Efficient Memory Encoding Explains the Interactions Between Hippocampus Size, Individual Experience, and Clinical Outcomes: A Computational Model

Andrea Stocco (stocco@uw.edu)
Department of Psychology, University of Washington, Seattle, WA 98195 USA

Briana M. Smith (brianam2@uw.edu)
Department of Bioengineering, University of Washington, Seattle, WA 98195 USA

Bridget Leonard (bll313@uw.edu)
Department of Psychology, University of Washington, Seattle, WA 98195 USA

Holly Sue Hake (hakehs@uw.edu)
Department of Psychology, University of Washington, Seattle, WA 98195 USA

Abstract

The relationship between hippocampal volume and memory function has produced mixed results in neuroscience research. However, an experience-dependent efficient encoding mechanism underlies these varied observations. We present a model that utilizes an autoencoder to prioritize sparseness and transforms the recurrent loop between the cortex and hippocampus into a deep neural network. We trained our model with the Fashion MNIST database and a loss function to modify synapses via backpropagation of mean squared recall error. The model exhibited experience-dependent efficient encoding, representing frequently repeated objects with fewer neurons and smaller loss penalties and similar representations for objects repeated equally. Our findings clarify perplexing results from neurodevelopmental studies: linking increased hippocampus size and memory impairments in ASD to decreased sparseness, and explaining dementia symptoms of forgetting with varied neuronal integrity. Our findings propose a novel model that connects observed relationships between hippocampus size and memory, contributing to the development of a larger theory on experience-dependent encoding and storage and its failure.

Keywords: Memory; Hippocampus; Autoencoder; Autism Spectrum Disorder; Alzheimer's Disease; Computational models; ACT-R

The hippocampus is a region of the medial temporal lobe that is critical for long-term memory storage and retrieval. The size of the hippocampus can vary significantly between individuals and these variations in size have been associated with corresponding differences in memory function (Pohlack et al. 2014; Hardcastle et al. 2020; Botdorf, Canada, & Riggins 2022). The relationship between hippocampus size and memory function, however, is complex and not always straightforward. On one hand, there is evidence that greater hippocampus volume is associated with better memory function. For example, greater hippocampus volume is associated with better spatial memory performance in a laboratory task (Erickson et al. 2011; Guderian et al. 2015). Conversely, reduced hippocampus size is associated with significant impairments in long-term memory. For example, in frontotemporal dementia and Alzheimer's disease, neuronal loss results in markedly reduced hippocampal volume, and the degree of volume loss positively correlates with the severity of amnestic symptoms (Dickerson et al. 2009).

The relationship between hippocampus size and memory performance is also, at least partially, mediated by experience. A notable case is the fact that London taxi-cab drivers have larger hippocampi than the normal population, likely due to the amount of information that cab drivers need to memorize ("The Knowledge") to pass the license test (Maguire et al. 2000). In fact, a follow-up study revealed that changes in hippocampus size follow, and do not precede, the amount of studying necessary to pass the test (Maguire, Woollett, & Spiers 2006). Similarly, changes in hippocampus size correlate with an individual's years of education (Nobis et al. 2019).

An intuitively appealing explanation for these effects might be that the hippocampus grows with the amount of data it needs to, or can, store. Thus, pressure to store more information results in the growth of the hippocampus, and a reduction in hippocampal size results in loss of memory. This simple explanation, however, is complicated by a number of other findings. Reductions in hippocampus size are observed in a variety of mental disorders, including post-traumatic stress disorder (PTSD) and anxiety. In these cases, significant reductions in hippocampal size are *not* accompanied by corresponding changes in long-term memory function (Karl et al. 2006). Conversely, larger hippocampal volume has been observed in autism spectrum disorder (ASD), where a corresponding increase in memory function was not observed (Varghese et al., 2017). In fact, evidence suggests that the prevalence of amnestic forms of dementia in ASD is up to four times higher than the neurotypical average, despite the fact that greater hippocampus size could have represented a buffering factor against neuronal loss (Fyfe, 2021). Thus, while it has been shown that experience drives changes in hippocampus size, changes have also been observed in clinical conditions *without* corresponding changes in memory.

