

## Barlow's 1972 paper

[Barlow H B, 1972 "Single units and sensation: A neuron doctrine for perceptual psychology" *Perception* 1 371–394. Original paper reprinted in the appendix.]

---

### Author's update

#### **Single units and sensation: A neuron doctrine for perceptual psychology?**

The original article had two roughly equal parts, a review of experiments, followed by a set of five dogmas intended to be provocative theoretical propositions. Here I shall comment on experiments and theory separately.

#### **Experiments**

E D Adrian and his colleagues succeeded in recording from single peripheral sensory nerve fibres in the 1920s, and by 1972 the activity of neurons in most sensory modalities had been studied at several levels up to and including the cerebral cortex. In the original paper this work was reviewed, and an attempt was made to bring out the main implications of the experimental findings.

Re-reading this part more than 30 years later, I do not think there is much that is actually wrong or very misleading: there is an unfortunate misprint at the top of page 273 ("explains" instead of "explores") that may have concealed my intended meaning from a good many readers, and parts of the section on modifiability over-emphasised the instructive as opposed to the selective role of visual experience in determining the properties of adult neurons in visual cortex. In addition, it has now become clear, from the use of awake, behaving animals, that the maintained discharge rates and peak impulse frequencies in cortical neurons are a good deal higher than those commonly found in anaesthetised preparations, and one of the implications of this new evidence will be considered below.

There have been many other very important additions to our knowledge of the subject since then, notably in our understanding of extrastriate visual areas (eg Zeki 1978; Maunsell and Newsome 1987), and in correlating single-unit results from these areas with behavioural performance by recording from neurons in the brains of awake, behaving animals performing psychophysical tasks (eg Newsome et al 1989 and later papers in this series). Studies in audition, and especially olfaction, have caught up with those in vision, and shown interesting similarities and differences. But, whereas there really was a revolution in our understanding of the capacity of single neurons as information processors in the years before the article was written, I believe subsequent work has, so far, mainly substantiated these changes. On the other hand, the outlook for the future has changed radically; not only have old techniques for investigating neural activity in the brain been greatly improved and new methods developed, but advances in cell and molecular biology, genetics, and development, seem to guarantee further major revolutions.

#### **Theory**

The term 'dogma' caused some confusion. As with Francis Crick's 'central dogma' of molecular biology (Crick 1957), my dogmas were meant to be challenged by freethinking argument and experiment, and were certainly not intended to be accepted as blind articles of faith. They were defined in different ways in different parts of the original article but here I shall use single phrases that, I hope, emphasise their main points.

(1) *The nodes of the brain's neural network are especially important*

The first proposition said the brain is a network whose nodes are the neurons but, contrary to the general ethos prevailing over the last few decades, it emphasised the nodes rather than the network. I still believe this emphasis is appropriate. Useful network models will require nodes that perform the full range of information-processing operations that real neurons perform, but our ideas about what these operations include may still be very incomplete. Bray (1990, 1995) has argued that intracellular mechanisms have computational capacities comparable with those previously thought to require a whole PDP network; neurons may therefore perform tasks even beyond the capabilities outlined in Koch's (1997) forward-looking article. As he pointed out, it is crucial which comes first in the sequence of synaptic excitation and activity in the post-synaptic neuron, for this determines whether a Hebbian synapse is strengthened or weakened (Markram et al 1997; Bi and Poo 1998). A mechanism that is selective for different timings of synaptic input relative to post-synaptic activity is potentially much more powerful than the classic Hebbian synapse; such super-Hebbian synapses, in which it is suggested that the switchover from potentiation to depression can occur at different delays for different reinforcing stimuli, would enable a single cortical neuron to learn to distinguish between complicated spatio-temporal patterns, such as those characterising syllables or short tunes (Barlow 1996).

Another potentially interesting mechanism is calcium-induced calcium release (CICR), which propagates intracellularly along the endoplasmic reticulum. Intracellular  $\text{Ca}^{++}$  is important in the mechanism of the starburst amacrine cell dendrites (Euler et al 2002), the sequence-selective elements that are almost certainly responsible for directional selectivity in rabbit retinal ganglion cells (Vaney and Taylor 2002). These facts, together with the demonstrated role of diurnal clock mechanisms in the cortex (Hastings and Maywood 2000), make one wonder whether the prediction of sequence and timing (obviously involved in speech, music, and much else) may be a major pre-occupation of the whole cortex.

(2) *High-level representations have few active neurons*

I described the nodes of high-level representations as 'cardinal cells', which were half way to being Sherrington's 'pontifical cells', or 'grandmother neurons' (Sherrington 1941; see Lettvin 1995). The idea is surely the same as, or very similar to, what is now described by the much better term 'sparse coding' (Field 1994; Olshausen and Field 1997), but is it actually correct that there are few active neurons at higher levels?

