OXFORD · PSYCHOLOGY · SERIES
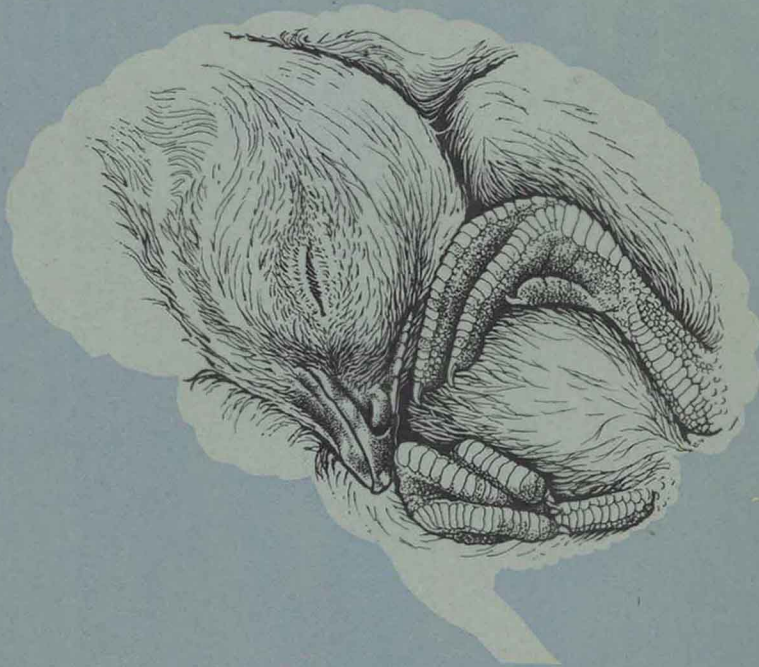
# Memory, Imprinting, and the Brain

## An inquiry into mechanisms



# Gabriel Horn

# Memory, Imprinting, and the Brain

## An Inquiry into Mechanisms

GABRIEL HORN

*Professor of Zoology,*
*University of Cambridge*

# OXFORD PSYCHOLOGY SERIES

EDITORS
DONALD E. BROADBENT
JAMES L. McGAUGH
NICHOLAS J. MACKINTOSH
MICHAEL I. POSNER
ENDEL TULVING
LAWRENCE WEISKRANTZ

# Preface

Learning is a complex set of processes involving the acquisition, storage, and retrieval of information. Acquisition is studied in various guises and includes the encoding, transmission, and processing of sensory information, behavioural arousal, and attention. This list of subprocesses is far from complete. Recent advances in several of these fields have been little short of spectacular. In contrast, little is known of the physiological basis of the retrieval process, and advances in our understanding of the mechanisms by which information is stored in the brain are recent and have been made on a limited front. The purpose of writing this book is to say something about these advances. The book is not, nor is it intended to be, a general review of this field, but gives an account of work in which I have been involved, over the past two decades or so, on habituation and imprinting. During that time modest success has been achieved in analysing habituation—a common change of behaviour, but a rather special form of learning. Some of this work, particularly as it impinges on the nature of information storage in nervous systems, is described in the first chapter. But the bulk of the book is devoted to imprinting. If the small amount of space allotted by some textbooks of psychology is used as a guide, then imprinting too is a rather special form of learning. I hope this book will clarify what is special about imprinting, and what it shares in common with other kinds of learning.

The problem of the neural bases of memory has taken a long time to yield; one of the reasons is that the problem lies at the boundary between the brain sciences and the behavioural sciences. I consider that the intellectual and technical resources of scientists from both fields need to be pooled to maximize the chances of resolving the problem. In so far as this hope has been realized in this study of imprinting, it is because such skills have been deployed. The study began in 1966. At that time, while I was working on habituation, I met Pat Bateson. His background was in the behavioural sciences, mine in the neurosciences. After discovering our common interest in learning in general, and in imprinting in particular, we began to collaborate in experimental work. It proved to be a very fruitful collaboration.

The experiments on imprinting which are described in this book employed many different techniques. It is inevitable, therefore, that the ground covered is wide, ranging from molecular biology through neurophysiology, neuroanatomy, and endocrinology to behaviour. The range is further extended because of the need to refer to work on humans and other

primates. I have attempted throughout to define the technical terms used so that anyone who has attended undergraduate courses in some of these subjects should have little difficulty in following the arguments. There is extensive cross-referencing in the text, but this is unlikely to meet all requirements and I hope that the reader will make full use of the index when the cross-references fail to meet a particular need. Because relatively few people have easy familiarity with the structure of the avian brain an appendix has been written to provide an outline of the major features of this part of the central nervous system.

