Journal of Modeling and Optimization 2020;12(1):1-11
https://doi.org/10.32732/jmo.2020.12.1.1

Artificial Neural Network Simulations of Human Learning Suggest the Presence of Metastable Attractors in Visual Memory

Philippe Chassy[1], Frederic Surre[2]

1. Mathematical Psychology laboratory, University of Liverpool, UK
2. School of Mathematics, Computer Science & Engineering, City University of London, UK
E-mail: philippe.chassy@liverpool.ac.uk

Received: 29 September 2019; Accepted: 27 October 2019; Available online: 15 January 2020

Abstract: The attractor hypothesis states that knowledge is encoded as topologically-defined, stable configurations of connected cell assemblies. Irrespective to its original state, a network encoding new information will thus self-organize to reach the necessary stable state. To investigate memory structure, a multimodular neural network architecture, termed Magnitron, has been developed. Magnitron is a biologically-inspired cognitive architecture that simulates digit recognition. It implements perceptual input, human visual long-term memory in the ventral visual pathway and, to a lesser extent, working memory processes. To test the attractor hypothesis a Monte Carlo simulation of 10,000 individuals has been run. Each simulated learner was trained in recognizing the ten digits from novice to expert stage. The results replicate several features of human learning. First, they show that random connectivity in long-term visual memory accounts for novices' performance. Second, the learning curves revealed that Magnitron simulates the well-known psychological power law of practice. Third, after learning took place, performance departed from chance level and reached a minimum target of 95% of correct hits; hence simulating human performance in children (*i.e.*, when digits are learned). Magnitron also replicates biological findings. In line with research using voxel-based morphometry, Magnitron showed that matter density increases while training is taken place. Crucially, the spatial analysis of the connectivity patterns in long-term visual memory supported the hypothesis of a stable attractor. The significance of these results regarding memory theory is discussed.
Keywords: Memory; Numerical cognition; Neural network; Chaos; Attractor; Magnitron; Monte-Carlo simulations.

1. Introduction

1.1 Objectives

The present paper falls within the long tradition of cross-fertilizing ideas between artificial intelligence, cognitive neuroscience and psychology. The purpose was to investigate learning-induced structural changes in memory. To this end, we developed a multimodular architecture that models human learning of Arabic digits. The simulations were to provide evidence supporting the hypothesis that information stored in visual memory creates a neural structure that can be formalized as a metastable attractor. The huge constraints we imposed on our simulations were that the results should be consistent with both the psychology findings at the behavioral level and the key neuroimaging finding that learning modifies grey and white matter density in biological neural networks [1-2]. These changes at the macroscale result from neural plasticity occurring at a microscale [3] and underpin the development of cognitive performance. We set the target that our model of memory formation should replicate the fact that the structure of the brain is modified by experience and that matter density increases in the relevant connections. This work thus constitutes an attempt at offering a theoretical ground that accounts for behavioral and biological findings. Many excellent models of mathematical cognition are available [4-6]. These studies, though, address a different question than ours, usually focusing on the mathematical knowledge acquired, rather than the impact of learning on the structure of memory in the visual areas. Other works have already addressed separately, and with great depth, the dynamical properties of neural networks and their applications [7-10]. Here too these works address a different question than the structure of memory at a macroscale in relation to the acquisition of mathematical knowledge. Finally, much research has focused on optimizing learning algorithms for machine learning [11]. These investigations too are beyond our objectives for we aimed at replicating human behavior, including its weaknesses, rather than optimizing learning. Our model uses chaos theory as a framework to bridge the gap between different disciplines interested in the changes of neural structure following the formation

of new memories. We have chosen to simulate digit recognition for it is of relevance not only to neuroscientists and psychologists but also to educational scientists.

1.2 Biological background

The main achievements of humanity, such as putting a man on the moon or building a tunnel under the sea, are due to scientific knowledge. Mathematics, as the language of science, constitutes the pillar of modern society. Understanding the development of mathematical skills up until expert stage is thus crucial to psychology and education. It is now established that the ability to count stands at the core of mathematical intuition [12]. Various lines of evidence indicate that our counting ability is a skill inherited from animals [13-15]. That basic capacity to estimate quantities, though useful for performing many basic tasks, is not sufficient to allow the development of formal mathematical thinking. Advanced skills require good mastering of a symbolic format that can be acquired only through years of training. The learning process will ultimately enable the emergence of an abstract form of thinking [16]; the hallmark of expertise in mathematics. Because of the many years of training required to achieve mastery, mathematics has naturally been made the core component of scientific education. Naturally, the first training stage in mathematics consists in learning Arabic digits. The present paper uses a multimodular neural network architecture to simulate digit recognition. The model, largely informed by biological data, is used to test the hypothesis of neural attractors.

Recognition of Arabic digits relies on a learning-dependent network involving several brain sites [17]; each site performing a specific cognitive function. Arabic digits have long been hypothesized to be stored in the ventral visual stream [18-19]. A recent study using metanalysis techniques has confirmed the theoretical predictions by identifying the location of the so-called number form area in the inferior temporal gyrus [20]. The temporal site stores representations of the physical stimuli but the semantic representation of numbers is held in the intraparietal sulci [21]. These parietal loci contain populations of neurons whose activity are tuned to respond to natural numbers [22]. An early stage in mathematical development consists in forming a network that connects Arabic digits in the number form area to the appropriate cell assemblies in the parietal sites. The connection is likely underpinned by the vertical occipital fasciculus, for it is the only major fiber tract connecting the ventral to dorsal visual streams [23]. Once a number is recognized by the visual stream, semantic information in the interparietal sulci can be consciously manipulated through working memory which is implemented as a frontoparietal loop [24]. A model of number mining thus requires an architecture made at least of a retinal input, a number form area to encode digits, a site to represent quantities, and a frontal site that enables the recognized digit to be consciously accessible.