There is a lengthy discussion of this point in the original paper, but, as noted already, the evidence has changed. The conclusions reached in "Single units" were almost all based on recordings from anaesthetised preparations, while later work on awake, behaving animals has shown that cortical neurons can respond in a much more lively fashion (see, for example, Rolls and Tovee 1995). Recent calculations (Lennie 1998, 2003; Attwell and Laughlin 2001) of the energy balance in the cortex and the energetic cost of impulses provide quite different evidence for the average firing rate being low, so it is probably correct that only a small proportion of cortical neurons are active at any one moment. Compare this with the retina, for example, where a high proportion are typically firing at substantial fractions of their peak rates. But I don't now think it is likely to be true that there is a progressive reduction as you go to higher levels within the cortex, and the changed evidence on this point seriously undermines the suggestion towards the end of the section that "a single visual scene can be represented quite completely by about 1000 such entities"—the entities being sparsely coding neurons. The passage gives a valid impression of how sparse coding and exploiting redundancy could form a useful representation, but I do not now think the actual figures suggested in that section are correct.

(3) *Trigger features are matched to redundancy*

This proposition asserted that the selectivity of sensory neurons was adapted to the redundancy of the environment. The topic has been the subject of four international conferences, so the dogma has probably been adequately discussed, though I'm not sure it has been proved to be true! Those interested can consult the published papers<sup>(1)</sup> from the conferences to form their own opinions.

(4) *High-level neurons directly and simply cause the elements of perception*

This claim is of course too simple and I now have difficulty even imagining anything scientific that would explain the personal and subjective aspects of perception. But I think there is one thing that it should be possible to do, both with the representation underlying one's subjective experience of sensory stimuli, and with the representation that is used for higher cognitive functions, and that is to determine how often the objects and events that it represents occur. Tony Gardner-Medwin and I (Gardner-Medwin and Barlow 2001) have compared how efficiently this can be done in direct and distributed representations, that is representations in which there are, and ones in which there are not, elements that are active when, and only when, the objects and events to be counted are present. From a distributed representation alone, one can count the occurrences of any member of those subsets of input events that activate each of the elements of the representation, but in general you cannot know how often a single input event, or a defined subset of input events, has occurred unless you resolve the distributed representation in the way your computer resolves ASCII coded characters. This is a drawback of distributed representations that is not often pointed out by those who vaunt their efficiency! One can make unbiased estimates of frequencies of single events and subsets, but unavoidable random errors are introduced that would give a considerable advantage to any competitor in a situation that required learning from experience.

There is a distinct advantage in having elements directly representing the events and objects that we need to learn about, and this may explain why higher level neurons in sensory systems, and perception itself, both appear to have such elements. But this is a long way from saying that sensory neurons "directly and simply" cause perception, which was too bold a claim.

(5) *High-impulse frequencies signal high certainty about trigger features*

The first neurons in sensory pathways signal the continuously graded physical quantities that strike the sensory surface, whereas those at higher levels encode a representation of the world around us that is used for controlling our actions, for learning, and for our subjective perceptions. Since Aristotle, the variables in this categorical representation have been held to be binary; they generally have the quality of duality—what is NOT (NOT a fish) must be a fish.

It is easy to convert a graded to a binary representation by introducing a steep nonlinearity and calling everything above it 1 and everything below it 0, but this destroys information about varying signal strengths, especially at signal strengths outside the threshold range. This would be very harmful, for it would effectively prevent the use of any categories that require the collection of many weak and variable signals to achieve reliable detection. So how and at what stage does the sensory system make the transition from graded to binary representation?

The fifth proposition suggested that this was done early, and that loss of information was minimised in the following way: neurons were thought to be elements that signal the certainty with which a particular hypothesis (usually the complement of its particular trigger feature) can be rejected on the current evidence (Barlow 1969).

<sup>(1)</sup> The first international conference on "Sensory Coding and the Natural Environment" was organised by Pamela Reinagel and held at Jiminy Peak, MA, in September 1997. Selected papers from the first three Gordon conferences on the same topic have been published in *Network: Computation in Neural Systems* 2001 12(3), 2003 14(3), and 2005 16(4).

This divides up the task differently from most models, for a particular neuron has to assess the degree of incompatibility of the current evidence with the hypothesis it represents, and to pass on a signal which, in my version, indicated the confidence with which the neuron's null hypothesis could be rejected. By setting a low threshold on incoming messages, one neuron could contribute messages frequently, but each of them would be of low significance by itself, while another neuron could set a high threshold and thus make a small number of contributions, each of high significance.

In the decades since 1972 belief networks incorporating many of the benefits I was striving for have been developed (Pearl 1988), and Bayesian models have become common (see, for example, Knill and Richards 1985). I believe that it is now widely recognised that, in networks like the brain, the quantities that need to be transmitted along the links are conditional probabilities of various kinds, not the quantities more closely related to stimulus intensities that physiologists traditionally consider.

