AMAZING PAPERS IN NEUROSCIENCE
Sparse Neural Representation of Odor Predicts Learning

James McManus*‡

** School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK, KY16 9JP; ‡ Present address: Department of Biomedical Science, The University of Sheffield, Sheffield , UK, S10 2TN.*

The way in which neurons encode information remains a hotly debated topic in neuroscience. Lin and colleagues in a 2014 article in the journal *Nature Neuroscience* demonstrate how sparse coding in the olfactory learning and memory center of *Drosophila* can influence learning behavior. Coding sparsity is the idea that only a small number of neurons in a network represent any given stimulus. Using neurogenetics, computational neuroscience, and cognitive approaches, they outline the discovery of an inhibitory feedback circuit responsible for differentiating the neuronal response to different odors. Manipulating this feedback circuit, they demonstrate how the sparseness in neural networks (how easily neurons are activated) can correspond to the ability to learn a sensory discrimination more easily. From a research perspective, this paper was important as it was the first causal demonstration of the role of sparseness in learning. From a teaching point of view, this paper is valuable because it is a simple but effective introduction to artificial neural network theory, where both the abstract theory and the importance of its application is apparent to those without a mathematical or computational background.

Key words: cognitive neuroscience; insect cognition; sparse coding; computational neuroscience

In 2014, Lin and colleagues published an article that demonstrates how a concept from artificial neural networks, sparse coding, applies to empirical neuroscience research. Given the increasing crosstalk between those studying artificial and biological neural networks, it is important to introduce neural network theory to neuroscience students in a way that is neither superficial nor esoteric. The present article walks through how Lin and colleagues achieved this balance.

Coding sparsity is a concept that was first used to describe neural representation in the 60s by David Marr and James Albus (Marr, 1969; Albus, 1971). While working on the cerebellum they observed that the number of neurons active at any given time was surprisingly small, this is referred to as a sparse code. Coding sparsity describes how many neurons are active in response to stimuli, and how many stimuli neurons are activated by. The sparsest networks are local codes in which single neurons respond to stimuli (Willmore and Tolhurst, 2001). On the other side of the spectrum lie dense codes, where many neurons are active during stimulus presentation, and many of the same neurons respond to different stimuli (Foldiak, 2002). A happy medium between these two extremes are sparse codes, where a small number of neurons are active in response to stimuli, and only a small number of stimuli activate any neuron.

Initially it was not clear that sparse coding had any special significance until it was implemented in theoretical models of learning in neural networks. These early 'perceptrons' were loose models of neurons connected by adjustable weights (see Anderson, 1995 for an introductory textbook). By adjusting these weights, perceptrons can learn simple patterns. However, as the number of patterns 'taught' to a perceptron increases, the time taken to teach the pattern increases exponentially (Albus, 1971).

Albus noted that pattern recognition in perceptrons could be made more powerful by restricting how many of the 'neurons' were allowed to be active at any one time. The introduction of a sparse code into the perceptrons resulted in faster learning of patterns (Albus, 1971) as well as an increased number of patterns that could be stored (Tsodyks and Feigel'man, 1988).

Since the advantage of sparse coding was theoretically proposed, other sparse codes have been reported for sensory areas (Crochet et al., 2011; Barth and Poulet, 2012) and in the motor cortex (Beloozerova et al., 2003). However, the paper by Lin and colleagues (2014) represents the first time a causal relationship has been drawn between sparse coding and learning *in vivo*. By using *Drosophila,* a less complex and more genetically tractable organism which nonetheless exhibits learning behavior, they were able to circumvent many of the difficulties linking neural representation to behavior in more complex organisms.

Olfactory information is processed in the mushroom bodies, two structures on each side of the *Drosophila* brain. The mushroom bodies (MB) contain Kenyon cells (KCs), which respond to olfactory information. Kenyon cells exhibit a sparse code, meaning that only a small percentage of them respond to a given odor (Turner et al., 2008).

Lin and colleagues hypothesized that KCs might contribute to their own sparse coding via feedback inhibition (where KC activity results in KC inhibition). This was based on findings from other sparse-coding systems which have suggested that recurrent feedback causes sparse coding (Papadopoulou et al., 2011), and the finding that GABA is required for sparse coding in the *Drosophila* MB (Lei et al., 2013). To test this, they expressed temperature-sensitive shibire (shi^{ts1}) in KCs. Shi^{ts1} is a mutation of the dynamin protein which is normally involved in the endocytosis of synaptic vesicles (Chen et al., 2002). At temperatures above

32°C, the mutated dynamin stops functioning, blocking endocytosis, and leads to a loss of synaptic transmission by neurotransmitter depletion (Poodry and Edgar, 1979; Koenig et al., 1989), without directly affecting the intrinsic excitability of the cell. At the same time, they expressed the calcium ion indicator, GCaMP3 in KCs to measure excitation in the Kenyon cells. They found that inhibiting KC transmission by raising the temperature reduced the overall activity of KCs compared to a control group without the shi^{ts1} gene. This suggested that the KCs inhibit their own activity.