Performance primarily relies on our ability to recognize numbers. Thus, how representations are stored in memory is a process of paramount importance. Long-term memory storage, referred to as learning in psychology literature, is a purely biological process. The key mechanism underpinning learning is experience-dependent neural plasticity [25]; A general definition of which is the ability of the brain to modify its structure. The process is different from short term memory in that it requires the synthesis of proteins and thus involves a genetic component [26]. At the cellular level, neurons activate genes to build new connections [3]. Learning thus constitute a perpetual reconstruction and adaptation of the brain's neural networks. There is now ample evidence of such reorganization of the neural circuitry [27-28]. Two features of experience-dependent neural plasticity are crucial to characterize the biological essence of long-term memory. First, learning generates more connections, which implies more neural tissue. This reasoning has been empirically confirmed by neuroimaging studies comparing matter density in targeted brain regions. For example, by using voxel-based morphometry, it has been shown that experts in mathematics display higher density matter than novices in areas performing mathematical tasks [29]. Second, learning is a synapse-specific process [30]. The consequence is that information is encoded into well-defined functional modules [31-32]. Visual objects are thus stored in the ventral pathway of the visual system [33]. Location specific memory encoding of perceptual knowledge has been evidenced for many objects including faces [34-36], cars [37] and chess patterns [38]. Altogether, the available evidence shows that visual objects are stored as highly-connected cell assemblies in location-specific neural networks [39-40]. The number form area [18,41], storing Arabic digits, constitutes one instance of this learning process.

If the brain location wherein numbers are stored is well-established, the biological structure of memory within the site remains an open debate. Many models in psychology circumvent the problem by considering knowledge at the symbolic level; that is, memory stores and manipulates internal representations of letters or words. If such level of description is sufficient to explain most of the behavioral data, it is far from acceptable if we want to reach an understanding of memory that would integrate both behavioral and biological data. In the present study we take a computational approach inspired from physics. The brain is a dynamical system that has the unusual ability to self-organize its internal structure. It is thus not surprising that behavior and brain display chaotic properties [42]. A key feature of dynamical systems is the so-called attractors. Attractors are originally defined as the set of values a system tends to take over time (for a more technical definition see [43]). In this context, the phase space is the space all the possible states of the system. With time elapsing, and regardless of its original position in the phase

space, the system will end in the attractor subspace. Considering any cell assembly [44], the phase space corresponds to the set of values that all neurons and their connections can take [45]. If the attractor is the specific state of neural activity that encodes a memory to be acquired then all learning experiences should lead to such state. This hypothesis is tested with a multimodular artificial neural network.

1.3 Magnitron

Magnitron, in its basic form, was introduced to simulate recognition of analogue and symbolic numerosities [46]. This initial implementation was oriented towards replicating human performance but left little room for biological validity and interpretation. The present version of Magnitron has been developed to account for behavioral results while meeting biological constraints.

Magnitron is composed of four modules, see Figure 1. We shall examine the structure of each module in turn.

(1) Retina module. This module implements the perceptual input of the cognitive architecture. The retina can be conceptualized as a matrix of cells capturing light. In the present article, we will consider all cells as passive receptors. The artificial retina is an 8 x 8 matrix that codes inputs as 0 and 1 and converts the 8 x 8 binary input into a 64-element vector. The top of Figure 2 shows the coordinate system that is used for the reconversion into a 64-element vector. The bottom panel illustrates conversion with the digit '5'. It is worth noting that conversion is a one-to-one mapping so that topological relationships are conserved. No learning is taking place in the retina module. The 64-element vector output is directly fed to the Ventral Visual Stream.

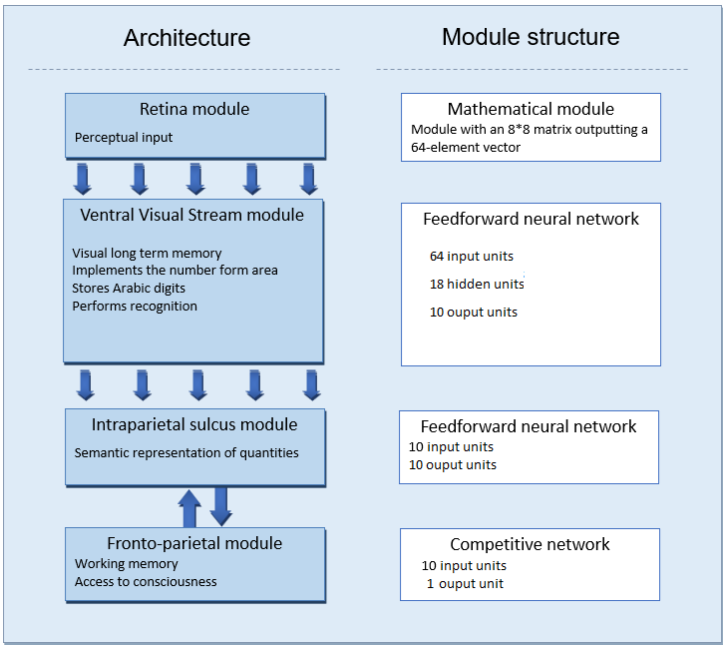

Figure 1. Architecture of Magnitron

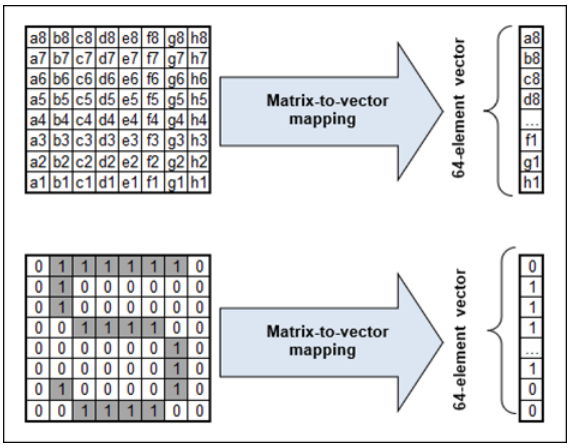

Figure 2. Mapping of vectors.