At least three possible explanations can be proposed to reconcile these findings. The first and most mundane is that changes in hippocampus size might not always reflect

underlying changes in the number of hippocampal cells or synapses. Virtually all of the studies assess hippocampal size through anatomical MRI, and the sheer volume of a region in an MRI scan can be affected by a variety of other factors, such as greater water density (Bansal et al. 2013).

A second explanation is that two or more biological mechanisms might be at play. Thus, while experience-dependent growth following intense memory training and dementia-related loss of memory function are connected to the number of cells and synapses, changes in other clinical domains might be related to other processes. For example, prolonged stress exposure causes neuronal death through the accumulation of cortisol. Thus, it is possible that the volume loss in PTSD and anxiety are due to cortisol-related pruning, which does not play a role in dementia or ASD (Kim, Pellman, & Kim 2015).

The third and last explanation is that these varied phenomena are indeed connected by experience-dependent efficient allocation of hippocampal cells and synapses to varying memory demands, but that this relationship is complex and non-linear.

In this paper, we put forward a neurocomputational framework that provides a possible account for the latter hypothesis. According to this framework, the need to store and retrieve memories demands an efficient allocation of neural resources, and the principles underlying this allocation can be understood in terms of information theory.

The remainder of the paper is structured as follows. First, we review previous computational attempts to model the relationship between memory function and hippocampus size. Specifically, we review a model that explicitly links resource allocation in the hippocampus to the information entropy of its memories, and how entropy is altered in PTSD. Second, we propose a possible neural network model of how such changes could happen. The model, based on the autoencoder architecture, shows that, under realistic conditions, the hippocampus can spontaneously learn to allocate neurons adaptively according to the demands. Finally, we speculate on the implications of this mechanism for two important memory-related phenomena, sleep, and spontaneous brain activity,

Previous Models

To the best of our knowledge, the first computational account of the relationship between memory demands and hippocampus size was put forward by Smith et al (2021). The authors proposed a mathematical model of memory storage and retrieval based on information entropy.

The model is based on the framework originally proposed by Anderson and Schooler (1991) and currently implemented in the ACT-R architecture (Anderson, 2009). According to this framework, each memory is a collection of *traces*, each corresponding to a specific episode in which the memory's contents were encoded. This makes the model broadly consistent with the Multiple Trace Theory of memory (Moscovitch et al. 2005). The strength of each trace decays over time according to a power function. The *memory*'s total strength, or *activation*, is the log of the sum of its traces. Thus, if a memory m is made of n traces encoded at times $t_1, t_2 \ldots t_n$, its activation at time t is:

$$A(m,t) = \log \sum_i (t - t_i)^{-d}$$

where d is an individual-specific decay rate (Sense et al., 2016; Zhou et al, 2021). Note that Equation 1 naturally captures the effects of recency (through the decay term d) and frequency (through the accumulation of traces). The probability $P(m)$ of retrieving a memory can be computed as a function of its activation, relative to all other memories:

$$P(m) = e^{-A(m,\,t)} / \sum_j e^{-A(j,\,t)} \qquad (1)$$

Smith et al. (2021) proposed that the distribution of probabilities across memories could be used to predict changes in hippocampus volume. The authors assumed that the hippocampus would use efficient coding, and allocate fewer resources to store information that is most likely to be retrieved. This is a common principle in lossless compression algorithms (Huffman, 1952). Consider, for example, the problem of efficiently encoding the quote "All those moments will be lost in time like tears in rain". Using standard ASCII coding, each character in the string would be represented by 8 bits and the entire string would take a total of 456 bits. To *efficiently* encode the string, however, one would first count the occurrence of each character in the string and then proceed to assign the shortest possible code to the most common character, the second shortest to the second, and so on. In this case, the letter "e", "i", and "l", which appeared six times each, would be assigned the three-bit codes 001, 010, and 011, while the letter "k", which appears only once, would be given the six-bit code 101001. This would result in the entire string being encoded with only 206 bits.