The analysis of the neural mechanisms of memory, and it almost goes without saying of imprinting, cannot be achieved by a single, 'critical' experiment. Once a change in the nervous system associated with training has been identified, it becomes necessary to determine to what aspect of behaviour that change relates. This determination is critical. The degree of confidence the experimenter has that, say, a biochemical change in the brain is related to learning, depends on the confidence he or she can place in the specificity of this relationship. Only if confidence is strong is it worth taking the analysis to a deeper level. This is how my colleagues and I have proceeded. And this book tells that story. I have not hesitated to speculate when attempting to interpret or to synthesize the results of the experimental analyses. I take the view that further advance in science depends critically on this synthetic operation. It is, however, essential to distinguish hard experimental evidence from speculation. I have made this distinction explicit in some chapters by including a separate section for discussion. This arrangement could not be adopted in all chapters, but I hope that I have made clear to the reader the points at which I have departed from the hard base of experimental evidence.

I am deeply indebted to my colleagues for their stimulating collaboration. It has been a pleasure to work with them. Besides Pat Bateson and Brian McCabe these include in alphabetical order Johan Bolhuis, Philip Bradley, Malcolm Brown, José Cipolla-Neto, Ceri Davies, Gordon Edge, Dick Hill, Ann Horn, Mark Johnson, Stephen Jones, Alex Milne, Jeremy Payne, Steven Rose, Hugh Rowell, Gerry Stechler, Zsuzsanna Wiesenfeld, and Mick Wright. The work benefited greatly from the skilled technical assistance of Barrie Fuller, Kevin Chapman, and Wendy Nix. I have received generous financial support from the Science and Engineering Research Council, the Medical Research Council, the Leverhulme Trust, the Wellcome Trust, the Nuffield Foundation, and the National Institutes of Health of the US Public Health Service, to all of whom I express my thanks.

I wish to acknowledge my indebtedness to those colleagues who read various parts of the typescript, Michael Berridge, Malcolm Brown, Malcolm Burrows, Ceri Davies, Nick Davies, Felicia Huppert, Mark Johnson, Brian McCabe, Jeremy Payne, Jean Thomas, Larry Weiskrantz, and Oliver Zangwill. Their critical comments have improved the book, but

its shortcomings are mine alone. John Rodford drew most of the illustrations and I am grateful to him. I also greatly appreciate the photographic work of Neal Maskell and Frances Pang.

I owe an especial debt of gratitude to Chris Percival who typed the whole of the manuscript and worked under difficult circumstances with dedication and good humour; to Robert Hinde who read all of the chapters, commented on them in detail and was a constant source of encouragement; and to my wife Priscilla Barrett who critically read all the chapters, drew the illustrations of the jungle-fowl and the cover design, and who for several years has had to put up with conversation which rarely ranged beyond memory, imprinting, and the brain.

*January 1985*                                                                                       G. H.
*Cambridge*

# Contents

# Contents

# 1
# Approaches to the analysis of the neural bases of memory

There is a long history of interest in the mechanisms by which information is stored in the brain. The explanations that have been offered have varied down the centuries, but with few exceptions they share a common theme. A particular experience or event leads to the formation or strengthening of particular pathways in the brain. Once formed or strengthened in this way the pathway is viewed as a 'trace' which 'represents' the particular experience or event. This trace may be linked to other traces so that the two representations become associated. If one trace becomes activated in some way, the other is likely to be activated as well. When this happens the events or experiences which the traces represent are 'brought to mind', or recalled. These ideas, although not stated in quite this form, can be traced to Descartes (1649; see Descartes 1961) and Locke (1690).

Whilst it was recognized that the retention of information over relatively short periods of time might be achieved by transitory activity, long-lasting memories were thought to involve strucural changes in the brain. Descartes (1649; see Descartes 1961) foreshadowed the idea of morphological change when he wrote that the traces were '... pores of the brain ... [which had] acquired a greater facility than the rest to be opened again ...' (Article 42). Freud (1895; see Freud 1966) envisaged the existence of barriers between neurones of a certain class. He suggested that the activation of these neurones led to an irreversible decline in the resistance of the barriers, and hence to the formation of a memory trace. For Cajal (1911) the critical change was the development of new connections, for Tanzi (1893) and for Hebb (1949) it was the strengthening of pre-existing connections. A number of more recent formulations are variants on these basic themes (see, for example, Milner 1957; Griffith 1966; Brindley 1967; Marr 1969).