(2) Ventral visual stream module. The ventral visual stream implements the flow of information between the retina and, crucially, the site in the temporal cortex that performs recognition. It is implemented as a feed-forward neural network with three layers. The input layer is made of 64 input units, i ∈ $\mathbb{Z}\{1;64\}$. These units do not have a transfer function so the input, noted I_i, is directly fed into the network. The hidden layer was made of 18 hidden units, j ∈ $\mathbb{Z}\{1;18\}$. Finally the output layer was made of 10 units, k ∈ $\mathbb{Z}\{1;10\}$. The number of hidden units was chosen to ensure good performance while keeping calculations within available computational power. All units in the hidden and output layers were designed with a tan-sigmoid transfer function. Equations 1 and Equation 2 below indicate how the output of each layer is calculated as a function of its input. For the hidden layer, all weighted inputs I_i are summed to provide I; which is used to compute the output signal of each hidden unit, noted G_j, as indicated in Equation (1). The value of I is specific to each unit j as it depends not only on the 64 input values i but also on the specific weights, noted V_{ji}, of the connections between unit j and every unit in the input layer. Similarly, the input of each unit, in the output layer is the sum of the output value from all hidden units multiplied by the specific weight, noted W_{kj}, representing the strength of the connection between hidden unit j and output unit k. That sum, termed J, is used to calculate the final value of each unit in the output layer, noted O_k with Equation (2).

$$G_j = \frac{2}{1+e^{-2I}} - 1, \quad \text{for j} \in \mathbb{Z}\{1,18\} \text{ with } I = \sum_{i=1}^{64} V_{j,i}.i \tag{1}$$

$$O_k = \frac{2}{1+e^{-2J}} - 1, \quad \text{for k} \in \mathbb{Z}\{1,10\}, \text{ with } J = \sum_{j=1}^{18} W_{k,j}.G_j \tag{2}$$

After the presentation of a stimulus in the input layer, the signal is cascaded to, first, the hidden layer and then to the output layer. The values of the output units reflect the degree at which symbols are recognized. For example, an output vector of **O** = [0 0 1 0 0 0 0 0 0 0] indicates that the symbol 3 was recognized and an output vector of **O** = [0 0 0 0 0 0 1 0 0 0] indicates that the symbol 7 was recognized. The ventral visual stream is the repository of knowledge; as such it is the target of the learning process.

Long-term memory storage requires a significant amount of time in biological networks for the rewiring of the network is controlled by complex mechanisms that include genetic activation. Newly-created synapses can be both inhibitory and excitatory. As a proxy, we consider that a weight of zero is the absence of a synapse and a weight different than zero is a synapse. If the weight is positive the synapse is excitatory and if the weight is negative the synapse is inhibitory. Within this framework, the adjustment of the weights in the matrix reflects the change in distribution of axons and number of synapses while the organism is learning.

(3) Intraparietal Sulcus module. This is a feedforward network where 10 units of the input layer are connected one-to-one to the 10 units of the output layer. All output units are defined by a hard limit that outputs zero if the input signal is higher than 0.5. This module implements the connection between the number form area in the ventral visual stream and the intraparietal sulcus where the semantic representation of quantity is held.

(4) Fronto-Parietal module. Working memory is biologically implemented as a frontoparietal loop. Neurons in the frontal cortex constantly refresh neural signal in posterior areas [47]. Fine grained analysis of animal data indicates the centrality of this network for the processing of digits [24]. The output unit of the module encodes the information that is stored in working memory and consciously accessible. This process is implemented as a competitive neural network of two layers that will select only the most active output unit to put its 'value' in working memory.

2. Methods

Magnitron was implemented using the neural network toolbox from Matlab® [48]. A Monte Carlo simulation was performed to test whether numerical knowledge is encoded as an attractor in long-term visual memory. We simulated 10,000 novices. For each novice, the weights in the ventral visual stream were randomly initialized using a uniform distribution with a value range [-1; 1]. Magnitron was then trained to recognize the ten Arabic digits (0, 1, 2, 3, 4, 5, 6, 7, 8, and 9). The following parameters were set for the training session of the ventral visual stream. Performance was measured by the mean square error, rated c. This is the most common indicator of performance in neural networks. The mean square error is the squared difference between the expected and actual output averaged across all units. A mean square error with a value of zero indicates that the stimuli is perfectly processed and recognized. The performance criterion c used to stop performance was $c = 0.0001$; which implies that the training would stop as soon as the mean square error of the output of the ventral visual stream would be equal or inferior to 0.0001. We also set a minimum learning gradient of $\eta = 10^{-7}$. The learning gradient is the minimum amount by which the ventral visual stream module's weight will be changed. A limit in the number of the maximum of epochs for learning was set to 200.

2.1 Stimuli and training cycle

Ten Arabic digits were coded as 8*8 matrices. Figure 3 below shows the stimuli used for the training of Magnitron. The corresponding 10 outputs were paired with their corresponding 64-element input vectors for training in the ventral visual stream.

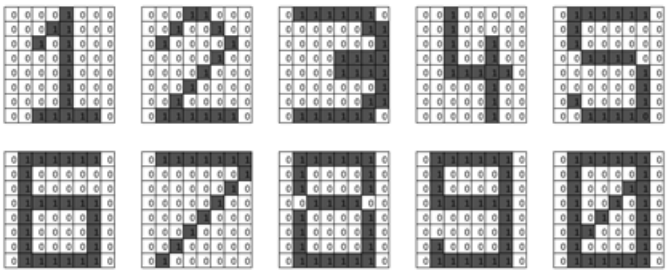

Figure 3. Stimuli sample used for the simulations

As indicated in the above, learning takes place in the ventral visual stream. Supervised learning was performed with the Levenberg-Marquardt algorithm [49-50], see Equation 3. The Levenberg-Marquardt algorithm changes the weight matrix x as a function of the difference between the expected and actual output. Considering Equation 3, $\boldsymbol{J}$ is the Jacobian of the matrix under consideration, $\boldsymbol{I}$ is the identity matrix, μ is the gradient, and e is the error.

$$x_{k+1} = x_k - [\boldsymbol{J}^T.\boldsymbol{J} + \mu\boldsymbol{I}]^{-1}.\boldsymbol{J}^T.e \tag{3}$$

The cycle of presenting an input vector, calculating the output, and adjusting the weight matrix, is termed an epoch. The epoch, in particular the phase of adjusting weights, implements the biological process of building up the network by creating new synapses. It thus represents a good artificial candidate to simulate a well-identified, biological process.