One candidate for inhibition was via a pair of neurons – one for each mushroom body – called anterior paired lateral neurons (APL). APL was chosen because: a) an analogue neuron performing the same role was discovered in locusts (Papadopoulou et al., 2011); and b) Only when shi^{ts1} was expressed in all lobes of the MB did inhibition occur. This suggests that the output from the KCs is integrated into a single signal. Candidate neurons would be predicted to widely innervate a large proportion of the KC population, as APL does.

To test the effect of KC activity on APL neurons, Lin and colleagues expressed the TRPA1 gene in KCs. This gene encodes a cation channel that opens at warm temperatures, causing excitation. Again, they expressed GCaMP3 to measure APL activity and found that inducing KC excitation caused APL excitation. Additionally, expressing shi^{ts1} in the APL neurons increased the Kenyon cell response, suggesting a negative feedback circuit where KCs excite APL and APL inhibits KCs (Figure 1).

If many of the same neurons were activated by odors (i.e., the code is not sparse) activity patterns between odors would be similar. On the other hand, if the code is sparse, the activity pattern should be distinct across different odors. When APL was inhibited, the activity pattern became more similar among different odors. In other words, inhibiting APL reduced coding sparsity.

At this point, they had a way to manipulate the sparsity of neural encoding in an awake, behaving animal, providing the ability to test the hypothesis that computational theory had predicted: that sparse codes lead to enhanced learning.

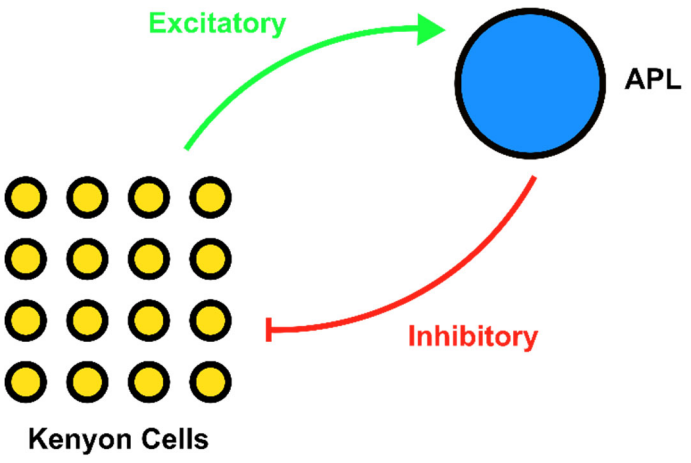

Figure 1. The results suggest that excitatory KC output is integrated into a single negative feedback signal that acts to make the KC code sparse.

They hypothesized that inhibiting APL neurons (using the same shi^{ts1} as before) would impair learning for similar odors, because the odor representations would interfere with each other but would have less of an effect for dissimilar odors or on odor coding in general. Dissimilar odors are represented by distinct sets of neurons, and so would enjoy less of a distinguishing effect from APL. While similar odors, without the sparse-inducing influence of APL, would have an overlapping representation, preventing differentiation. To test this, they presented flies with an odor shortly before giving them a small electric shock. They put flies into a T-maze with one odor on one side, and another odor on the other to see if the flies would avoid the odor that is associated with shock. When flies were given the choice between an odor that was previously paired with shock, and another dissimilar odor, inhibiting APL had no effect. However, when the two odors were similar, flies for whom the APL neurons were inhibited were impaired in their ability to avoid the odor that had not been paired with shock compared to controls. This implies that the feedback inhibition - and by extension the sparser coding - resulted in an increased ability to learn to discriminate between similar sensory stimuli.

VALUE

This paper was the first to draw a casual link between sparsity of neural coding and learning. It also demonstrates the application of findings from artificial networks to biological ones. Transplanting ideas and concepts from artificial neural networks has been a fertile ground for current neuroscientific insight (e.g., Whittington and Bogacz, 2019). It is important that students are exposed to the value of this kind of work, and also to the computational mindset in neuroscience. There are many different applications of artificial neural network theory to neuroscience, but the current paper represents a uniquely relevant case in that: a) the underlying theory is based on perceptron models, which are intuitively understandable to a non-expert and do not require a strong grasp of mathematics; and b) it is one of only a small number of papers that clearly and causally tests the predictions of artificial network theory, as opposed to an interpretation of correlational data. In a classroom, discussions could take place around whether the simple perceptron is an appropriate simplification of a biological neural network, as well as what this finding in the *Drosophila* odor system means for the sparse codes found elsewhere in more complex organisms (Foldiak, 2002; Quiroga et al., 2005; Poo and Isaacson, 2009).