Of course there has been some shift of ground and change of emphasis throughout this long period of speculation. For some writers the enduring trace was a dynamic one involving, for example, reverberating activity in chains of neurones. But for many the trace was structural. These 'structural' theories have included not only those referred to above, but also the following: changes in the 'supporting' cells of the brain, in the molecular organization of ribonucleic acid (RNA) in neurones, and in specific lipo-peptide patterns in nerve-cell membranes. This list is far from exhaustive.

1

Despite their diversity these views have two things in common: the neural changes were inferred from behaviour, and direct experimental evidence for these changes was at best tenuous and at worst non-existent (see, for example, Lashley 1950). The mismatch between the plethora of theories and the dearth of experimental evidence is not, however, remarkable. The great interest in the neural processes that underlie 'higher mental functions' inevitably generated speculation as to what these processes might be; and in so far as these speculations related to the real nervous system, they were necessarily constrained by the state of knowledge at the time that the views were expressed. If the deep understanding that derives from experimental evidence seems unduly to have lagged behind speculation, it is salutory to reflect that less than a hundred years have past since Waldeyer, drawing on the work of the great neuroanatomists of the nineteenth century, enunciated the 'neurone theory', that the terminal swellings of axons were first clearly described in 1897 by Held, and that, in the same year, the term 'synapse' was coined by Sherrington to refer to the junctions that these swellings make with the cell body (Foster 1897). At that time the most effective strategy for advancing knowledge of the actions of the nervous system in behaviour was to investigate physiological processes which are relatively stable over a wide range of experimental conditions. Certain spinal reflexes have some of these characteristics and by 1906 Sherrington, in a masterly analysis, had laid bare the basic mechanisms by which these reflexes are integrated within the spinal cord. Learned patterns of behaviour do not have these characteristics in any simple way. Behaviour is modified during the course of learning and so too, it may be inferred, is the state of the nervous system. Before any viable attempt could be made to understand the nature of this change of state, it was essential to know a great deal about the fundamental organizations of the system, and much of the work on the nervous system conducted in the present century was directed to this end. Viewed in this light, the discrepancy between the scarcity of experimental evidence and the abundance of theories about the neural mechanisms of learning and information storage becomes more comprehensible.

Certain benefits have accrued from the delay in investigating the neural bases of learning and memory. On the one hand, a multitude of powerful, highly sophisticated techniques have recently become available, and the application of these techniques to the study of the nervous system has radically changed our understanding of its structural and functional organization. On the other hand, there have been major advances in the ways in which behaviour is described, analysed, and modelled. All of these advances have conspired to deepen and refine understanding of behavioural processes. Here and there the interests of the behavioural scientist and the neuroscientist have converged. One result of this convergence has been that speculation about the ways that information is stored in the brain

is now being informed by experimental evidence. The first step in providing that evidence came from the study of habituation.

## 1.1 Memory and habituation

### 1.1.1 Some behavioural characteristics of habituation

Humphrey (1930, 1933) showed that the snail, *Helix albolabris,* retracted its antennae when the platform on which it was moving was vibrated once every few seconds. Gradually the withdrawal response diminished and ultimately the snail no longer appeared to be affected by the vibrations. This gradual waning of a response is a component of a behavioural change known as habituation, a phenomenon which is widespread in the animal kingdom (Humphrey 1933; Thorpe 1956). The properties of habituation have been investigated extensively in the laboratory by studying the 'orientation response' (Pavlov 1927; Sokolov 1960, 1963). This response is elicited by a novel stimulus and gradually wanes if the stimulus is delivered repeatedly. For example, if a sound is suddenly presented to a cat, it may direct its gaze to the source of the sound. If the sound is repeated a few times it gradually ceases to elicit a response from the cat. Habituation is, within limits, stimulus-specific, so that when the response to one stimulus has waned the same response may be elicited by a different stimulus. The response to the repeated stimulus may be restored if this stimulus is withheld for a time before being reapplied.

Habituation is a learning process which confers some obvious benefits. An animal ceases to respond to a repeated stimulus which is, and may continue to be, of no consequence to it (Humphrey 1933; Thorpe 1956). But the animal is still able to detect and respond to a novel stimulus which may signify danger, or food, or perhaps a mate. The memory that is implied by habituation is a curious one. It is a memory not to respond; if an habituated response returns the animal may, in some sense, be said to have 'forgotten'. The neural analysis of this memory would greatly be facilitated if changes in some aspect of neural function occurred which paralleled the behavioural changes of habituation. The first evidence for such neural changes came from an unlikely source, a series of studies of an animal with an incomplete nervous system.

