, I might never finish it. Nevertheless, if it is raining outside I can pluck my umbrella from this plethora in a matter of seconds (minutes if I have put it in a safe place). I find the umbrella because that is what I am looking for; my visual system has been preset to detect umbrellas. The object of my attention was selected by my desire to go out without getting too wet. My visuomotor system could have been set to point to purple Xs on a computer monitor, in which case the object of my attention would probably have been selected either out of curiosity, or to be polite to an experimental psychologist.

In our everyday life it is stimuli that we mostly attend to—the smell of gas, the weather forecast, the mail, the newspaper—rarely a point on the wall 2° to the left of fixation. The mail does not occupy a particular part of the visual field (except that almost everything we attend to pretty soon occupies the region around central vision). It can be identified properly only when it activates an engram, which implies that its location in the visual field is no longer a factor in recognition. In fact, if I want to grab my mail it is no good facilitating mail-shaped patches all over my retina; facilitation has to be applied to a circuit that recognizes the abstract concept of mail, in other words, to the mail engram.

The problem with that idea, as I pointed out in an earlier paper (Milner, 1974), is that when the motor system gets involved it has to have information about location; the hand cannot pick up the concept of mail. Fortunately the necessary information is available in the early stages of the visual system, the primary visual cortex (V1), for example, if it can only be sorted out from all the other input that is of no immediate interest. This, of course, is what attention is supposed to do: the only question is how does the engram tell V1 which of its neurons are carrying information about mail? An obvious answer is that the message must be carried by the reciprocal connections that are plentiful throughout the sensory pathways. It is not quite as simple as that, however. The ascending signal has no record of where it came from: It had to throw away its tags to enter the conceptual level. To find the neurons in V1 that are processing mail input, the returning signals must follow a sort of paper-trail left by the ascending input signal (see Fig. 1).

Once this return route has been opened up, any intensification of the engram activity produced by attention (i.e., need) is transmitted back, perhaps as far as the retina itself, to amplify the signal coming from mail. The consequent increase of lateral inhibition no doubt reduces interference from other inputs. The most important function of this intensification, however, is to allow the selected input to gain control of the motor system. When my arm and hand get instructions to grasp something, they are not required to recognize what it is, mail or whatever; all they have to do is operate with the