Horace B Barlow

Physiological Laboratory, Cambridge CB2 3EG, UK; e-mail: hbb10@cam.ac.uk

### References

- Attwell D, Laughlin S B, 2001 "An energy budget for signaling in the grey matter of the brain" *Journal of Cerebral Blood Flow and Metabolism* **21** 1133–1145
- Barlow H B, 1969 "Pattern recognition and the responses of sensory neurones" *Annals of the New York Academy of Sciences* **156** 872–881
- Barlow H B, 1996 "Intraneuronal information processing, directional selectivity and memory for spatio-temporal sequences" *Network: Computation in Neural Systems* **7** 251–259
- Bi G, Poo M-M, 1998 "Synaptic modification in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and post-synaptic cell type" *Journal of Neuroscience* **2** 32–48
- Bray D, 1990 "Intracellular signalling as a parallel distributed process" *Journal of Theoretical Biology* **143** 215–231
- Bray D, 1995 "Protein molecules as computational elements in living cells" *Nature* **376** 307–312
- Crick F H C, 1958 "On protein synthesis", in *Symposia for the Society for Experimental Biology* Number 12 *The Biological Replication of Macromolecules* (Cambridge: Cambridge University Press) pp 138–162
- Euler T, Detwiler P B, Denk W, 2002 "Directionally selective calcium signals in dendrites of starburst amacrine cells" *Nature* **418** 845–852
- Field D J, 1994 "What is the goal of sensory coding?" *Neural Computation* **6** 559–601
- Gardner-Medwin A R, Barlow H B, 2001 "The limits of counting accuracy in distributed neural representations" *Neural Computation* **13** 477–504
- Hastings M, Maywood E S, 2000 "Circadian clocks in the mammalian brain" *BioEssays* **22** 23–31
- Knill D, Richards W (Eds), 1985 *Perception as Bayesian Inference* (Cambridge: Cambridge University Press)
- Koch C, 1997 "Computation and the single neuron" *Nature* **385** 207–210
- Lennie P, 1998 "Single units and visual cortical organization" *Perception* **27** 889–935
- Lennie P, 2003 "The cost of cortical computation" *Current Biology* **13** 403–407
- Lettvin J Y, 1995 "J Y Lettvin on grandmother cells", in *The Cognitive Neurosciences* Ed. M S Gazzaniga (Cambridge, MA: MIT Press) chapter 26 appendix, pp 434–435
- Markram H, Lübke J, Frotscher M, Sakmann B, 1997 "Regulation of synaptic efficacy by coincidence of post-synaptic APs and EPSPs" *Science* **275** 213–215
- Maunsell J H R, Newsome W T, 1987 "Visual processing in monkey extra-striate cortex" *Annual Review of Neuroscience* **10** 363–401
- Newsome W T, Britten K H, Movshon J A, 1989 "Neuronal correlates of a perceptual decision" *Nature* **341** 52–54
- Olshausen B A, Field D J, 1997 "Sparse coding with an overcomplete basis set: a strategy employed by V1?" *Vision Research* **37** 3311–3325
- Pearl J, 1988 *Probabilistic Reasoning in Intelligent Systems; Networks of Plausible Inference* (San Mateo, CA: Morgan-Kaufmann)
- Rolls E T, Tovee M J, 1995 "Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex" *Journal of Neurophysiology* **73** 713–726
- Sherrington C S, 1941 *Man on His Nature* (Cambridge: Cambridge University Press) p. 277
- Vaney D I, Taylor W R, 2002 "Directional selectivity in the retina" *Current Opinion in Neurobiology* **12** 405–410
- Zeki S, 1978 "Functional specialisation in the visual cortex of the rhesus monkey" *Nature* **274** 423–428

---

## Comments

### **From maverick to mainstream**

The 1972 Barlow paper has been with many of us for almost all our working life in perceptual research. It lays out a programme for how to conduct research that expects the investigator to work at several levels of explanation simultaneously (minimally, behavioural and neurophysiological). In doing so, it set the standards for a generation of studies, many of which Horace discusses in his retrospective.

The roots of the 1972 paper reach back more than 20 years from its publication date. The programme began with studies of activity in peripheral nerves and its relation with sensation. Later, as work progressed into the cortex, the ambition to explain more complex, cognitive aspects of perception grew proportionately.

One striking aspect of the push forward into the cortex has been the extent to which cortical neurons, like their peripheral inputs, convey statistically reliable signals in their firing patterns. In 1972, it was far from clear that this would be the result, but present-day studies have amply demonstrated this principle. Equally, the development of combined neuronal and behavioural studies, particularly in awake behaving monkeys, has enabled us to understand a great deal about the influence of processes such as attention, memory, and expectations on perceptual judgments.