We currently do not know with sufficient precision how information is encoded in the hippocampus. However, independently of the specific code, the degree of compression allowed by any adaptive scheme of this sort is functionally related to the information entropy H of the data:

$$H = -\sum_i p(i) \log p(i) \qquad (2)$$

Smith et al. (2021) showed that the reduced hippocampus size in individuals suffering from PTSD could be predicted by calculating the entropy of the retrieval probabilities (Equation 2) associated with every memory in the model. Specifically, when the model was modified to simulate emotional trauma, the persistence of intrusive memories had a significant effect on the probability distribution of the memories that could be retrieved. The more likely the intrusive memory was to be retrieved, the lower the entropy of the model's memory system, and as a correlate, the lower the volume of the hippocampus.

Limits of the Model

The original model by Smith et al (2021) was noteworthy but did not address a number of limitations. First, it provided no biological mechanisms by which neurons could be efficiently allocated to different representations. In fact, it could not solve the problem of how the hippocampus could

form an efficient engram without knowing in advance its future activation level.

A second limitation of the original model was its scope: it only addressed changes in hippocampal volume due to PTSD. The same framework can be arguably applied to anxiety, which shares with PTSD the transdiagnostic symptom of intrusive thoughts that are ruminated upon. It could be possibly extended to include experience-dependent changes as well (such as the effects of education). It does *not* address, however, other findings, such as the greater hippocampus volume in ASD and the association between smaller hippocampus size and memory loss in dementia.

A Neural Network Model of the Hippocampus

To address these limitations, we examined the behavior of a neural network model of the hippocampus and conducted a series of simulations to test whether (a) Efficient coding spontaneously emerges in more biological models, and (b) Whether the model can account for the diversity of findings relating hippocampus size and memory function.[1]

The connections between the cortex and the hippocampus form a recurrent loop. The exact set of synapses varies slightly across regions; as an example, this paper will consider the connectivity between the inferior temporal lobe and the hippocampus. This specific circuit is well understood and underlies memory for higher-level visual objects, which will be used as experimental stimuli. Projections from the inferior temporal cortex pass through the entorhinal cortex and the dentate gyrus before reaching area CA3 of the hippocampus, which is considered the initial seat of an engram (Tonegawa et al., 2015).

During recall, memories are then reactivated in the cortex (Danker and Anderson, 2010) through a series of connections that originate in CA3 and progress through area CA1, the entorhinal cortex again, and finally return to the temporal cortex.

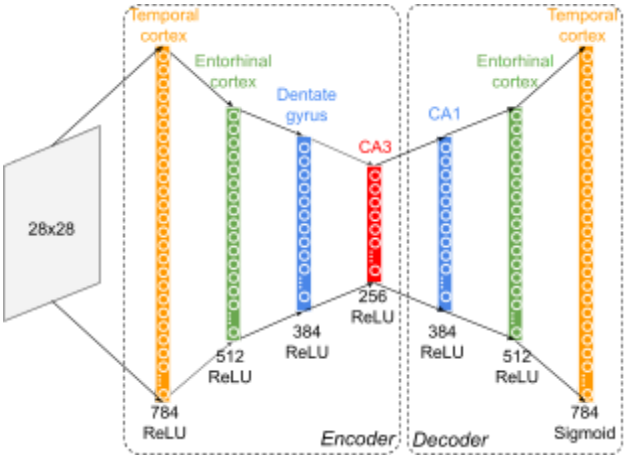

Figure 1: Architecture of the neural network model

For convenience, this recurrent loop can be "unrolled" and transformed into a feedforward deep neural network with multiple layers. The first half of the model corresponds to the neural populations encountered from the cortex to the hippocampus, and the second half to the neurons encountered from the hippocampus back to the cortex. In this design, the cortex is both the input and the output of the network, and the hippocampus is the central bottleneck. This architecture is technically known as an *autoencoder* (Kramer, 1991) and is used, in deep-learning applications, to learn a set of features that would efficiently compress the original input so that its output is minimally different from its input. To a large extent, the application of autoencoders can be construed as exactly the function of episodic memory and, by extension, of the hippocampus.

