Journée du 4 juillet 1969

Première séance

INFORMATION ET SYSTÈMES BIOLOGIQUES

II - ASPECTS THÉORIQUES

PRÉSIDENT S. L. SOBOLEV

---

* M. P. Schutzenberger's report has been presented orally. We have not received the manuscript.

De la Physique théorique à la Biologie, C.N.R.S., 1971.

NON-HOLOGRAPHIC ASSOCIATIVE MEMORY

D.J. WILLSHAW, O.P. BUNEMAN, H.C. LONGUET-HIGGINS

Department of Machine Intelligence and Perception,
University of Edinburgh

The remarkable properties of the hologram as an information store have led some people [1, 2] to wonder whether the memory may not work on holographic principles. There are, however, certain difficulties with this hypothesis if the holographic analogy is pressed too far; how could the brain Fourier-analyse the incoming signals with sufficient accuracy, and how could it improve on the rather feeble signal-to-noise ratio [3] of the reconstructed signals? Our purpose here is to show that the most desirable features of holography are manifested by another type of associative memory, which might well have been evolved by the brain. A mathematical investigation of this non-holographic memory shows that in optimal conditions it has a capacity which is not far from the maximum permitted by information theory.

Our point of departure is Gabor's observation [4, 5] that any physical system which can correlate (or for that matter convolve) pairs of patterns can mimic the performance of a Fourier holograph. Such a system, which could be set up in any school physics laboratory, is shown in Figure 1. The apparatus is designed for making "correlograms" between pairs of pinhole patterns, and then using the correlogram and one of the patterns for reconstructing its partner. One of the pinhole patterns is mounted at A, and the other at B. The distance between them equals f, the focal length of the lens L. A viewing screen is placed at C, at a distance f from the lens, and a diffuse light source is mounted behind A. The pattern of bright dots appearing at C is the correlogram between the pattern at A and the pattern at B. Formally, $C = \overline{A} * B$, where the asterisk stands for convolution and A is the result of rotating the pattern A through half a turn round the optical axis. If $\overline{A}$ and B were interchanged, the pattern at C would be $\overline{B} * A = A * \overline{B} = \overline{C}$, so that the correlogram would be inverted. This is clear enough if B is a pinhole, and shows that the order of the patterns is important.

To recover pattern A from pattern B we convert the correlogram into a pattern of pinholes in a black card and place the light source behind it, so that the light shines through C and B on to a viewing screen at A (Fig. 2). A pattern of spots now appears on the viewing screen. All the spots of the original pattern A are present, but a number of spurious spots as well. If the

pinholes were infinitesimal and there were no diffraction effects the reconstructed pattern would be $\overline{C} * B = A * \overline{B} * B$, just as in Fourier holography. If B were a random pattern, one could argue, $\overline{B} * B$ would approximate to a delta function at the origin, so that the reconstructed pattern would look like a slightly bespattered version of the original pattern A. How can we pick out the genuine spots from the others ?

To solve this problem let us simplify the set-up by removing the lens (Fig. 3). Suppose, for example, that A has two holes and B has three. Then the pattern C will consist of six bright spots (barring coincidences). When these spots are converted into pinholes and illuminated from the right, a total of 18 ($= 6 \times 3$) rays will emerge from B and impinge on the screen at A. But we shall not see eighteen spots on this screen, because six of the rays will converge, in sets of three, on to the two points of the original pattern. The other twelve rays will give rise to spurious spots, but (again barring coincidences) these spots will be fainter than the genuine ones. We can therefore expect to be able to pick out the wheat from the chaff with a detector with a threshold slightly less than three units of brightness.

This reasoning applies equally to the " correlograph ", with lens, illustrated in Figs. 1 and 2. So, having found how to get rid of the unwanted background in reconstructing A from B and C, we can now envisage the possibility of constructing multiple correlograms, comprising all the spots present in $C_1 = A_1 * B_1$ or in $C_2 = A_2 * B_2$, and so on. The presentation of B_1 should evoke A_1, presentation of B_2 should evoke A_2, and so on, up to the limite set by the information capacity of the system. But what is this limit ?