So the programme of explaining perceptual function in terms of neuronal responses has become the standard method. Indeed the concept is now firmly part of mainstream early education—see the BBC children's TV show "Nina and the Neurons" (at <http://www.bbc.co.uk/cbeebies/nina/>), during which science is presented as a process in which evidence must be gathered by the five senses, each of which is assisted by its own single neuron!

Two issues that were underestimated in 1972 were cortical anatomy and the importance of intervention. In moving from peripheral to central nervous system, we forgot about the cortical circuitry. Perhaps this was a helpful simplification, but the cerebral cortex is not just a bag containing a jumble of the neurophysiologists' single units. The cortex has neuronal types, specific connections, interneurons, projection neurons, etc. Obviously, these different elements must have different roles and Horace's commentary encourages us to think more about the nodes of the brain's network.

Most of the experimental papers discussed in the 1972 paper measured neuronal correlates of perception. Methods for intervening causally in perceptual processing at the level of single neurons were limited in 1972 and remain so. Surgical lesions of brain areas often ask the right question, but cannot be reversed and are too crude to examine single neurons. Electrical microstimulation has been used to great effect in recent years, but its ascendancy is probably temporary. One may predict that in another 36 years the advent of new molecular techniques will provide a far superior set of approaches for intervening causally in the link between neuronal function and perceptual decisions.

Andrew J Parker

Department of Physiology, Anatomy and Genetics, University of Oxford, Sherrington Building, Oxford OX1 3PT, UK; e-mail: [andrew.parker@dpag.ox.ac.uk](mailto:andrew.parker@dpag.ox.ac.uk)

---

### **Seminal papers don't age**

It was a touching experience to re-read Horace Barlow's paper after more than 30 years, a paper that had been written when I started my scientific career. I was struck by two discoveries: first, how much had already been known at that time and is no longer cited, leaving one with the impression that most of the data that we build our hypotheses on are of much more recent origin. Second, how many of Horace's

---

intuitions withstood the test of the time and have since then turned into facts. Of the five dogmas listed on pages 380 and 381 the first three have remained undisputed. The remaining dogmas have either not been proven yet or need some extension.

The fourth dogma says that active high-level neurons directly and simply cause the elements of our perception. While it is certainly true that the responses of these neurons are a necessary prerequisite for perception, their mere activation is probably not a sufficient condition. Recent studies on the neuronal correlates of consciousness emphasise the fact that not only the activation of neurons, but the temporal coherence of widely distributed discharge patterns, is required to assure access to conscious experience. This fourth dogma ignores the possibility that a substantial amount of information may be encoded in the relations defined by the discharges of distributed assemblies of neurons. In this context, it was fascinating to read that Horace Barlow had been stimulated by Gerald Westheimer to write this article with the argument that something similar to the genetic code should be established for the brain. Geneticists are now realising that an enormous amount of the information stored in the genes is contained in the extremely complex relations established by the interactions among genes, ie by the syntax through which the words are bound into sentences. Similar principles of relational coding are now discussed for the brain but at the time Horace's paper was written, neither the geneticists nor the neurobiologists were ready to conceive of the possibility of dynamical distributed codes. Systems theory was essentially dominated by linear coding schemes derived from cybernetics and the von Neumann architecture of computers. Accordingly, it was held that all the information that a neuron could convey would have to be encoded in its discharge frequency and not in the temporal relations defined by the firing patterns of distributed neurons.

In line with this concept, and in analogy to the coding properties of peripheral sensory neurons, the fifth dogma stated that the frequency of neural impulse codes for certainty. This dogma still awaits experimental testing, as we still ignore the criteria according to which the brain can determine when a computation has come to an end, when it has obtained a result, and how valid this result is. If a result, for example the representation of a polymodal perceptual object, does not simply consist of the firing of a given set of cardinal cells but consists of a coordinated dynamical pattern of large numbers of neurons that may be distributed across different cortical areas, then discharge rates may be an ambiguous signature of validity. An additional problem is that recognition times are often so short that the information an individual neuron can convey must be contained in only a few discharges, which leaves not enough time for a frequency code. Another problem is that the representation of results may actually be associated with lower firing rates than the computations leading towards these results, and finally, it has since been established that discharge rates are modulated by numerous, context-dependent mechanisms, and above all, by attention. Thus, the modulation of discharge frequency appears to serve a number of different purposes.

Adopting the format of a review, I shall comment on several points in the sequence in which they occur in the text following the summary of the dogmata on page 381.