The model's final architecture is shown in Figure 1. Its input is a 28×28 matrix that contains a visual representation of an object. This representation is then flattened to a layer of 784 neurons, which represents the object as encoded in the inferior temporal cortex. This representation is passed through a smaller layer of 512 neurons, representing the entorhinal cortex, and an even smaller one of 384 neurons, representing the dentate gyrus. It finally reaches a layer of 256 neurons that hold a compression representation of the original content and stands for the hippocampus' CA3 field. The output of the hippocampus is then passed through a mirror series of layers representing CA1, the entorhinal cortex, and the temporal cortex again (784 neurons), generating a reconstructed version of the original stimulus. All of the neurons in the model are Rectified Linear Units (ReLUs), with the exception of the very last layer, which uses a sigmoid function to ensure that all of the predicted pixel values are, like the inputs, between 0 and 1.

The model was used in five different simulations, each of which addresses a different facet of the relationship between hippocampus size and memory function.

Materials and Methods

Model Implementation

The model was implemented in Keras with an underlying TensorFlow engine. In addition to those of Figure 1, the model contains four additional layers that perform purely technical operations such as reshaping inputs and outputs and computing penalty terms for the cost functions (see below); although necessary, these layers are not functionally relevant. Altogether, the model has a total of 1,347,282 trainable parameters.

Training and Testing Data

The model was trained on a selection of objects from the Fashion MNIST database (Xiao, Rasul, & Vollcraf, 2017), a collection of 70,000 28×28 black-and-white images from 10 clothing categories. A subset of 1,111 images was randomly selected at every run. The images were repeated with varying frequencies across simulations (see below) but always formed a training set of 4,000 stimuli.

[1] All data and code are available at https://osf.io/wxh2r/

Model Training and Loss Function

In all simulations, the model was trained on all of the training set items for five consecutive epochs using stochastic gradient descent with adaptive moment estimation (Adam: Kingma & Ba, 2014). We used a combined loss function L that included two terms:

$$L = \sum_{i,o}^{N} (y_i - y_o)^2/N + \lambda \sum_{h \in CA3} |y_h|$$

The first term is the *accuracy cost* of the network's recall function, and is the mean squared difference between the activations y_i and y_o of each input neuron i and corresponding output neuron o. The second term is the *resource cost* and is the sum of the activation y_h of each hippocampal neuron h in the CA3 layer. (Note that the penalty cost only interests the CA3 neurons). The hyperparameter λ regulates the weight of the penalty and was set to 0.00001 throughout these simulations; pilot tests showed that the results did not qualitatively change for different λ values, as long as λ was below a critical threshold of 0.0001, above which the penalty became too severe.

Note that the resource cost penalty is equivalent to the L1 penalty used in regularization methods, such as LASSO (Tibshirani, 1996). Unlike other penalties, the L1 penalty can force its terms to zero, thus reducing the number of active neurons within a representation.

Dependent Variables

For each of the 1,111 objects in the training set, four dependent variables were computed. Two variables measured the sparseness of the hippocampus representation: the value of the resource cost *L1 penalty* and the total *number of active neurons* (that is, with activation $y_h \neq 0$) in the hippocampus The other two variables measured the model's recall accuracy, and they were the value of the *error penalty*, i.e. the squared sum of differences between target and predicted activations in the output layer, $\sum_{i,o} (y_i - y_o)^2$, and the Pearson *correlation coefficient* between the encoded and recalled image. Because of their constrained range, correlation coefficients were normalized using Fisher's r-to-Z transform: $Z = [\log(1 + r) - \log(1 - r)] / 2$.

Results

Simulation 1: Emergence of Efficient Coding

In the first simulations, the set of 1,111 objects was used to create a 4,000-item training set in which different objects were repeated with different frequencies. Specifically, 1,000 objects occurred only once; 100 objects occurred 10 times, 10 objects occurred 100 times, and a single image occurred 1,000 times. If the model is learning a form of efficient coding, the internal hippocampal representation of an object should depend on its frequency in the training set, and, therefore, objects that are repeated the most should have representations with fewer neurons and smaller L1 penalties than objects that are repeated the least.