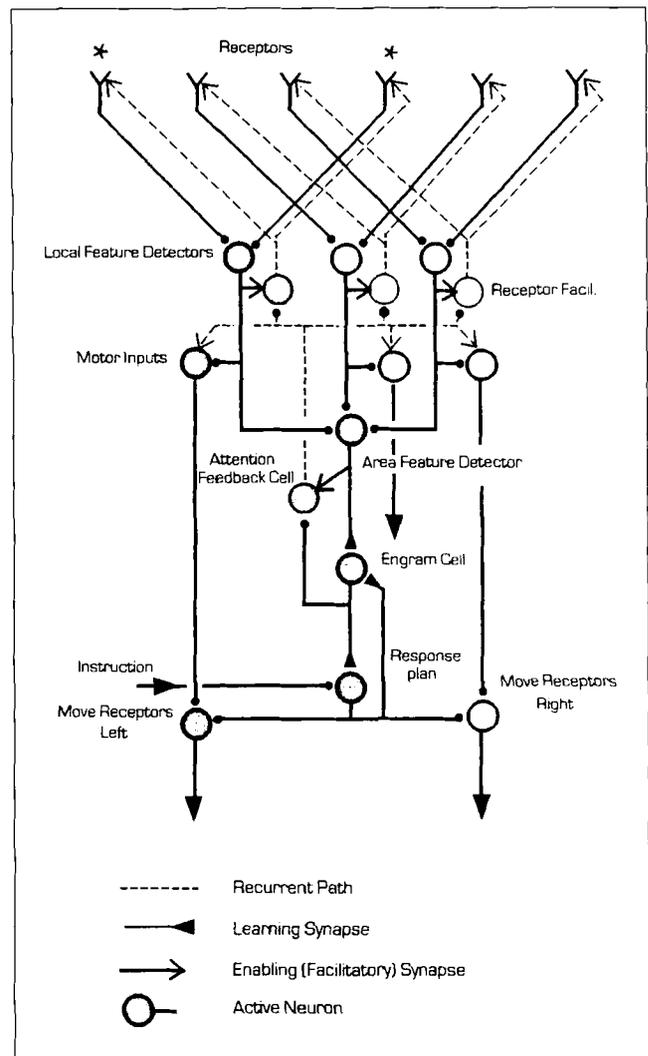


Figure 1. Diagram of attention circuits. This figure illustrates how a very simple stimulus (two spots three pixels apart) may fire the same engram (via a general feature detector) as it moves relative to the receptors. It is to be understood that other stimuli are being simultaneously detected by other circuits that, to avoid making the figure even more difficult to follow, are not shown. If execution of a response requires the stimulus (e.g., to carry out an instruction to fixate it), the response plan intensifies the engram activity. The combined input from the response plan and the general feature detector fires the engram vigorously, enabling it to fire the recurrent attention neuron. This neuron sends branches to relay neurons in the vicinity of all the inputs to the general feature detector. At that level, only the branch that is close to the active local feature detector finds a facilitated relay neuron that it can fire. The fired neuron delivers attentional facilitation to the receptors currently being stimulated. The intensified output from these receptors is delivered to the motor system, dominating the weaker signals of receptors not benefiting from attention, ensuring that the response plan is applied to the correct stimulus.

intensified part of the visual input. Other perhaps incidental effects of attentional facilitation are increased speed of reaction to changes in the selected input, improved memory for it, and, of course, increased awareness of it. The latter two are probably related; it is

difficult to see how one could know that something was experienced if one does not remember it. The strength of a memory probably depends on how intensely the engram fired.

MULTITASKING

At least some writers on attention have been struck by examples of divided attention, but as James (1890) points out, usually only one task demands attention; additional tasks require attention only intermittently. For example, it is possible to perform certain routine tasks, such as shaving, or walking to work, while attending to the news on the radio or composing a letter in one's head. Some behavior can become so automatic that it is impossible to remember it a few minutes later, or even seconds later if no immediate cues can be referred to. This does not seem consistent with the need for intensification of the stimuli that guide a response—avoiding obstructions such as lampposts, cars, and other pedestrians when walking for example.

If we accept that a stimulus or event may have several quasi-independent representations in different parts of the brain, however, responding to a stimulus without remembering it seems less mysterious. We may speculate that as permanent associations develop (presumably in the neocortex, cerebellum and basal ganglia, as they are unimpaired by limbic lesions), engrams no longer automatically form associations at episodic memory synapses (presumably in paleocortical structures where lesions do impair episodic memory). Neocortical engrams for the routine task are still facilitated by need, but they do not engage the neurons responsible for medium-term memory, which are thus free to respond to other unrelated engrams.

Broadbent's (1956) dichotic listening experiments are interesting in this respect. It is evident that the material reaching the unattended ear must activate its engram (nonsense material is much less likely to be recovered), but if attention is not switched quickly to that ear the information is lost for ever. Probably the initial activation leaves neocortical traces in the form of briefly potentiated synapses, and if the neurons are facilitated by attention before the potentiation has decayed, the activity is revived and amplified to embrace neurons or synapses whose potentiation is more persistent.

A somewhat similar process must go on when something one is not attending to suddenly becomes interesting for some reason. A short sentence, or a few words, may still be "replayed" by a switch of attention, though without it one might be completely unaware that they had been heard. Probably control of attention is not just a simple on-off, but involves facilitation of specific systems, the motor system or the episodic memory system of the hippocampal region, for example.

SELECTIVE ATTENTION AND RESPONSES

The traditional fascination of visual phenomena for experimental psychologists has had the unfortunate effect of diverting research away from the business end of the nervous system. Sensory systems evolved to enhance the effectiveness of the motor system, not to keep us entertained. Hebb was masterful in his criticism of the linear, sensory to motor picture of the nervous system that was widely held until 40 or 50 years ago (and is still not dead), but I do not believe he pursued the idea to its logical conclusion, which is that the response system has as much influence on what is perceived as the sensory system has on what is being performed. Probably more. An important function of the response system of the brain is to determine what sensory information is needed to support the ongoing response and direct attention to it.