When contrasting the atomists and the globalists, Horace mentioned that interest in evoked potential and electroencephalography had faded because of the slow progress compared with single-unit recording and because the rationales for the use of these techniques were undermined. Hence, it is fascinating to see how the pendulum swings. Since the discovery of synchronous oscillations in the brain and the formulation of hypotheses that attribute significance to these phenomena, there has been a real renaissance of approaches exploiting the virtues of field potential, MEG and EEG recordings. Capitalising on the globality of these methods and the fact that they detect preferentially synchronised activity, much has been learned from these investigations about network properties supporting attentional mechanisms, subsystem integration, sensory motor coordination,

---

and short- and long-term memory. However, it is still true that deciphering the neural code will require cellular resolution.

When dealing with the number of selective categories that need to be represented Horace equates the complexity of the content represented by a cardinal cell with the description provided by a word and predicts that a very large number of words is required to describe all possible visual scenes. Interestingly, however, there is no mention of the cortical areas where all these cardinal cells should be localised. It is only mentioned towards the end of the paper that a large fraction of the neurons contained in the cerebral cortex should be influenced in one or the other way by visual stimuli. This prepares the notion of distributed coding but there is no mention of Hebbian assemblies, of the superposition problem that arises if these were defined only by increases of discharge rate nor is there any hint that assembly codes can effectively reduce the number of neurons required for the representation of different contents. Horace emphasises that economy is an important criterion for effective representational strategies and formulates the prediction that codes should be sparse—a prediction that is now confirmed. However, it appears as if the enormous success of single-cell analysis had drawn attention away from the possibility of assembly codes.

When discussing the fourth dogma, the origin of perception and the conundrum of consciousness, Horace suggests that we are not consciously aware of much that goes on in our brains, that not every cortical neuron's activity has its perceptual correlate. Here it is realised that something is missing in the coding scheme, but Horace remains vague and attributes the missing link to interpersonal communication. Surprisingly, no mention is made here of attention and the possibility of phenomenal awareness in animals.

In section 12, 'Criticisms and alternatives', Horace lists the many problems associated with the generation of invariant representations and predicts that these problems can be solved within hierarchically organised feed-forward networks, extrapolating from the invariant response properties of the neurons known at that time. This scenario is then contrasted with the alternative of combinatorial or holographic representations. Surprisingly, however, these alternatives are not considered in the context of coding efficiency or their possible superiority for the generation of invariant representations, but mainly in the context of robustness. And here it is argued that a combinatorial code would be more susceptible to the loss of some of the components of the distributed code than a code based on cardinal cells. Apparently, the robustness of associative networks and their ability to reconstitute full representations on the basis of partial information had not been known yet, even by the most competent scholars of neurophysiology.

The most profound impact that this re-reading exercise had on me has been triggered by the lecture on the last two pages. Here, Horace clearly advocates the necessity of a distributed relational code for the representation of visual scenes and actually proposes an inverted hierarchical organisation with divergence rather than convergence as the ruling principle. Although he does not refer to association or feedback connections and does not further specify how these relational graphs might be defined neuronally, he clearly does not diverge much from current concepts on distributed coding. Although one is inclined, when reading the text, to equate cardinal cells with the representation of individual perceptual objects, it is never said explicitly and it could well be that Horace admitted the possibility that objects, particularly if they are polymodal objects, are represented by assemblies of cardinal cells. The lesson to be learned from re-reading this text is that we have come to associate Horace Barlow's name with the caricature of a coding scheme that has been presented in a much more differentiated and thoughtful way than the references to this seminal paper make one believe. This shows impressively the fallacy of citing citations. Finally, there is this ingenious intuition at the end of the paper that sensory processes probably cannot be

---

understood unless one had models that also comprise the executive functions. Here, Horace's intuition has received full confirmation by the notion that perceiving and acting are intimately related. Had the techniques been available to record from more than one neuron at the time, Horace would certainly have used these techniques in order to test his intuitions about relational codes and distributed representations.

Wolf Singer

Max Planck Institute for Brain Research, Deutschordenstrasse 46, Frankfurt am Main, D 60528 Frankfurt am Main, Germany; e-mail: singer@mpih-frankfurt.mpg.de

---

### **Single units and sensation: still just as relevant today**

Horace Barlow's seminal 1972 paper has definitely stood the test of time. His clear and provocative statement of the link between single-unit activity and perception goes to the heart of the debate between 'local' and 'distributed' coding scheme that has continued to rage through the intervening decades (Thorpe 2002). Localist coding would certainly be good news for the neurophysiologist doing single-unit recording because the key information can be obtained by recording from individual cells in isolation. On the other hand, technical developments in the last couple of decades have made it possible to record from large numbers of cells simultaneously, and such studies have provided some support for distributed coding. One example is the evidence that synchrony across neurons could provide a mechanism for binding features together (Engel and Singer 2001). Likewise, I myself have been arguing that the relative order of firing across a population could be important for fast sensory processing (Thorpe et al 2001). However, even if there are stages in sensory processing where some form of distributed code is in use, this does not preclude the possibility that the top end of a sensory system contains large numbers of "high-level neurons, each of which corresponds to a pattern of external events of the order of complexity of the events symbolised by a word".