To answer this question let us evade the complicated (and basically irrelevant) issues raised by the finite wavelength of light, edge effects and so on, and pose the question in terms of a discrete, and slightly more abstract, model. We suppose A, B and C to be discrete spaces, each containing N points, a_1 to a_N, b_1 to b_N, and c_1 to c_N. The point-pair (a_i, b_j) is mapped on to the point c_k if $i - j = k$ or $k - N$. Conversely, the point-pair (c_k, b_j) is mapped

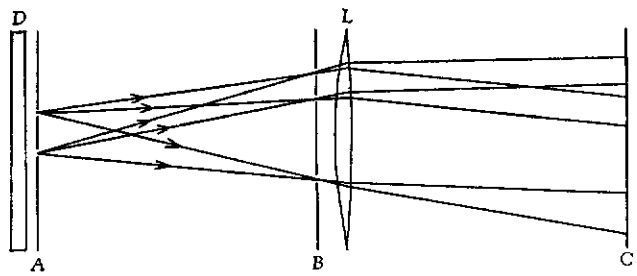

Fig. 1. Constructing a correlogram. D is a diffuse light source, L a lens and C the plane of the correlogram of A with B.

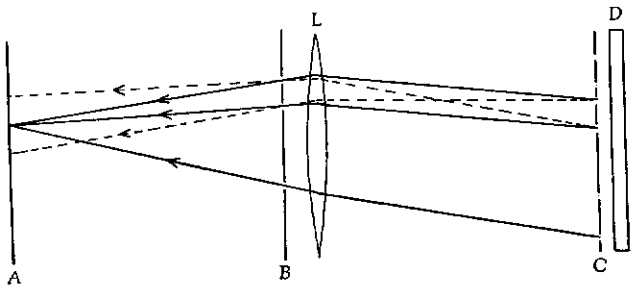

Fig. 2. Reconstructing a pattern. ——, Paths traversed in Fig. 1; - - -, paths not traversed in Fig. 1.

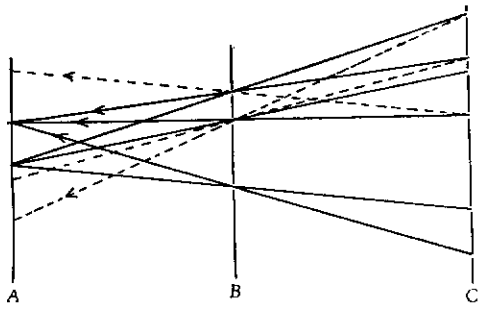

Fig. 3. Showing that original spots are generally brighter.

on to a_i if the same condition is met. Imagine now that we have R pairs of patterns which we wish to associate together, each pair consisting of M points selected from A and another M selected from B. The total number of point-pairs determined by all the pairs of patterns will be RM^2, and we may think of this number of " rays " striking C. If they impinge at random, the probability of any point c_k not being struck will be

$$\exp(-RM^2/N) = 1 - p, \text{ say}$$

The correlogram for the whole set of R pairs will then consist of the remaining pN points of C.

Now consider the reconstruction process. One of the B-patterns, comprising M of the points b_1 to b_N, is selected, and combined with the correlogram to produce pNM " *rays* " impinging on A. Each point of the original A-pattern will receive exactly M rays, so that we should set the threshold of our detector at M if we want to pick up all the original points. Now consider any one of the $N - M$ other points in A. It may receive a ray through any one of the M " holes " in B; the probability that it receives a ray through

a given hole is just p, for this is the chance that the point on C " behind " the hole belongs to the correlogram. The chance of an unwanted point reaching the threshold is thus p^M, and the probable number of spurious points of brightness M is consequently $(N - M)p^M$. If M is a fairly large number, this will be a sensitive function of p, and for given N and M the critical value of p above which spurious points begin to appear may be found from the relation

$$(N - M)\, p^M = 1$$

Alternatively, this may be viewed as a relation which sets a lower limit to the value of M for given values of N and p. A slightly safer estimate is given by

$$\text{(i)} \qquad Np^M = 1, \text{ or } M = -\log N/\log p$$

If M falls below this value, the reconstruction will be marred by spurious points.