2.2 Memory attractor

Considering that both biological and artificial networks learn by building or modifying connections, connectivity was used to examine whether a specific pattern of connectivity underpins performance. As indicated in the above, non-zero weights reflect the existence of either an excitatory or inhibitory synapse. In line with this biological reasoning, we aggregated the absolute values of positive and negative connections to calculate the average matrix of connections. For any simulated individual n, we thus have two matrices characterizing their long-term memory structure. One matrix is the weight matrix V_{ji}^n that reflects the weight of the connection between the i^{th} input units and j^{th} hidden units for novice n. The other matrix, W_{kj}^n, is the weight of the connection between the jth units in the hidden layer and the kth units in the output layer for novice n. Equation 4 shows how the density matrix, $\mathbf{D_1} = \left(d_{1,ij}\right)$, was calculated for the weights connecting the input units to the hidden units and Equation 5 shows the calculations for the density matrix, $\mathbf{D_2} = \left(d_{2,ij}\right)$, between the hidden and output units. Learning can be quantified by calculating the average change in connectivity in the whole network; which basically translates mathematically as adding the averaged absolute weights all two density matrices.

$$D_{1,ij} = \frac{1}{N}\sum_{n=0}^{N-1}\left|V_{ji}^n\right| \tag{4}$$

with $i \in \mathbb{Z}\{1,64\}$, $j \in \mathbb{Z}\{1,18\}$ and N the number of novices, here $N = 10^4$

$$D_{2,ij} = \frac{1}{N}\sum_{n=0}^{N-1}\left|W_{kj}^n\right| \tag{5}$$

with $j \in \mathbb{Z}\{1,18\}$, $k \in \mathbb{Z}\{1,10\}$ and $N = 10^4$

To disentangle the memory structure of each output units, the density matrices for each output unit, noted $\boldsymbol{D}_{network}^k$, with k the index of the unit, was computed by adding the weights of the second density matrix that relate to unit k only to the density matrix D1.

$$\boldsymbol{D}_{network}^k = \left(d_{ij}^k\right) \text{ with } d_{ij}^k = d_{1,ij} + d_{2,jk} = \frac{1}{N}\sum_{n=0}^{N-1}\left(\left|V_{ij}^n\right| + \left|W_{jk}^n\right|\right) \tag{6}$$

Density matrices not only quantify the degree of connectivity but also offer a mean to analyze the distribution of weights. Since they are the phase space that reflects the degree of knowledge stored by the network. If a pattern

of connectivity is present in the phase space for expert and absent for novices, then we could conclude that the learning process tends to reach a specific pattern of connectivity. To test this assumption, we used Global Moran I index of spatial autocorrelation. This statistical test has been developed to determine whether elements in a spatial map are related to each other, thus showing if they are distributed randomly or clustered in specific locations.

The Global Moran I index for matrix D_1 (resp. D_2) is calculated using the formula:

$$I_D = \frac{N}{W} \frac{\sum_m \sum_n w_{mn}(d_m - \bar{d})(d_n - \bar{d})}{\sum_m (d_m - \bar{d})^2} \quad (7)$$

where dm and dn are elements of matrix $\mathbf{D_1}$ (resp. $\mathbf{D_2}$), $\bar{d}$ is the mean value of all elements in $\mathbf{D_1}$ (resp. $\mathbf{D_2}$), N is the total number of elements in $\mathbf{D_1}$ (resp. $\mathbf{D_2}$), w_{mn} is the element of the hollow matrix of spatial weights and W is the sum of all these elements.

Moran's I values in the novice and expert density matrices will reveal the degree of randomness in the distribution of weights. The neural attractor hypothesis predicts no spatial autocorrelation in the novice matrices for the weights will be assigned randomly. But, it predicts spatial autocorrelations to be revealed in the expert matrix as the attractor hypothesis predicts location-specific connections to encode knowledge. Both results combined would demonstrate that values in a network converge towards a specific pattern of connectivity: the attractor.

3. Results

We shall present our results in four sections. The first section consists in reporting the characteristics of the simulations at the novice stage. We report density matrices and compare connectivity across the input and output units to ensure that randomization of the weight matrix was uniform. In the second section, we report on the learning process of the 10,000 simulations. A mathematical model of the learning curves is then described. In the third section, we report the characteristics of the network after learning has taken place. We show that Magnitron replicates findings at the biological level. Finally, by analyzing the topological distribution of connections we test the existence of an attractor.

3.1 Novices' performance

Following the randomization procedure, the average density was *M* = .500 (*SD* = .003) for layer 1 and also M = .500 (*SD* = .003) for layer 2. The 10,000 novices performed poorly in recognizing digits (9.97%; *SD* = 1.75%). Their performance cannot be distinguished from chance performance, $t(9999) = -.33$, $p = .74$. This result confirms that connectivity in the network was random. Figure 4 shows the average connectivity of the input and output units averaged across their connection to the common hidden layer. As Figure 4 suggests, connectivity was similar across units of the same layer, an intuition that is statistically confirmed for both layer 1 ($F(1,63) = 1.087$, $p = .304$, MSE < 0.01) and layer 2 ($F(1,9) = 1.723$, $p = .087$, MSE < 0.01).