Evidence for the existence of such coding has recently come from some remarkable studies of single-unit activity in the human temporal lobe, where neurons have been described that show invariant responses to particular people (Quiñan Quiroga et al 2005). One neuron responded selectively to a varied set of photographs of actress 'Jennifer Aniston', and not to a wide range of other images including other blond women. Another responded to photographs of 'Halle Berry', another Hollywood actress, even when dressed as 'Catwoman', and even to the letter string 'Halle Berry'. This remarkable convergence from a wide range of sensory patterns that all 'mean' the same thing really does seem to argue for a form of localist coding. Furthermore, the data argue that neural representations at higher levels of sensory processing may indeed be quite sparse, as Barlow suggested (Waydo et al 2006).

It is important that this sort of selectivity was seen for objects that were highly familiar. We can certainly reject the idea that the brain comes equipped with neurons pre-tuned to respond to Jennifer Aniston. But if someone spends a large amount of time watching a particular TV series, is it so unreasonable to imagine that the brain could devote neural hardware to making information explicit at the single-cell level? Barlow's more recent work with Tony Gardner-Medwin has shown that localist coding has clear computational advantages (Gardner-Medwin and Barlow 2001). Furthermore, there are various biologically plausible mechanisms that can lead to the formation of neurons selective for frequently presented stimuli (see, for example, Masquelier and Thorpe 2007). So perhaps a consensual view would be to say that the brain simultaneously uses a range of different representations, and that this could be a key to its remarkable flexibility.

Simon J Thorpe

Centre de Recherche Cerveau et Cognition, Toulouse, France; e-mail: simon.thorpe@cerco.ups-tlse.fr

---

**References**

- Engel A K, Singer W, 2001 “Temporal binding and the neural correlates of sensory awareness” *Trends in Cognitive Sciences* **5** 16–25
- Gardner-Medwin A R, Barlow H B, 2001 “The limits of counting accuracy in distributed neural representations” *Neural Computation* **13** 477–504
- Masquelier T, Thorpe S J, 2007 “Unsupervised learning of visual features through spike timing dependent plasticity” *PLoS Computational Biology* **3** e31 doi:10.1371/journal.pcbi.0030031
- Quian Quiroga R, Reddy L, Kreiman G, Koch C, Fried I, 2005 “Invariant visual representation by single neurons in the human brain” *Nature* **435** 1102–1107
- Thorpe S, Delorme A, VanRullen R, 2001 “Spike-based strategies for rapid processing” *Neural Networks* **14** 715–725
- Thorpe S J, 2002 “Localized versus distributed representations”, in *The Handbook of Brain Theory and Neural Networks* Ed. M A Arbib, 2nd edition (Cambridge, MA: MIT Press) pp 643–645
- Waydo S, Kraskov A, Quian Quiroga R, Fried I, Koch C, 2006 “Sparse representation in the human medial temporal lobe” *Journal of Neuroscience* **26** 10232–10234

---

## Author's response

### What's next with single neurons?

First I want to emphasise that, if my article was a success, this was through the work of the skilled experimentalists who opened up the brain to single-unit recording in the decades from 1940 to 1970; without the facts they established, my speculations would not have had much value. Perhaps this is a moment to look quizzically at the results of work on single units since then, and to make some more guesses.

Far more investigators have been involved since 1972, they have been better trained and no less skilled, they have written many more papers, which have been more critically reviewed, and many brilliant new techniques have been developed and applied. Just to spell out some of these: we now use awake, behaving preparations instead of anaesthetised ones; we can do multi-unit recording; we have available much improved anatomical techniques and background knowledge; we have fMRI and other means of locating groups of active neurons; and we have computers to control our experiments and to perform sophisticated statistical assessments of the results. Many of the papers I originally cited would be rejected now, not just because the results would no longer be novel, but on the grounds of technical weaknesses that were excused as unavoidable in the bad old days. But it would be a brave advocate of new-age neuroscience who would seriously claim that, with all these advantages, our conceptual insight into the brain has advanced much more than it did in those early years. Partly this is because, in 1940, the field was new, and it was relatively easy to discover facts with genuine new conceptual implications—but have we also hung on to old concepts needlessly, or failed to develop and adopt new ones?

I think there is a bit of both. Today we tend to put too much trust in new technology, thinking that it will automatically reveal new facts and lead to new insights. We also hang on to old buzz-words, thinking they refer to well-established facts and ideas that should not be ditched. Here are two examples.