Next we enquire about the amount of information stored in the memory when R pairs have been memorized and M satisfies the aforementioned condition for accurate retrieval. We can evoke any one of R A-patterns by presenting the appropriate B-pattern. There are $\binom{N}{M}$ possible A-patterns altogether, so the amount of information needed to store any one of them is $\log \binom{N}{M}$, which is roughly $M \log N$ natural units of information. The total amount of information stored is, therefore, approximately

$$\text{(ii)} \qquad I = RM \log N \text{ natural units}$$

But according to our original calculation of p

$$\text{(iii)} \qquad RM^2 = -N \log (1 - p)$$

and if we are working at the limit of accurate retrieval

$$\text{(iv)} \qquad M = -\log N/\log p \simeq \log_2 N \text{ (see below)}$$

It follows immediately that

$$\text{(v)} \qquad I = N \log p \log (1 - p)$$

As one might have anticipated, this expression has its maximum value when p is 0.5—when the correlogram occupies about half of C.

What is remarkable is the size of I_{max}.

$I_{max} = N (\log 2)^2$ natural units $= N \log 2$ bits. The maximum amount of information that could possibly be stored in C is N bits. So the correlograph, in this discrete realization, stores its information nearly ($\log 2 = 69$ per cent) as densely as a random access store with no associative capability.

As described, the discrete correlograph, like the holograph, will " recognize " displaced patterns. If an A-pattern $\{a_i\}$ and a B-pattern $\{b_j\}$ have

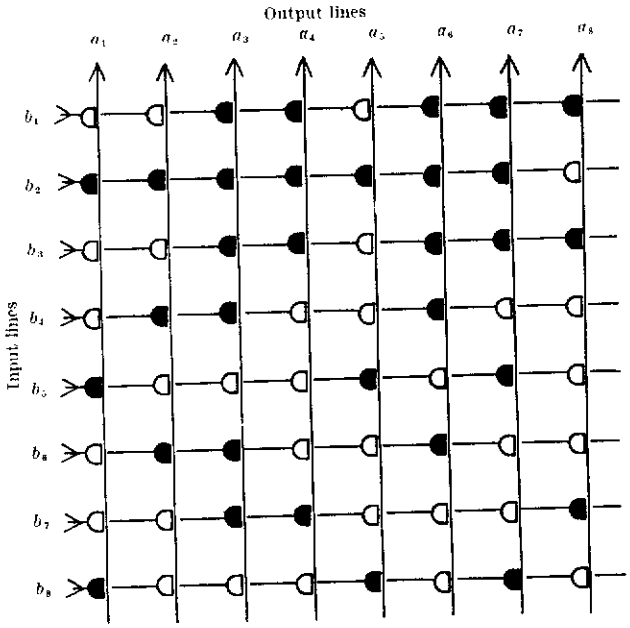

Fig. 4. An associative net.

been associated then presentation of the displaced *B-pattern* $\{b_{j+d}\}$ will evoke the displaced A-pattern $\{a_{i+d}\}$.

But the resemblance does not cease there. Just as in holography, the information to be stored is laid down (i) in parallel, (ii) non-locally and (iii) in such a way that it can survive local damage. In parallel, because each mapping $(a_i, b_j) \rightarrow c_k$ can be effected without reference to any other; the same applies to the reconstructive mappings $(c_k, b_j) \rightarrow a_i$. Non-locally, because the presence of a_i in an A-pattern is registered at M separate points on the correlogram, one for each point of the B-pattern. And robustly, because if the system is not stretched to its theoretical limit it can (as we shall show elsewhere) be used for the accurate reconstruction of A-patterns even when some of the correlogram is " ablated " and/or the B-patterns are inaccurately presented. But it can only be made secure against such contingencies by sacrificing storage capacity—as one would expect.

In our discussion of the process of reconstruction we had occasion to note that a point c_k might owe its presence on the correlogram to the joint occurrence of (a_i, b_j); but that if a pattern were presented containing the point b_{j+d}, the " ray " (c_k, b_{j+d}) would light up the point a_{i+d}, which might never have occurred in any A-pattern. It was this feature which underlay the ability of the system to recognize displaced patterns; but the same feature is a slight embarrassment when one comes to consider how a discrete correlograph, with the reconstructive facility, could be realized in neural tissue. We will not dwell

on this point, except to acknowledge that it was drawn to our attention by Dr F. H. C. Crick, to whom H. C. L.-H. is indebted for provocative comments. But it led us on to a further refinement of our model, in which a given point c_k is admitted to the correlogram only if the particular pair (a_i, b_j) occurs in one of the pairs of patterns, and not otherwise. On this assumption there might be as many as N^2 separate point-pairs to take into account, and a correspondingly large number of points in the space C.