Figure 2 illustrates the results of one such simulation. The top row shows four example objects from one specific simulation, chosen from the sets of stimuli repeated 1, 10, 100, or 1,000 times, respectively. The middle row represents the corresponding responses of the simulated CA3 layer, with the activations of its 256 neurons arranged in a 16x16 grid. The dependent variables for hippocampus sparseness (L1 penalty and number of neurons) are also reported. Finally, the bottom row depicts the recalled memory.

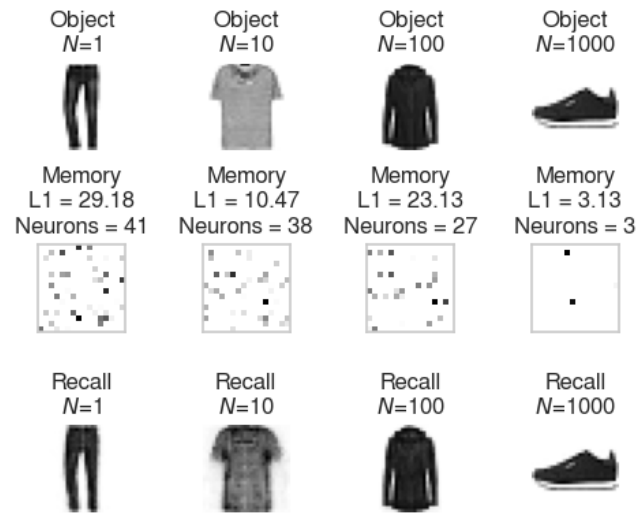

Figure 2: (Top) Four example stimuli that were repeated 1, 10, 100, or 1,000 times in the training set. (Middle) Corresponding CA3 representations of the stimuli; (Bottom) Recalled stimuli reconstructed by the decoder from the CA3 representations.

Although representative, Figure 2 only illustrates four examples from a single run. A complete overview of all simulations is instead reported in Figure 3, where the mean L1 penalty and the mean number of neurons are reported as the blue lines in the two panels. As the figure shows, higher frequency results in a dramatic reduction in the number of neurons needed to represent an object.

This change in representation has no consequences for the model, which has a fixed and immutable structure in which all synapses exist all the time, even when they are connected to silenced neurons. In a biological hippocampus, however, synapses and neurons change over time: synapses with a value of zero do not exist, and those connected to mute cells would simply be pruned. Thus, the sparse representations in Figures 2 and 3 could be associated with a smaller hippocampus size.

But how closely does the reduction in the CA3 representations match the predictions of information theory? According to Huffmann (1952), efficient codes are such that the length of a code for an object x matches its information content $I(x)$, which is the negative log of its probability: $I(x) = -\log_2 p(x)$. The value of $p(x)$ can be calculated from the number of occurrences of stimulus x in the training set. Figure 4 compares the relationship between the number of neurons used to encode an object in the CA3 layer and its corresponding information content. As the figure shows, the number of neurons closely mirrors ($r = .97$) the information content, a hallmark of efficient coding.

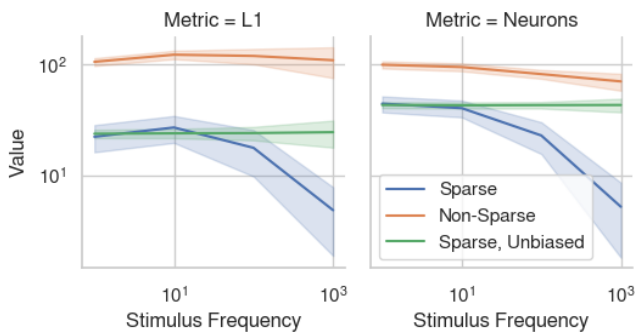

Figure 3: A summary overview of sparseness metrics (left: L1 penalty; Right: Number of neurons) across stimuli of different frequencies, averaged over 50 simulations. Colors represent different model conditions (*Blue* = Sparse; *Red* = Non-Sparse; *Green* = sparse with an unbiased training set). Lines and ribbons represent means +/- SD.