What does this mean in neural terms? How can a response know what stimuli it needs? Where do responses come from anyway, if not from an ego that controls all our behavior? Responses originate in a need. Organisms are equipped to detect their needs, and use the information to generate appropriate responses. Needs may be based on internal stimuli, such as the constitution of the blood, or detected by skin receptors for itch, pain, temperature, and pleasurable sensations. Social animals like ourselves have complicated needs to dominate or be placatory, and almost all animals that can learn have a strong urge to investigate unfamiliar objects.

When an animal experiences a need, engrams for stimuli and responses that have previously been associated with the need are facilitated, i.e., they become foci of attention. If a stimulus whose engram is thus facilitated by a need is present, the chance that it will fire its engram is increased, and the resulting activity may select one of the response plans being simultaneously facilitated by association with the need input.

The need input now facilitates both the response plan and one or more of the engrams associated with the performance of the response, generating activity that returns along reciprocal connections to sensitize all the relay stations of the relevant sensory path. Thus the required stimulus dominates the sensory input and probably inhibits competing inputs. Its increased vigor of firing ensures that it is selected for delivery to the motor system to guide the selected movement. Completion of the movement changes the stimulus situation and modifies the response plan, the new version of which facilitates engrams for the next stimulus required to continue with the plan.

"Top-down" attention may therefore be somewhat of a misnomer. The assertion that responses, through attention, determine what is perceived, must be qualified. Stimuli indicating, or associated with, needs have a direct

pipeline to the response system. They are the ultimate source of attention and ensuing behavior. As I have argued elsewhere (Milner, 1961, 1977) there are two major routes through the nervous system, a direct, basically innate path available to the naive animal, through which stimuli that interact positively with a need input produce responses, and an indirect path through memory, available, of course, only to experienced animals. Need stimuli have access to the highest level of the memory, and it is true that from that point attention flows down from the top. But there is no little man sitting at the top dispensing attention. It is delivered to the top quite legitimately by sensory input related to needs, in much the same way as a flashing light or a noise induce "bottom-up" attention.

To illustrate the proposed system with a specific example, if you need a spoon to eat your soup or stir your coffee and one is already in the visual field, feedback from the aroused "spoon" engram travels back along parallel reciprocal paths and facilitates neurons in the visual path carrying input ascribed to the spoon. The rate of firing of these neurons is thereby increased, giving them control over the motor system, so that the hand is guided to the spoon, not to an adjacent fork for example. If no spoon is visible an engram representing an associated object that *is* visible, such as the drawer in which spoons are kept, is fired instead.

If nothing immediately associated with spoons is visible, the facilitation is relayed to associates of associates of spoons, and so on until some response that can be made with available sensory input is facilitated. If the train is long, with many branches that are eventually rejected or postponed, this process constitutes problem solving. One may, for instance, make a mental note to buy spoons at the next opportunity, which involves setting up a temporary association between shopping and spoons, an association that one hopes will persist for an indefinite time until the purchase has taken place, whereupon it promptly vanishes, or at least is rendered ineffective.

Engrams are thus the channels through which needs recruit responses and the stimuli necessary for their performance, as well as the links between stimuli and the numerous responses with which each is associated. Biological needs establish their own engrams, which acquire associations with all other engrams that frequently accompany their activation.

NESTED ENGRAMS, HOW DO THEY SURVIVE?

We return now to one of the paradoxes raised earlier; that the features of a face, or the components of a house or a perfume maintain control over independent engrams even though they are also represented by the engram of the face or other object of which they form a part. They may be perceived either as the whole object,

or as a component of that object, which appears to contradict the classical cell assembly's property of completion. Under some circumstances it is clear that engrams are aroused by partial sensory input; I see only the backs of the books on my shelf, but that is enough to arouse the complete book concept. Turning to the somewhat simpler chemical senses, most people can analyze complex tastes and smells into distinct components; the orange may have a sweet or a bitter taste, but that does not stop it tasting of orange. Expert wine tasters and blenders of tea or whisky have special sensitivity to complex odors, and *their* skills pale into insignificance when compared with that of a well-trained tracking dog.