In this form our associative memory model ceases to be a correlograph, having lost the ability to recognize displaced patterns, but its information capacity is now potentially far greater than before. To show this, we will adopt a rather different type of representation, in which the points of A become N_A parallel lines, and those of B become N_B parallel lines. The points of C are the $N_A N_B$ intersections between the lines a_i and the lines b_j.

In this network model, as before, a particular point of C is included in the active set if the pair of lines (a_i, b_j) which pass through it have been called into play in at least one association of an A-pattern with a B-pattern. Let us suppose that R pairs of patterns have been associated in this way, each pair comprising a selection of M_A lines from A and M_B lines from B. Then the chance that a given point of C has not been activated by the recording is

$$\text{(vi)} \qquad \exp(-RM_A M_B/N_C) = 1 - p, \text{ say}$$

where we have written N_C for $N_A N_B$. If B-patterns are being used to recall A-patterns, then there will be a mniimum value of M_B such that if the threshold on the A-lines is set at M_B (so as to detect all the genuine lines) spurious lines will begin to be detected as well. (The argument is just the same as that applied to the correlograph earlier on.) This minimum value of M_B is given by

$$\text{(vii)} \begin{cases} N_A p^{M_B} = 1 \\ \\ \text{or} \quad M_B = -\log N_A / \log p \simeq \log_2 N_A \end{cases}$$

Now the amount of information stored in the memory when R pairs of A-patterns have been memorized is roughly

$$\text{(viii)} \qquad I_A = RM_A \log N_A$$

But from our equation for $1 - p$

$$\text{(ix)} \qquad RM_A M_B = -N_C \log (1 - p)$$

therefore

$$\text{(x)} \qquad I_A = N_C \log p \log (1 - p)$$

showing that, as in the correlograph, the density with which the associative net stores information is 69 per cent of the theoretical maximum value. We may

note, in passing, that I_B, defined as $RM_B \log N_B$, is also equal to $N_C \log p \log (1 - p)$.

An associative network of this kind also operates (i) in parallel (ii) non-locally and (iii) in such a way that local damage or inaccuracy is not necessarily disastrous. We intend to go into the details of (iii) elsewhere. We now succumb to the temptation of indicating how such an associative memory might be realized in neural tissue though, as Brindley has pointed out [6], function need not determine structure uniquely.

The system we have in mind is represented diagrammatically in Fig. 4. The horizontal lines are axons of the N_B input neurones $b_1, b_2, \ldots$, while the vertical lines are dendrites of the N_A output neurones $a_1, a_2, \ldots,$. At the intersection of b_j with a_i is a modifiable synapse c_{ij}. This synapse is initially inactive, but becomes active after a coincidence in wich a_i and b_j are made to fire at the same time by some external stimulus. Such a coincidence is supposed to occur if an A-pattern containing a_i is presented in association with a B-pattern containing b_j. After the activation of c_{ij} (which we regard as a permanent effect) the firing of b_j will locally depolarize the membrane of a_i. The output neurone a_i is then supposed to fire if M_B or more input cells depolarize it simultaneously.

In Fig. 4 we indicate what the state of the network would be atfer it had learned to associate the following pairs of patterns :

(xi)

B-pattern	A-pattern
1,2,3	4,6,7
2,5,8	1,5,7
2,4,6	2,3,6
1,3,7	3,4,8

The synapses indicated by solid semicircles would be active, those indicated by open semicircles being still inactive. In this particular example, N_A and N_B are both 8, and M_A ($\simeq \log_2 N_B$) and M_B ($\simeq \log_1 N_A$) are both 3. R, the number of pairs of patterns associated, has been chosen so as to make p, the proportion of synapses active, close to 0.5; in fact p equals 0.5 exactly. These various numbers illustrate the system working near its maximum capacity. The reader may verify that every B-pattern except the first evokes the correct A-pattern at a threshold of 3; the only mistake the system makes is that when supplied with the B-pattern 1, 2, 3 it responds with an A-pattern 3, 4, 6,7 containing four elements.