Furthermore, this paper highlights the usefulness of genetically tractable, model organisms like *Drosophila*. Though likely familiar to students taking a more biological neuroscientific approach, the use of *Drosophila* is less common in cognitive neuroscience. Indeed, cognitive approaches often neglect the significance of seemingly simple organisms which nonetheless are capable of demonstrating relatively complex behavior. This paper stands as an elegant counterpoint to this way of thinking.

AUDIENCE

The paper, along with some general background on sparse

coding would present a challenging yet accessible paper for a third or fourth-year undergraduate. Featuring work from computational, genetic, and cognitive neuroscience, this is a truly interdisciplinary paper that could appeal to students interested in any or all of those areas.

REFERENCES

Albus JS (1971) A theory of cerebellar function. Math Biosci 10:25–61.

Anderson JA (1995) An Introduction to Neural Networks. MIT Press. Cambridge, MA.

Barth AL, Poulet JFA (2012) Experimental evidence for sparse firing in the neocortex. Trends Neurosci 35:345–355.

Beloozerova IN, Sirota MG, Swadlow HA (2003) Activity of different classes of neurons of the motor cortex during locomotion. J Neurosci Off J Soc Neurosci 23:1087–1097.

Chen M-L, Green D, Liu L, Lam YC, Mukai L, Rao S, Ramagiri S, Krishnan KS, Engel JE, Lin JJ-C, Wu C-F (2002) Unique biochemical and behavioral alterations in Drosophila shibire(ts1) mutants imply a conformational state affecting dynamin subcellular distribution and synaptic vesicle cycling. J Neurobiol 53:319–329.

Crochet S, Poulet JFA, Kremer Y, Petersen CCH (2011) Synaptic mechanisms underlying sparse coding of active touch. Neuron 69:1160–1175.

Foldiak P (2002) Sparse coding in the primate cortex. In: The Handbook of Brain Theory and Neural Networks (Arbib, MA ed) pp 1064-1068. 2nd Edition MIT Press, Cambridge.

Koenig JH, Kosaka T, Ikeda K (1989) The relationship between the number of synaptic vesicles and the amount of transmitter released. J Neurosci 9:1937–1942.

Lei Z, Chen K, Li H, Liu H, Guo A (2013) The GABA system regulates the sparse coding of odors in the mushroom bodies of Drosophila. Biochem Biophys Res Commun 436:35–40.

Lin AC, Bygrave AM, de Calignon A, Lee T, Miesenböck G (2014) Sparse, decorrelated odor coding in the mushroom body enhances learned odor discrimination. Nat Neurosci 17:559–568.

Marr D (1969) A theory of cerebellar cortex. J Physiol 202(2):437–470.1.

Papadopoulou M, Cassenaer S, Nowotny T, Laurent G (2011) Normalization for Sparse Encoding of Odors by a Wide-Field Interneuron. Science 332:721–725.

Poo C, Isaacson JS (2009) Odor representations in olfactory cortex: “sparse” coding, global inhibition and oscillations. Neuron 62:850–861.

Poodry CA, Edgar L (1979) Reversible alteration in the neuromuscular junctions of Drosophila melanogaster bearing a temperature-sensitive mutation, shibire. J Cell Biol 81:520–527.

Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. Nature 435:1102–1107.

Tsodyks MV, Feigel'man MV (1988) The Enhanced Storage Capacity in Neural Networks with Low Activity Level. Europhys Lett EPL 6:101–105.

Turner GC, Bazhenov M, Laurent G (2008) Olfactory Representations by Drosophila Mushroom Body Neurons. J Neurophysiol 99:734–746.

Whittington JCR, Bogacz R (2019) Theories of Error Back-Propagation in the Brain. Trends Cogn Sci 23:235–250.

Willmore B, Tolhurst DJ (2001) Characterizing the sparseness of neural codes. Netw Bristol Engl 12:255–270.

Received April 4, 2020; revised April 24, 2020; accepted May 24, 2020.

The author would like to thank Dr. Stefan Pulver and the MRes neuroscience course at St. Andrews for the advice and encouragement that contributed to this piece of work.

Please address correspondence to James McManus, Department of Biomedical Science, The University of Sheffield, Western Bank, Sheffield, United Kingdom S10 2TN. Email: james.alexander.mcmanus@gmail.com