### Neural assemblies

One concept that I think had a negative effect on our understanding of the brain was the 'neural assembly', which was introduced by Donald Hebb in 1949, but is still praised by some, even today. Hebb's admirable aim in this wonderful book (Hebb 1949) was to advocate and build up explanations for psychological and behavioural facts in terms of properties of the brain's principal components—nerve cells or neurons. He was baffled by the apparently short duration of a neuron's influence on other neurons, for it was thought at the time to last no more than a few milliseconds or tens of milliseconds, whereas he knew that the psychological and behavioural effects of sensory messages often lasted for hundreds of milliseconds or much longer. He therefore developed Lorenté de No's concept of self re-exciting chains of neurons in which the activity might reverberate and last the required time—he called this a neural assembly.

We now know that neural activity can initiate much-longer-lasting processes in other, quite different ways: for example, it can switch on a gene—an effect that could have consequences lasting a lifetime. It can also initiate a wave of calcium-induced calcium release from the endoplasmic reticulum (Jaffe and Brown 1994); if this occurred at the tip of an apical dendrite in a large pyramidal cell, it would take 50 s for the wave to reach the cell body, 2 mm away. I've no doubt that if Donald Hebb had known of these persistent effects of neural activity he would have forgotten about his reverberating cell assemblies, and I think we should do this too, especially since, after nearly 60 years, I don't believe anyone has found convincing evidence for them.

---

### **Distributed representation**

Another term that seems liable to arouse false hopes as a clarifying principle is that of 'distributed representation'. My view is that it occurs, and the brain has to deal with it, but this is part of the problem and can rarely if ever be regarded as a solution. Let me make three remarks to illustrate this.

(1) It's important for an animal's brain to know about certain objects and events in its environment, but information about them naturally comes to be coded in a distributed form. This arises from the way these objects and events cause the physical signals that impinge on the animal's sensory receptors and cause the sensory signals that the brain must use to detect the presence of the objects and the occurrence of the events. Distributed representations arise naturally and are not in the first place specially devised by the brain.

(2) In engineering it's different: using 8 bits to code 255 keyboard characters for transmission to another place in the system provides economy, and nothing is lost (except the cost of resolving the distributed representation at the receiving end, if this is needed) because no noise is added. Coding messages from 100 million photoreceptors onto 1 million or so optic nerve fibres might be thought to involve similar principles, and there probably is considerable compression in the sense that the optic nerve fibres habitually operate nearer their limiting capacity than do the photoreceptors. But for the most part this does not involve newly devised combinatorial coding, as the ASCII code does. For instance we can distinguish hundreds of different hues in the spectrum, but this results from trichomacy imposed by the fact that there are only three different photopigments in the cones. The nonlinear operations that are later required to distinguish a huge number of different patterns of activity of the three primaries are certainly not carried out in the retina. We must, however, avoid being too dogmatic about this distinction, for directional selectivity requires nonlinear operations to discriminate the spatio-temporal correlations that are the hallmark of motion; this *does* occur in the retina, and there are other nonlinear mechanisms in the retina whose significance is not understood.

(3) Distributed representation complicates the tasks the brain has to perform, but they always represent some features of the input directly, a fact that is often neglected. To illustrate this, consider the ASCII code and ask this question: What can you say about the input if you look at just a single bit? The answer is obvious: if it's a 1, the input must belong to the subset of all possible inputs for which that bit is active, and if it's 0 then it does not belong to that subset. In the ASCII code itself, this information is very rarely of any use whatever, but this need not be the case. If one arranged all the ASCII characters in a list in order of their frequency of occurrence, one would find a large group at the bottom that occurred only very rarely. Now devise a new code in which one bit is active for a large subset of such very rare characters, but not active for any others. There is a great deal of latitude in the way bit patterns are assigned to characters, so this set of assignments is unlikely to make it impossible to devise an unambiguous code for all the other inputs, and assuming this is the case one now has a code in which a single bit acts as a kind of alarm signal: if it's active, something unusual has happened, and if it's repeatedly active something very unusual has happened. That's a kind of 'novelty detector' that might be useful for a brain.

It seems likely that the brain will ensure that the subsets of its inputs that are directly represented in this way are worthy of the advantages conferred by direct representation.

### **The principle of local computation**

If cell assemblies and distributed representations have failed to give the insights expected of them, are there others that could be more helpful? Tony Gardner-Medwin introduced me to one when we were working on the problem just mentioned above.

---

This is the '*Principle of Local Computation*', which would say "*If the brain performs a computation, such as combining a number of input variables to form a new output variable, then all the variables must be represented by biophysical quantities that are present at a single locus in the brain*" (Gardner-Medwin and Barlow 2001). It was suggested (Barlow and Gardner-Medwin 2000) as an additional argument for localist representation in our comment on Page's BBS article on that subject, and I think it expresses an intuition that many share, namely that the nonlinear interactions that are needed for most computations cannot occur unless the interacting quantities are all brought together at the same place.