To summarize, we have attempted to distil from holography the features which comment it as a model of associative memory, and have found that the performance of a holograph can be mimicked and actually improved on by discrete non-linear models, namely the correlograph and the associative net just describde. Quite possibly there is no system in the brain which corresponds

exactly to our hypothetical neural network; but we do attach importance to the principle on which it works and the quantitative relations which we have shown must hold if such a system is to perform, as it can, with high efficiency.

[1] P.J. Van Heerden, *App. Optics*, **2,** 393 (1963).

[2] K.H. Pribram, *Sci. Amer.*, **220,** 73 (1969).

[3] D. Willshaw and H.C. Longuet-Higgins, *Machine Intelligence* 4 (edit. by D. Michie) (Edinburgh University Press) (1969).

[4] D. Gabor, *Nature*, **217,** 1288 (1968).

[5] D. Gabor, *Nature*, **217,** 584 (1968).

[6] G.S. Brindley, *Proc. Roy. Soc.*, B, **168,** 361 (1967).

(Reprinted from *Nature*, **222,** 960 (1969)).

De la Physique théorique à la Biologie, C.N.R.S., 1971.

DO SPECIFIC BIOCHEMICAL CORRELATES TO LEARNING PROCESSES EXIST IN BRAIN CELLS ? *

HOLBER HYDÉN and PAUL W. LANGE

Institute of Neurobiology, Faculty of Medicine, University of Göteborg, Göteborg, Sweden

The present paper will deal mainly with the biochemical changes observed in neurons during three different learning experiments. First, however, some experiments showing differences in the protein composition of nerve and glia cells with be described. Of the learning experiments, the first is a case of instrumental learning in rats, in which changes in the synthesis of three acidic neuronal proteins and in the RNA base composition of neurons occurred; the arguments that these changes are specifically related to the training and that they are an expression of increased gene activation will be presented. In the next study, the protein changes observed in brain cells during simple sensory conditioning in rats will be described, and it will be argued that these are due to an increased level of attention rather than to learning, *per se*. Finally, some RNA data on neurons from monkeys performing a visual discrimination test will be reported.

Neuronal and glial proteins

Four years ago, Moore and collaborators [1, 2] described a brain-specific protein, called S100, because it is soluble in saturated ammonium sulfate. It is an acidic protein, has a molecular weight of around 20,000 constitutes 0.1 % of the brain proteins and moves close to the anodal front in electrophoresis. It develops after 12 days postnatally in the rat and is present only in nervous tissue. Thirty per 100 moles of its amino acids are acidic. It contains 30 % glutamic acid and no tryptophane. S100 can be further separated into at least 3 fractions, of which 2 have a high turnover and react immunologically with antiserum against S100 [3]. The S100 protein is not linked to carbohydrates (Fig. 1).

Hydén and McEwen [4] have shown by antiserum precipitation reactions supported by the Coons'technique [5] that S 100 is mainly a glial, protein which

* The main content of this paper has also been published in *Proceed. Sympos. on Biology of Memory*, Sept. 1, 1969, Tihany, Ed. G. Adám.