This survival of concepts within other concepts is related to the problem raised earlier of why cell assemblies do not coagulate into one massive cell assembly, but it is somewhat more subtle. In this case, even sensory inputs within the same modality can elicit different but overlapping engrams. The smell of beer not only activates a "beer" engram; individual components of the smell such as alcohol, yeast, and hops can still, depending on circumstances, activate their own independently acquired engrams. Clearly sensory input is not the only determinant of engram activity. If a task requires a particular input for its performance, the engram network is designed to find that input, even if it is embedded in a more complex pattern.

As you will have gathered, feedback in the direction opposite to the conventional sensory to motor path is an important part of the attention mechanism. This feedback seems a likely explanation for the phenomenon described above, the ability of component parts of an object to maintain separate and independent concepts. Let us suppose that the smell of hops excites a pattern of receptors that gives rise to an olfactory engram; call it "hops-olf." This is associated, in both directions, with a word engram, "hops-name," representing the word hops. Now suppose we sniff some beer, exciting receptors, some of which are also excited by the smell of hops. The total pattern establishes an engram, "beer-olf," that is strongly associated with the auditory engram "beer-name."

The *name* engrams, "hops" and "beer," are not at all similar, and do not overlap, even though their olfactory associates do. The smell of beer presumably excites beer-olf and beer-name preferentially, but if hops-name has already been activated, either by someone pronouncing the name or by some other association, the hops-olf engram would be prefacilitated, amplifying the sensory input from the hops when beer is smelled. The hops part of the pattern would thus dominate, allowing hops to be detected. As long as the meanings or motor associations of overlapping sensory patterns are distinct, then any overlap at the sensory level will be sorted out by the attentional feedback.

CONCLUSION

That ideas and images must have a physical basis has occurred to some philosophers from ancient times, but until recently not enough was known about the nervous system to permit any plausible explanation of how it was done. Hebb made ingenious use of the anatomy of his day, but his attempts to explain association were frustrated by the almost complete lack of information about the physiology of neurons and synapses. The remarkable achievements of neurochemists during the last few years in elucidating the very complex series of reactions that accompany and follow synaptic transmission provide an opportunity for better founded speculation (Schwartz & Greenberg, 1987; Rose, 1991; Edelman & Gally, 1992; Sacktor et al., 1993; Colley & Routtenberg, 1993). Some synapses undergo immediate potentiation when they fire, accounting for short-term associations; some of the same synapses may also accumulate long-term increases in effectiveness during multiple firings over a long period of time, accounting for more permanent associations.

It seems, however, that short-term (episodic) associations can involve only a small fraction of the neurons in an engram, and are unlikely to activate the engram without help from attention or other convergent associations. They can never become strong enough to weld engrams together into an undifferentiated nexus.

Except for very primitive organisms, only a fraction of the total sensory input is used to control behavior at any moment. Even that part usually has many associations, both sensory and motor. Behavior would be chaotic without attention to select the associations appropriate for the task in hand. Traditionally we talk about "top-down" and "bottom-up" attention, but no one is very specific about the summit from which top-down attention flows. I suggest it originates in a need, which may be detected by internal or external receptors, or it may arise spontaneously, perhaps in neurons that have been unused for a long time. Its main function is to find and intensify relevant input that may be used to guide responses. The intensified input also makes a stronger impression on memory, and it can be isolated from complex patterns of which it is a component.

Reprint requests should be sent to Dr. Peter M. Milner, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec, Canada, H3A 1B1. e-mail: PS64@musica.mcgill.ca.