### 1.1.2 Physiological analysis

If the skin of the back behind the shoulder of a 'spinal' dog is rubbed or tickled or tapped, a series of rhythmic scratching movements is evoked in which a foot is brought towards the site of stimulation. Sherrington (1906) found that when the mechanical stimulus to the skin was maintained, the individual beats of the 'scratch-reflex' usually became slower and their amplitude smaller. The scratch-reflex recovered if the stimulus was with-

drawn for a few seconds, though recovery was not always complete with such a short remission time. Within certain limits the response decrement was specific to the area of skin stimulated: when the reflex had waned it could easily be elicited by stimulating the skin a few centimetres away. The waning of the scratch-reflex has some features of habituation: the response decrement is relatively specific to the site at which the stimulus is applied, the decrement is not attributable to fatigue of the muscle and probably not attributable to a change in the receptors in the skin and the response recovers following a lapse of time.

Sherrington's experiments involved stimulating the skin or sensory nerves, and recording the movements of a hind-limb. Events taking place within the spinal cord were inferred, but not directly measured. Recordings of neural activity during habituation came some years later; and the activity first studied was of the brain, not of the spinal cord. One of the first approaches was to examine the changes which occur in the rhythms of the electroencephalogram when a novel stimulus is repeatedly applied. These investigations were first conducted on humans (Berger 1930; Knott and Henry 1941; Sokolov 1960), unanaesthetized rabbits (Ectors 1936), and cats (Rheinberger and Jasper 1937). The changes were studied in detail by Sharpless and Jasper (1956). In their experiments, unanaesthetized cats were allowed to fall asleep, as indicated by both behavioural and electro-encephalographic criteria. A 500 hertz tone was used to awaken the cat. When this happened the slow-waves, which characterized the electroen-cephalogram during sleep, were replaced by runs of low-voltage fast activity; that is, the electroencephalogram was 'activated' (Fig. 1.1(a)). As soon as the cat fell asleep and the electroencephalographic sleep pattern had returned, the tone was sounded again. In the course of twenty or thirty presentations of the tone, the stimuli were progressively less effec-tive in waking the cat and in activating the electroencephalogram (Fig. 1.1(b)). When the animal had ceased to be aroused by the repeated 500 hertz tone, another tone with a frequency of 100 hertz (Fig. 1.1(c)) or of 1000 hertz (Fig. 1.1(e)) was presented. The novel tones awakened the animal and activated the electroencephalogram. Once this cortical activa-tion response had become habituated to a repeated tone, it was often possible to elicit the response again, with the same tone, provided that 15 minutes or more were allowed to pass before presenting the stimulus again. The duration of this pause depended, in part, on the number of times the stimulus had previously been presented. In some cases it was necessary to interpose several days of rest between successive recordings for a response to be elicited once more by the repeated tone.

Although the changes in the electroencephalogram possessed many of the characteristics of stimulus-specific habituation, Sharpless and Jasper were of the view that these studies shed little light on the ultimate nature of the habituation process. Nevertheless, the work was important in at least two
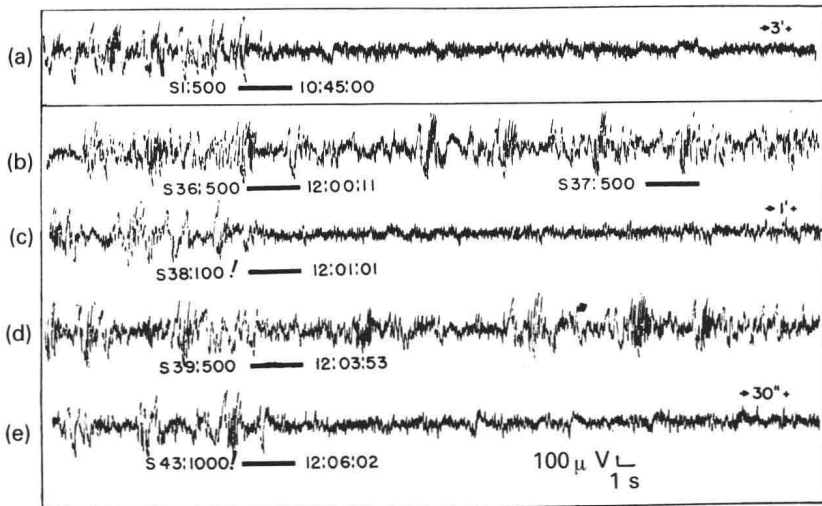
**Fig. 1.1.** Cortical electrograms from the suprasylvian gyrus of a normal cat showing typical habituation of the arousal reaction to a 500 Hz tone after about thirty trials. In the first tracing (a) the response to the first presentation of the 500 Hz tone is shown (S1: 500). The solid bar shows the duration of the stimulus followed by the time at which it was delivered, given in hours, minutes, and seconds (10:45:00). When the sound was presented the electroencephalogram was 'activated'. In the second tracing (b) are shown the thirty-sixth and thirty-seventh trials (S36 and S37). (c) A novel tone (!) of 100 Hz is presented in the thirty-eighth trial (S38: 100!) followed (d) by a repetition of the habituating tone (S39: 500) and then (e) another novel tone (S43: 1000!). The figures at the right above the electroencephalogram traces indicate the duration of the electroencephalographic activation in each trial. (After Sharpless and Jasper 1956).