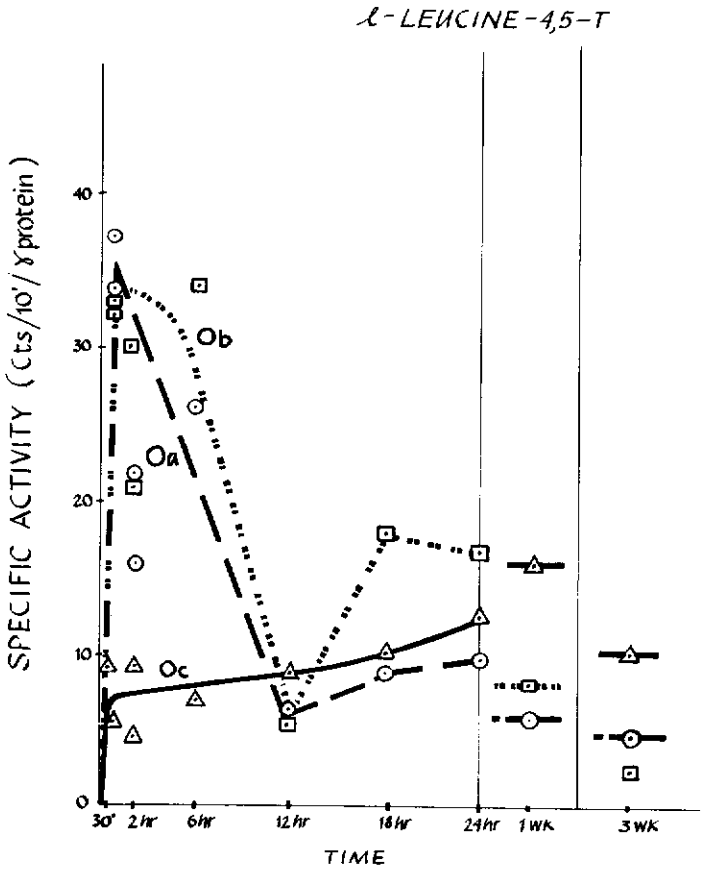

Fig. 1. Specific radioactivity of bands Oa, Ob, and Oc separated on 11,2 % polyacrylamide gels as a function of time between isotope injection and sacrifice. Radioactivity was determined after combustion of slices of the polyacrylamide gels by liquid scintillation counting. Isotope : 1-leucine-4,5 T.

in nerve cells is found only in the nuclei. Recently, Benda and collaborators [6] confirmed its presence in glia and showed its 10-fold growth in a clonal strain of glial tumors. Perez and Moore [7] have also presented evidence that S100 is mainly a glial protein. Moore and Perez [8, 9] have described another brain-specific protein which seems to be localized exclusively to the nerve cells and which has been named the 14-3-2 protein.

There is evidence for the existence of still other brain-specific soluble proteins. MacPherson [10] has described one in the β-globulin range, Kosinski [11] has described five soluble proteins, and Warecka and Bauer [12, 13] recently described an α-glycoprotein rich in neuraminic acid, which develops three months after birth in man and is probably derived from glia.

Bennett and Edelman [14] have purified and characterized still another acidic brain-specific protein.

An immunological study of Deiters' nucleus

We have examined the properties of antibodies prepared against neurons and glia [15] obtained from Deiters' nucleus in the continuing attempt to identify brain-specific proteins in them. The antigens in brain cells presumably number in the order of hundreds; Huneeus-Cox [16, 17], for instance, successfully prepared antisera against eleven antigens in preparations of squid axoplasm that did not include the external membranes. In our study antigen consisted of glial material dissected from the Deiters' nucleus of the rabbit by the free-hand technique previously described [18]. The dissection was carried out at 4 °C, with careful removal of capillaries and nerve cell bodies and processes; in this way, 3.2 mg of Deiters' nucleus glia was collected from 40 rabbits. The other antigen consisted of 1.3 gm of whole Deiters' nucleus, containing both neurons and glia, dissected from 100 rabbits.

Each of these antigens was homogenized and mixed with both complete and incomplete Freund's adjuvant. A group 6 six rhesus monkeys weighing 3-3.5 kg was injected intramuscularly with 0.6 ml of one or the other emulsion once a week for 4 weeks. None ever showed neurological symptoms, or signs of tuberculosis. The animals were bled after one week. On day 44 each monkey received a booster injection of 0.2 ml of its antigen emulsion precipitated with $Al_2(SO_4)_3$, and was bled one week later. These sera were tested on Ouchterlony

TABLE I

A. *Gel precipitation reactions* (+) *between anti nucleus Deiters' antiserum* (1:512) *and and a homogenate of nucleus Deiters.*

Antigen (μg/μl)	Reaction	Antigen	Reaction	Antigen	Reaction	Antigen	Reaction
8.20	-	0.80	+	0.30	+	0.05	-
4.10	-	0.50	+	0.20	+	0.02	-
2.10	-	0.40	+	0.10	+	0.01	-
1.00	-						

B. *Gel precipitation reactions* (+) *between anti Deiters' glia antiserum* (1:512) *and an antigen homogenate of Deiters' glia.*

Antigen (μg/μl)	Reaction	Antigen	Reaction
0.67	+	0.08	-
0.60	+	0.04	-
0.16	+	0.02	-

TABLE II

Gel precipitation reaction (+) between anti Deiter's glia antiserum and 0.9 g of protein extracted from nerve and glia cells dissected from Deiters' nucleus. Normal serum controls negative in each case.