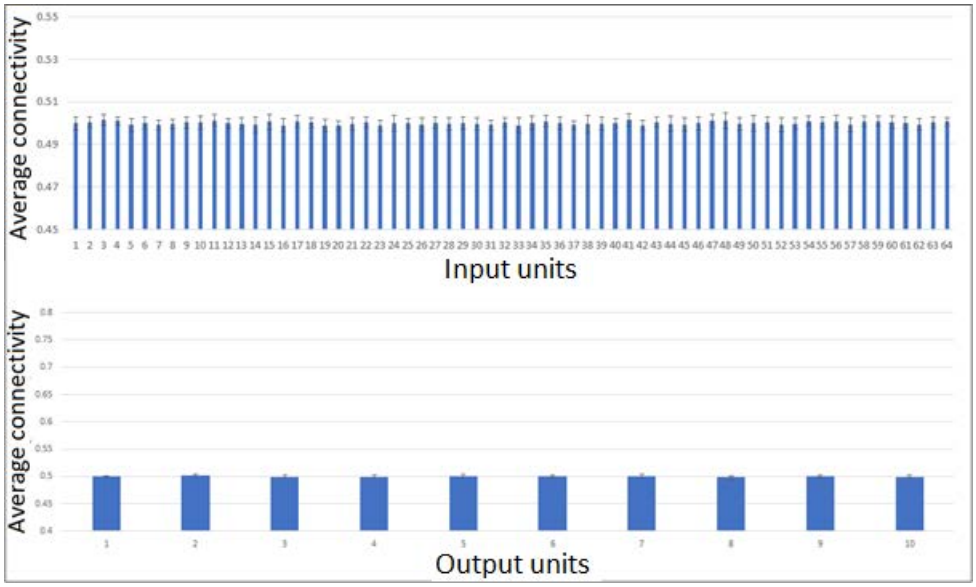

Figure 4. Average connectivity of the units in the first layer

3.2 Learning

Figure 5 shows the evolution of the performance criterion, mean error of the network, as a function of the number of epochs as averaged across the 10,000 simulations. Regressing the number of epochs to the average performance, as measured by mean error, provides an equation that significantly accounts for 92.5% of variance, the best fitting function was Performance = 11.247*epoch-3.64, $F(1, 34) = 408.486$, $p < .001$.Evolution of performance is thus adequately captured by a power law. Though all 10,000 simulations, learners reached expert stage in 12.169 epochs (SD = 3.383) on average. There is a noticeable variance in the number of epochs needed to reach expertise, ranging from 7 to 56 with a median value of 11.

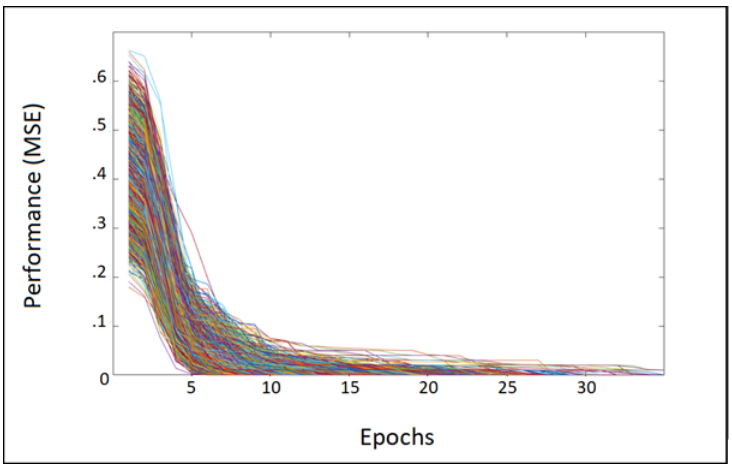

Figure 5. Learning curves

3.3 Experts' performance

After learning, the 10,000 simulations yielded an average recognition performance of 99.71% (SD = 1.75%). Learning has successfully changed the pattern of connectivity. Figure 6 displays the averaged connectivity for each unit of the input and output layers. In contrast with the connectivity pattern of novices, connectivity at the expert stage is not evenly distributed across units. Units in layer one have an average connectivity of $M = 0.507$ (SD = .006) but the distribution is not even across all units, $F(1,63) = 96.97$, $p < .01$, MSE < 0.01. The mean density in connectivity of layer 2 is $M = 0.726$ ($SD = 0.040$). Similarly, it is not evenly distributed, $F(1,9) = 1013.96$, $p < .01$, MSE < 0.01. These two results constitute key evidence that learning has significantly changed the pattern of connectivity within the network.

The overall density of the network has undergone a significant change during training. The average density of the first layer has significantly increased with training as indicated by an analysis of variance comparing the mean density input units across connections to the hidden layer $t(1151) = -35.91$, $p < .01$. Similarly, the analysis of variance on output units connectivity indicates that the substantial increase in average density of the second layer is significant $t(179) = -75.92$, $p < .01$. The question remains of whether these changes display a topological pattern.

3.4 Memory structure

The topological reorganization of the network can be seen in the density matrices D1 and D2, see Figure 5. The left column pictures D1 density matrices with the input layer in abscissa and the hidden layer in ordinate. The right column depicts the density matrices between the hidden (abscissa) and output layers (ordinate) for D2. The connectivity map before training reflects the novice stage, wherein weights have been allocated randomly. The novices' mean pattern of connectivity is pictured in the top density matrices. These matrices merely reflect the random distribution of weights over the 10,000 simulations and as such reveal no internal structure. After training there is a clear trend indicating that some pathways between units are more favored than others. The pattern of connectivity as revealed by the two matrices at the bottom of Figure 5 shows in topological maps the averaged connectivity matrices that the 10,000 simulated experts use to recognize and successfully classify the 10 basic digits.

Moran's I statistics were used to estimate the amount of spatial stratification for each connectivity matrix. The tests were not significant for both the first layer (observed = -0.001562738, expected = -0.0008688097, sd = 0.001519137, p = 0.6478213) and the second layer (observed = -0.008058135, expected = -0.005586592, sd = 0.006730431, p = 0.7134556) of the novice density matrices. But, both expert matrices yielded statistically significant spatial distributions, showing that they display a significant stratification of the distribution of connectivity over space; for Layer 1 (observed = 0.0853238, expected = -0.0008688097, sd = 0.00151734, p < .01) and for layer 2 (observed = 0.07809298, expected = -0.005586592, sd = 0.00676179, p < .01). The patterns of

connectivity of the expert matrices visible in Figure 7 are not the object of a random distribution but reflect spatial autocorrelations. The attractor hypothesis is statistically supported.