It is this intuition that makes me worry about Wolf Singer's oscillations (Singer 2007; see also this volume). If I understand correctly, he suggests that when two separate regions of the brain oscillate together, they are somehow experienced as a unity: but what mechanism decides that they are oscillating together, and where is it located? It's easy enough for the experimenter to decide they are synchronous, but for the brain to experience unity, surely the brain itself must decide about the synchrony: Where is this supposed to be done? Without a suggested answer, I think the hypothesis is fatally incomplete.

Nevertheless, Wolf's idea has focused attention on a crucial problem, one that Simon Thorpe (this volume) also raises, that of determining the sequence and relative timing of events in the brain. There is ample evidence that the timing of visual events can be done with millisecond accuracy (see Westheimer and McKee 1975; Burr 1979; Greene 2006) and, like Wolf and Thorpe, I do not see how this would be possible by using neural codes that average the spike rate over tens or hundreds of milliseconds. And, as well as fast events, we also need to question how slow ones, such as those involved in directional selectivity or diurnal rhythms (see for example Hastings and Maywood 2000) are timed. I think the answers to many of these problems are more likely to be found in cell biology than in neural-network theory (see Bray 1995). There is a lot more to discover about what single neurons can do and how they do it.

## Conclusions

If I were to update my single-units paper, I might include a section on the importance of throwing out unnecessary old ideas, though I'm afraid I would not include any of my original five dogmas among the discards (or not yet, anyway). In fact I would be tempted to add more, such as the principles of exploiting redundancy (Barlow 2001), sparse coding (Olshausen and Field 1996), and local computation (Barlow and Gardner-Medwin 2000). I would also emphasise the need for finding both how precise timing and sequence are coded, and for clarifying the role of intracellular mechanisms, especially in timing slower events.

But I would be very optimistic, for, after quite a long period during which it was mainly the techniques that advanced, I believe there are now healthy signs that we are on the track of facts and concepts that will revolutionise our understanding of cortical function. We may be on the way to formulating models of neural interactions (eg Rust et al 2006) that, when performed serially, might explain the mechanisms and logic underlying the extraordinary psychological facilities, such as for facial discrimination (Tsao 2006) and reading text, that most of us have, together with the more exotic discriminatory skills of the bird watcher or car buff. When this is achieved, the results I reviewed in 1972 will be seen to be a step in the right direction, but only a small one. It will also mean that the physiological explanation of psychological phenomena has advanced much further than I ever expected it would when I started trying to find such explanations 60 years ago, and possibly Hebb himself would also be surprised!

---

**References**

- Barlow H B, 2001 "Redundancy reduction revisited" *Network: Computation in Neural Systems* **12** 241–253
- Barlow H B, Gardner-Medwin A R, 2000 "Localist representation can improve efficiency for detection and counting. (BBS Commentary on article on Connectionist Modelling by Mike Page)" *Behavioral and Brain Sciences* **23** 467–468
- Bray D, 1995 "Protein molecules as computational elements in living cells" *Nature* **376** 307–312
- Burr D C, 1979 "Acuity for apparent vernier offset" *Vision Research* **19** 835–837
- Gardner-Medwin A R, Barlow H B, 2001 "The limits of counting accuracy in distributed neural representations" *Neural Computation* **13** 477–504
- Greene E, 2006 "Simultaneity in the millisecond range as a requirement for effective shape recognition" *Behavioral and Brain Functions* **2** 38
- Hastings M, Maywood E S, 2000 "Circadian clocks in the mammalian brain" *BioEssays* **22** 23–31
- Hebb D O, 1949 *The Organisation of Behavior* (New York: John Wiley)
- Jaffe D B, Brown T H, 1994 "Metabotropic glutamate receptor activation induces calcium waves within hippocampal dendrites" *Journal of Neurophysiology* **72** 471–474
- Rust N C, Mante V, Simoncelli E P, Movshon J A, 2006 "How MT cells analyze the motion of visual patterns" *Nature Neuroscience* **9** 1421–1431
- Olshausen B A, Field D J, 1996 "Emergence of simple-cell receptive-field properties by learning a sparse code for natural images" *Nature* **381** 607–609
- Singer W, 2007 "Large-scale temporal coordination of cortical activity as a prerequisite for conscious experience", in *The Blackwell Companion to Consciousness* Eds M Velmans, S Schneider (Oxford: Blackwell) pp 605–615
- Tsao D, 2006 "A dedicated system for processing faces" *Science* **314** 72–73
- Westheimer G, McKee S P, 1975 "Visual acuity in the presence of retinal-image motion" *Journal of the Optical Society of America* **65** 847–850

ISSN 0301-0066 (print)

ISSN 1468-4233 (electronic)

# PERCEPTION

VOLUME 38 2009

[www.perceptionweb.com](http://www.perceptionweb.com)

**Conditions of use.** This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.