respects. It established a link between a change in a physiological measure and a change in behaviour, and it drew attention to the process of habituation at a time when techniques were becoming available which would allow the process to be analysed at the cellular level with microelectrodes.

One such analysis involved recording the activity of single neurones, or units within the optic tectum (Horn and Hill 1964, 1966; Horn 1965; Masland *et al.* 1971; Cynader and Berman 1972), an important visual area in all vertebrates. In mammals the output neurones of the retina, the ganglion cells, send axons to the lateral geniculate body, or nucleus. Cells in this nucleus project to the visual cortex. Another target of ganglion cells is the optic tectum. The optic tectum contains two elevations, the superior colliculi. Single neurones in the optic tectum respond to visual stimuli if an image of the stimulus is focused on the appropriate part of the retina. This region of the retina is referred to as the receptive field of the tectal neurone. Some cells in the superior colliculus respond only to visual stimuli. Contrary, however, to the expectations of function implied by its name, the optic

tectum receives also input from other sensory pathways (Horn and Hill 1964, 1966; Jassik-Gershenfeld 1966; Gordon 1973; Dräger and Hubel 1975).

The response of some neurones recorded in the optic tectum of the anaesthetized rabbit wane when a stimulus is repeatedly presented (see Fig. 1.2). The unit whose responses are plotted in this figure discharged 'spontaneously', in the absence of any deliberate sensory stimulation, and responded with a burst of impulses, or 'spikes' when a puff of air was delivered to the shoulder. When this stimulus was repeated once every few seconds the response was stable for the first half-dozen or so presentations, and thereafter began to decline. After approximately twenty presentations, the stimulus failed to evoke a discharge that could be distinguished from the spontaneous background activity. For many such units the response recovered if the stimulus was withdrawn for some time before being pre-
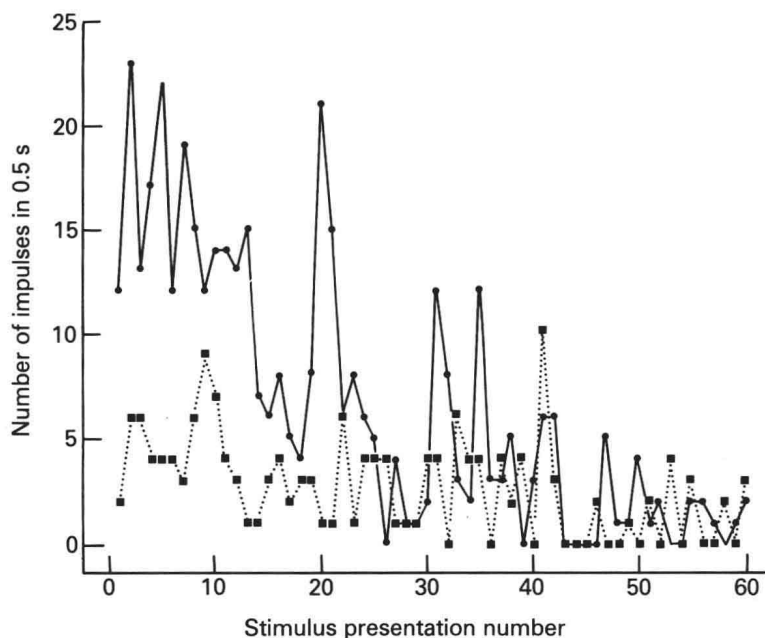


**Fig. 1.2.** Responses of a unit in the optic tectum of an anaesthetized rabbit. The unit responded to a puff of air blown on to the right shoulder. The puff of air had a duration of 2 s and was presented at 3-s intervals. The number of spontaneous impulses or spikes (squares) present in the 0.5 s before each stimulus was delivered, together with the number of evoked impulses (circles) in the first 0.5 s following application of the stimulus were plotted against stimulus presentation number (abscissa). The response gradually became less brisk as stimulation proceeded. After the twenty-third presentation, the number of impulses present during the time the stimulus was being applied was not consistently different from the number present during the prestimulus period. (After Horn and Hill 1966.)