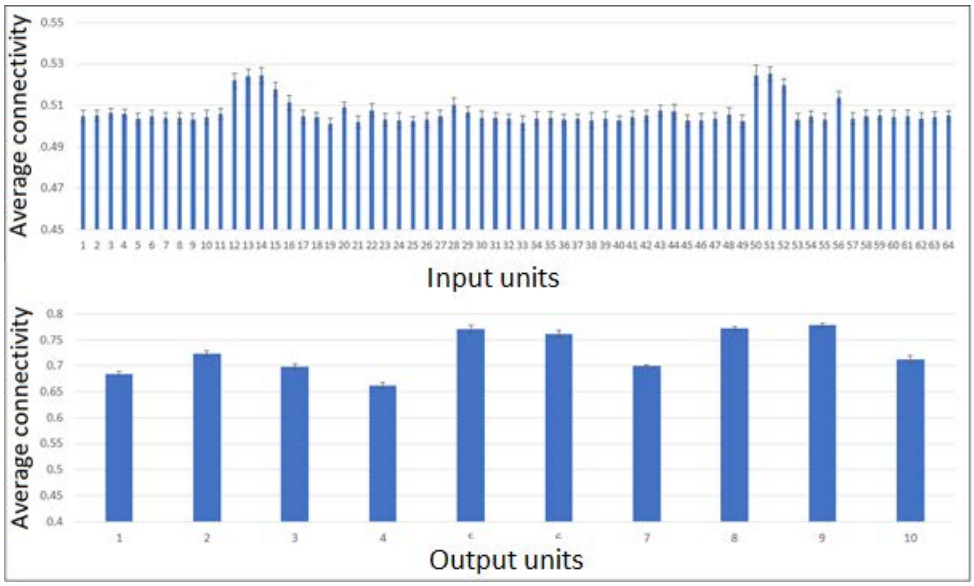

Figure 6. Average connectivity of the units in the second layer

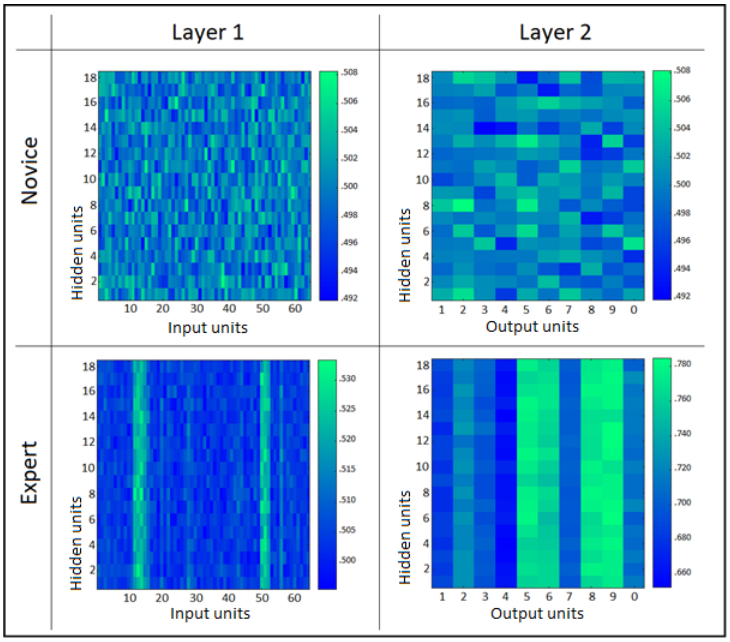

Figure 7. Matrices of connectivity of novices and experts

4. Conclusion and discussion

To test whether attractors are a general feature of visual memory we simulated learning processes with a multi-modular artificial neural architecture in 10,000 virtual human agents. The pattern of results validates our model, Magnitron, as a viable biologically-inspired model of digit recognition. Magnitron simulates many key features of human learning. First, the low level of novices' performance is shown to be adequately accounted for by randomized connectivity in the ventral visual stream. Second, experts' performance is simulated with accuracy. Third, the model simulates the learning curve of human agents showing its ability to replicate not only performance but also its evolution over time. Finally, in addition to simulating human behavior, we found key support for our attractor hypothesis. We shall discuss these results individually for each constitutes a piece that helps solving the puzzle of human memory.

Magnitron accounts for novice's performance in a digit recognition task by modelling their visual long-term memory as a network of random connections. This poses a theoretical problem as there could be no random connectivity in biological networks for it would imply a biological cost to serve no function. The question posed is therefore the interpretation of the randomized connectivity in the ventral visual stream module. Prior to the learning of numbers, neurons are engaged in performing other tasks. For example, neurons in the ventral visual areas are recruited early in the course of development to store the representations of many different types of items such as faces, building, and other everyday objects [51]. So, when the time comes to learn digits, the ventral visual stream is made of pre-existing connections. This pre-existing network, though, is not wired to properly process digits. Random connectivity in Magnitron should be interpreted as connections that do not contribute to digit recognition but reflect existing organization in the ventral visual stream of the individual. In our model, the state of knowledge of the simulated individuals is determined by the 1,332 connections (64x18 + 18x10) of their visual long-term memory. These connections, randomized with a uniform function, do reproduce the level of performance displayed by novices.

The learning curves yielded by the simulations have been shown to be best modelled by a power function. A function that has been long known to model human skill acquisition [52]. Magnitron thus simulates a key finding in the literature on learning and expertise acquisition. The learning curves also revealed that novices displayed a significant amount of variance in reaching expert status; a finding further confirmed by the fact that the range of epochs needed to reach expertise across the 10,000 novices can vary from 7 to 56 epochs. Our simulations thus replicate the fact that the time necessary to reach expertise can vary by a factor up to 8 [53]. The number of epochs necessary to reach expertise is determined by the distance between the initial position of the agent in the phase space and the attractor. In this context, the difficulty with which an individual performs a task reflects the chance that the neural circuitry necessary to perform that task has been modified by previous learning experiences.