Antiserum Dilution	Protein from	
	Nerve cell	Glia cell
1 : 64	–	+
1 : 128	–	+
1 : 256	–	+
1 : 512	–	+
1 : 1024	–	–
1 : 2048	–	–

plates against extracts of glia and of Deiters' nucleus, and their precipitation activities against sucrose-Triton X-100 extracts of both glia and of Deiters' nucleus material were also evaluated. In addition, the micromethod for double diffusion in one dimension in glass capillaries previously described [4] was used as an assay system, the Coons' [5] multiple layer indirect method for immunofluorescence applied to cryostat sections through the Deiters' nucleus, with evaluation of the specific fluorescence appearing in the nerve and glia cells was

TABLE III

Precipitation reaction between anti Deiters' glia antiserum (1:512) and homogenates homogenates of Deiters' nerve cells and corresponding volumes of glia. Neuronal protein estimates based on 12,000 µg of protein per cell. Glial protein per unit volume estimated at 50 % neuronal.

Deiters' neurons			Deiters' glia (same volume as nerve cells)	
Number of nerve cells	Calcul. protein in 10^{-6} g*	Precip.	Calcul. protein in 10^{-6} g*	Precip.
300	3.6	–	1.8	+ 2 ppt
150	1.8	–	0.9	+ 2 ppt
70	0.9	–	0.45	+ 2 ppt
60	0.72	–	0.36	+ 1 ppt
30	0.36	–	0.18	+ 1 ppt
15	0.18	–	0.09	–
6	0.09	–	0.045	–
3	0.045	–	0.022	–

also used. Some samples of the antisera were absorbed in two or three steps with sucrose-Triton X-100 homogenates of glia and of rabbit spleen, while others were twice absorbed with rabbit spleen and then absorbed with glia.

Tables I to IV summarize some results of these studies. Both the anti-Deiters' nucleus and the anti-glia sera formed well-defined precipitates with μg/μl amounts of their respective antigens (Table I). Table II shows that the antiglia serum formed precipitates with the glia but not with nerve cells obtained from Deiters' nucleus, and that no precipitates formed when normal rabbit serum was used against these antigens.

Table III shows the results of an antigen dilution study : homogenates of isolated nerve cells and of the same volumes of glial cells were tested against the antiglia antiserum in the dilution 1 : 512. Even when 300 isolated nerve cells were used no precipitation was obtained, but glial homogenates gave well-defined precipitates.

Precipitates were obtained when the anti-Deiters' glia antiserum was tested against glia dissected from other parts of the brain, e.g. from the hypoglossal nucleus and from the spinal cord and cerebral cortex, but none appeared against homogenates of motor neurons, pyramidal nerve cells of the hippocampus and granular cells from the cerebellum, all containing from 3.5 to 0.01 μg of protein per microliter.

Antiserum against the whole Deiters' nucleus gave two precipitation lines with both glia and nerve cells as antigens. However, when this antiserum was absorbed with glia or with spleen, only the nerve cell homogenates gave precipitates (Table IV).

TABLE IV

Number of precipitation lines after absorption of anti Deiters' nucleus antiserum. Antigen : homogenates from 120 *isolated nerve cells and corresponding amount of glia containing* 1.6 μg *used in each case. All dilutions tested* (1:2, 1:4, 1:8, 1:16) *gave the same result.*

	Protein form	
	Nerve cells	Glia cells
Unabsorbed	2	2
Absorbed with glia	1	0
Absorbed with spleen	1	0

The results with the fluorescence technique matched those obtained with the immunodiffusion technique as summarized in these Tables. Experiments were carried out according to the multiple layer method of Coons [5, 19].