REFERENCES

- Braitenberg, V. (1989). Some arguments for a theory of cell assemblies in the cerebral cortex. In L. Nadel, L. A. Cooper, P. Culicover, & R. M. Harnish (Eds.), *Neural connections, mental computation* (pp. 137-145). Cambridge, MA: MIT Press.
- Broadbent, D. E. (1956). Successive responses to simultaneous stimuli. *Quarterly Journal of Experimental Psychology*, 8, 145-162.
- Colley, P. A., & Routtenberg, A. (1993). Long-term potentiation as synaptic dialogue. *Brain Research Reviews*, 18, 115-122.
- Edelman, G. M., & Gally, J. A. (1992). Nitric oxide: Linking space and time in the brain. *Proceedings of the National Academy of Sciences U.S.A.*, 89, 11651-11652.
- Hebb, D. O. (1949). *The organization of behavior* (pp. 1-335). New York: Wiley.
- Hetherington, P. A., & Shapiro, M. L. (1993). Simulating Hebb cell assemblies: The necessity for partitioned dendritic trees and a post-not-pre LTD rule. *Network*, 4, 135-153.
- Hilgard, E. R., & Marquis, D. G. (1940). *Conditioning and learning* (pp. 1-429). New York: Appleton-Century.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences U.S.A.*, 79, 2554-2558.
- James, W. (1890). *The principles of psychology*. New York: Henry Holt.
- Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, 37, 43-75.
- Kaplan, S., Weaver, M., & French, R. (1990). Active symbols and internal models: Towards a cognitive connectionism. *AI & Society*, 4, 51-71.
- Lashley, K. S. (1942). The problem of cerebral organization in vision. In H. Klüver (Ed.), *Visual mechanisms. Biological symposium* (Vol. 7, pp. 301-322).
- Milner, B. (1972). Disorders of learning and memory after temporal lobe lesions in man. *Clinical Neurosurgery*, 19, 421-446.
- Milner, B., Corkin, S., & Teuber, H.-L. (1968). Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of H. M. *Neuropsychologia*, 6, 215-234.
- Milner, P. M. (1957). The cell assembly: Mk. II. *Psychological Review*, 64, 242-252.
- Milner, P. M. (1961). The application of physiology to learning theory. In R. A. Patton (Ed.), *Current trends in psychological theory* (pp. 111-133). Pittsburgh, PA: University of Pittsburgh Press.
- Milner, P. M. (1974). A model for visual shape recognition. *Psychological Review*, 81, 521-535.
- Milner, P. M. (1977). Theories of reinforcement, drive, and motivation. In L. L. Iversen, S. D. Iversen, & S. H. Snyder (Eds.), *Handbook of psychopharmacology. Vol. 7: Principles of behavioral pharmacology* (pp. 181-200). New York: Plenum Press.
- Müller, G. E., & Pilzecker, A. (1900). Experimentelle Beiträge zur Lehre vom Gedächtniss. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane, Ergänzungsband*, 1, 1-288.
- Nadel, L. (1968). Dorsal and ventral hippocampal lesions and behavior. *Physiology and Behavior*, 3, 891-900.
- Rose, S. P. R. (1991). How chicks make memories: The cellular cascade from c-fos to dendritic remodelling. *Trends in Neurosciences*, 14, 390-397.
- Sacktor, T. C., Osten, P., Valsamis, H., Jiang, X., Naik, M. U., & Sublette, E. (1993). Persistent activation of the zeta isoform of protein kinase C in the maintenance of long-term potentiation. *Proceedings of the National Academy of Sciences U.S.A.*, 90, 8342-8346.
- Schacter, D. L. (1992). Semon, Richard. In L. R. Squire (Ed.), *Encyclopedia of learning and memory* (pp. 586-588). New York: Macmillan.

Schwartz, J. H., & Greenberg, S. M. (1987). Molecular mechanisms for memory: Second-messenger induced modifications of protein kinases in nerve cells. *Annual Review of Neuroscience*, 10, 459-476.

Squire, L. R. (1987). Memory: Neural organization and behavior. In F. Plum (Ed.), *Handbook of physiology—Higher*

functions of the brain, V. Section I. *The nervous system*, 4th Ed. (pp. 295-371). Bethesda, MD: American Physiological Society.

Teitelbaum, H., & Milner, P. M. (1963). Activity changes following partial hippocampal lesions in rats. *Journal of Comparative and Physiological Psychology*, 56, 284-289.