The results of simulations at the expert stage confirm that learning has been efficient. The 10,000 simulated individuals perform significantly better when they have reached the expert stage as compared to their performance as novices. Comparisons between novice and expert density matrices has indicated that the expert network is characterized by higher connectivity. This result is in line with neuroimaging studies using voxel-based morphometry that tested the same hypothesis in biological networks [29]. Magnitron thus not only replicates human performance but also simulates a key result at the biological level. Analyzing the spatial distribution of the expert network has revealed that the network has self-organized into a well-defined, topologically-determined pattern of connectivity. Such pattern, encapsulating knowledge in visual long-term memory, is an attractor. This finding is crucial in that it shows the attractor emerged as a consequence of learning. In this context the increasing amount of time necessary to maintain a constant rate of progress in learning a task is basically explained by the fact that the fine-tuning of the connectivity pattern requires minute adjustments. The many thousands of training hours reported to reach expertise ([54]) are thus justified by the necessity to adjust the relative weight of millions of neurons in biological networks. This logic holds true as much for biological organisms as it does for artificial systems. Behavioral, neuroimaging, neurobiological and now computational evidence suggest that, to encode information in the long term, neural networks converge to create a specific pattern of connectivity. The simulations in this paper have demonstrated that new knowledge is encoded in the phase space of memory as an attractor.

The interpretation of the present work is subject to limits. First is the fact that we used only one single topology and set of parameters. Future works will explore the manipulation of the number of hidden layers and learning parameters to examine how these variables affect the structure of the attractor. The results presented here though constitute a first step towards the demonstration that structural changes in memory reflect an attractor in the phase space of visual long-term memory. A second limit is the fact that choices regarding the learning functions and transfer functions might not reflect biological complexity to its full extent. It is worth reminding that our purpose was to model changes at a macroscopic level, rather than mesoscopic or microscopic scales. In this respect the model has been highly successful in replicating the increase in matter that is generated by long-term memory encoding. One main limit of the present simulations is the available processing power offered by current technologies. Computers, however powerful, do not allow simulating the number of neurons that the brain contains and thus a degree of simplification is necessarily involved in any neural network simulation. Even though our neural network simulations are perfectly in line with biological evidence at the cell and brain level to indicate the presence of attractors in memory, the concept remains a theoretical assumption that will find its ultimate proof only with appropriate biological imaging. A second, related limit is that the attractor revealed in the present study is necessarily a simplification compared to an attractor in biological networks. Memory attractors in biological systems might involves millions of neurons. The typical pattern of connectivity of a memory attractor in biological system might be highly complex and thus display properties that our model does not capture. A third limit is that the difference between biological networks and the simulation is that, in biological systems, the number of neurons varies from one individual to the next. Hence, when considering different biological agents, we actually consider different neural structures and thus attractors should differ across individuals.

Beyond accounting for human behavior, Magnitron shows that learning in artificial and biological processes

display similar features insofar as the mechanisms that are modelled depend upon the same principles of connectivity. The present paper crucially contributes to our understanding of human learning processes by introducing chaos theory in the fields of learning and expertise acquisition. It has demonstrated, through a Monte Carlo simulation that the behavioral and neuroscientific features of learning in the early stages of expertise acquisition in mathematics are accounted for by the notion of attractors.

5. References

[1] Groussard M, Viader F, Landeau B, Desgranges B, Eustache F, Platel H. The effects of musical practice on structural plasticity: the dynamics of grey matter changes. Brain and Cognition. 2014; 90:174-180.
[2] Bengtsson SL, Nagy Z, Skare S, Forsman L, Forssberg H, Ullén F. Extensive piano practicing has regionally specific effects on white matter development. Nature Neuroscience. 2005; 8(9):1148.
[3] Kandel ER. The molecular biology of memory storage: a dialogue between genes and synapses. Science. 2001; 294(5544):1030-1038.
[4] Dehaene S, Changeux JP. Development of elementary numerical abilities: A neuronal model. Journal of Cognitive Neuroscience. 1993;5(4):390-407.
[5] Ratcliff R, McKoon G. Modeling numerosity representation with an integrated diffusion model. Psychological Review. 2018;125(2):183.
[6] Verguts T, Fias W. Representation of number in animals and humans: A neural model. Journal of Cognitive Neuroscience. 2004;16(9):1493-1504.
[7] Ganguli S, Huh D, Sompolinsky H. Memory traces in dynamical systems. Proceedings of the National Academy of Sciences. 2008; 105(48):18970-18975.
[8] Miller P. Dynamical systems, attractors, and neural circuits. F1000Research, 5. 2016.
[9] Lomp O, Richter M, Zibner SK, Schöner G. Developing dynamic field theory architectures for embodied cognitive systems with cedar. Frontiers in Neurorobotics. 2016; 10:14.
[10] Schoner G, Kelso JA. Dynamic pattern generation in behavioral and neural systems. Science. 1988;239(4847):1513-1520.
[11] Arel I, Rose DC, Karnowski TP. Deep machine learning-a new frontier in artificial intelligence research. IEEE Computational Intelligence Magazine. 2010; 5(4):13-18.
[12] Dehaene S. Origins of mathematical intuitions. The Year in Cognitive Neuroscience. 2009; 1156: 232-259.
[13] Sella F, Berteletti I, Lucangeli D, Zorzi M. Spontaneous non-verbal counting in toddlers. Developmental Science. 2016; 19(2): 329-337.
[14] Dehaene S, Izard V, Spelke E, Pica P. Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. Science. 2008; 320(5880):1217-1220.
[15] Tudusciuc O, Nieder A. Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. Proceedings of the National Academy of Sciences. 2007; 104(36):14513-14518.
[16] Chassy P, Grodd W. Abstract mathematical cognition. Frontiers in Human Neuroscience. 2016; 9:4-5.
[17] Chassy P, Grodd W. Comparison of quantities: Core and format-dependent regions as revealed by fMRI. Cerebral Cortex. 2012; 22(6):1420-1430.
[18] Dehaene S, Cohen L, Sigman M, Vinckier F. The neural code for written words: a proposal. Trends in Cognitive Sciences. 2005; 9(7):335-341.
[19] McCandliss BD, Cohen L, Dehaene S. The visual word form area: expertise for reading in the fusiform gyrus. Trends in Cognitive Sciences. 2003; 7(7):293-299.
[20] Yeo DJ, Wilkey ED, Price GR. The search for the number form area: A functional neuroimaging meta-analysis. Neuroscience Biobehavioral Reviews. 2017; 78:145-160.
[21] Nieder A, Dehaene S. Representation of number in the brain. Annual Review of Neuroscience. 2009; 32:185-208.
[22] Tudusciuc O, Nieder A. Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. Proceedings of the National Academy of Sciences. 2007; 104(36):14513-14518.
[23] Yeatman JD, Weiner KS, Pestill F, Rokem A, Mezer A, Wandell BA. The vertical occipital fasciculus: a century of controversy resolved by in vivo measurements. Proceedings of the National Academy of Sciences. 2014; 111(48):E5214-E5223.
[24] Nieder A, Miller EK. A parieto-frontal network for visual numerical information in the monkey. Proceedings of the National Academy of Sciences. 2004; 101(19):7457-7462.
[25] Kandel ER, Dudai Y, Mayford MR. The molecular and systems biology of memory. Cell. 2014; 157(1):163-186.
[26] Mayford M, Kandel ER. Genetic approaches to memory storage. Trends in Genetics. 1999; 15(11):463-470.
[27] Holtmaat A, Svoboda K. Experience-dependent structural synaptic plasticity in the mammalian brain. Nature Reviews Neuroscience. 2009; 10(9):647.

[28] Trachtenberg JT, Chen BE, Knott GW, Feng G, Sanes JR, Welker E, Svoboda K. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. Nature. 2002; 420(6917):788.
[29] Aydin K, Ucar A, Oguz KK, Okur OO, Agayev A, Unal Z, Yilmaz S, Ozturk C. Increased gray matter density in the parietal cortex of mathematicians: a voxel-based morphometry study. American Journal of Neuroradiology. 2007; 28(10):1859-1864.
[30] Martin KC, Casadio A, Zhu H, Yaping E, Rose JC, Chen M, Bailey CH, Kandel ER. Synapse-specific, long-term facilitation of aplysia sensory to motor synapses: a function for local protein synthesis in memory storage. Cell. 1997; 91(7):927-938.
[31] Poldrack RA, Desmond JE, Glover GH, Gabrieli J. The neural basis of visual skill learning: an fMRI study of mirror reading. Cerebral Cortex. 1998; 8(1):1-10.
[32] Schwartz S, Maquet P, Frith C. Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination. Proceedings of the National Academy of Sciences. 2002; 99(26):17137-17142.
[33] Ishai A, Ungerleider LG, Martin A, Schouten JL, Haxby JV. Distributed representation of objects in the human ventral visual pathway. Proceedings of the National Academy of Sciences. 1999; 96(16):9379-9384.
[34] Chang L, Tsao DY. The code for facial identity in the primate brain. Cell. 2017; 169(6):1013-1028. e1014.
[35] Keller CJ, Davidesco I, Megevand P, Lado FA, Malach R, Mehta AD. Tuning face perception with electrical stimulation of the fusiform gyrus. Human Brain Mapping. 2017; 38(6):2830-2842.
[36] McCarthy G, Puce A, Gore JC, Allison T. Face-specific processing in the human fusiform gyrus. Journal of Cognitive Neuroscience. 1997; 9(5):605-610.
[37] Ross DA, Tamber-Rosenau BJ, Palmeri TJ, Zhang J, Xu Y, Gauthier I. High-resolution functional magnetic resonance imaging reveals configural processing of cars in right anterior fusiform face area of car experts. Journal of Cognitive Neuroscience. 2018;30(7):973-984.
[38] Bilalić M, Langner R, Ulrich R, Grodd W. Many faces of expertise: fusiform face area in chess experts and novices. Journal of Neuroscience. 2011; 31(28):10206-10214.
[39] McGugin RW, Gatenby JC, Gore JC, Gauthier I. High-resolution imaging of expertise reveals reliable object selectivity in the fusiform face area related to perceptual performance. Proceedings of the National Academy of Sciences. 2012; 109(42):17063-17068.
[40] McGugin RW, Newton AT, Gore JC, Gauthier I. Robust expertise effects in right FFA. Neuropsychologia. 2014; 63:135-144.
[41] Shum J, Hermes D, Foster BL, Dastjerdi M, Rangarajan V, Winawer J, Miller KJ, Parvizi J. A brain area for visual numerals. Journal of Neuroscience. 2013; 33(16):6709-6715.
[42] Barton B. Chaos, self-organization, and psychology. American Psychologist. 1994; 49(1):5-14.
[43] Milnor J. The theory of chaotic attractors. New York, NY: Springer;1985.
[44] Hebb DO. The organization of behavior: A neuropsychological approach. John Wiley Sons;1949.
[45] Chassy P. A neural network test of the expert attractor hypothesis: Chaos theory accounts for individual variance in learning. In: International Conference on Innovative Techniques and Applications of Artificial Intelligence. Springer. 2016. p.151-162.
[46] Chassy P, Buckley N, Reid D. Magnitron: A spiking neural network model of numerical cognition. Keedwell, 2nd Symposium on Nature-Inspired Computing and Applications. University of Exeter. 2013. p. 4-7.
[47] Linden DE. The working memory networks of the human brain. The Neuroscientist. 2007; 13(3):257-267.
[48] Beale M, Hagan M, Demuth H. Neural network toolbox™ reference. The MathWorks. Inc., R2017a. 2017.
[49] Levenberg K. A method for the solution of certain non–linear problems in least squares. Quarterly of Applied Mathematics. 1944; 2:164-168.
[50] Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics. 1963;11(2):431-441.
[51] Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of Neuroscience. 1997; 17(11):4302-4311.
[52] Newell A, Rosenbloom PS. Mechanisms of skill acquisition and the law of practice. Cognitive Skills and Their Acquisition. 1981; 1: 1-55.
[53] Gobet F, Campitelli G. The role of domain-specific practice, handedness, and starting age in chess. Developmental Psychology. 2007; 43(1):159-172.
[54] Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. Psychological Review. 1993; 100(3): 363-406.