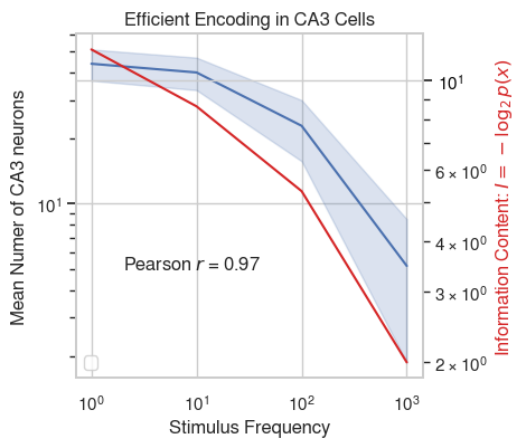

Figure 4: Relationship between the number of neurons that encode a stimulus (blue line and ribbon, representing mean +/- SD) and its information content (red line).

Simulation 2: Frequency Drives Efficient Coding

To ensure that the effect is driven by frequency and not by other confounding factors, a second series of simulations were run. In these series, the model was trained with a dataset of identical size (4,000 items) and containing the same 1,111 objects but with an *unbiased* number of repetitions per object, i.e., with each object being repeated 3 or 4 times.The results of these simulations are shown as the green line in Figure 3. Under these conditions, both the L1 penalty and the number of neurons remain invariant and equal to the values of the least-frequent memories in Simulation 1. Note that, for the unbiased training model, the frequencies on the *x*-axis do not actually reflect the frequencies of the simulation training set; instead, they are used to identify the corresponding group of objects in Simulation 1.

The results of this simulation could be used to explain the increased hippocampal volume in London taxi cab drivers compared to bus drivers (Maguire et al., 2006): as noted by Smith et al. (2021), taxi cab drivers, unlike bus drivers, have to rehearse the streets of London with comparable frequency to prepare for the license test.

Simulation 3: Enlarged Hippocampus in ASD

In the model, sparseness is achieved by adding a penalty to the loss function. In a biological network, however, sparseness must be achieved through some neural mechanism. The most straightforward candidate is lateral inhibition, that is, inhibitory synapses between neurons belonging to the same region. Inhibitory synapses typically express GABA receptors, and abnormally low expression of GABA receptors is a key characteristic of ASD (Cellot & Cherubini, 2014), one of the disorders also characterized by abnormalities in hippocampus size. Recent studies estimate that individuals with ASD express as much as 40% fewer GABA receptors than healthy controls. Thus, we hypothesized that the reduced availability of GABA receptors in ASD may lead to a decrease in lateral inhibition, resulting in less efficient coding and thus the larger hippocampus observed in ASD.

To test this hypothesis, a series of simulations were carried out using the biased training set but with the λ parameter set to λ=0, allowing for a minimum amount of sparseness based solely on the thresholds of the ReLU units. The results of the simulations with such a Non-Sparse model are shown in the red lines of Figure 2. As it can be seen, without the resource cost penalty, the model now uses a disproportionately large number of neurons and incurs in large L1 penalties. Furthermore, both measures remain remarkably stable even when encoding extremely high-frequency stimuli, indicating that the hippocampus is not using efficient coding.

Simulation 4: Hippocampal Damage in Dementia

As noted in the introduction, some reductions in hippocampus size *are* associated with distinctive deficits in memory. This is the case, for example, of neurodegenerative diseases such as Alzheimer's Disease. In these cases, neuronal loss afflicts long-term memory by harming the engram associated with a specific memory.

To simulate the effects of dementia, we ran a fourth series of simulations, identical in nature to Simulation 1 but with an additional manipulation. After completing the training phase, the model's hippocampus was artificially damaged by applying a binary mask to the activation of its units. Binary masks were generated by creating a null vector of 256 elements, and randomly setting a percentage of its elements to 1. The proportion of units set to 1 represents the *neuronal integrity* of the hippocampus and was parametrically varied from 0.1 to 0.9. After every simulated lesion, the model's recall was tested again, and the two accuracy measures (squared recall error and recall correlation) were recorded. Figure 5 illustrates these results.

Interestingly, and consistent with the observed symptoms of dementia, less frequent memories are more affected, even at higher levels of neuronal integrity, than the more frequent ones, which remain comparatively well preserved even at lower levels of neuronal integrity.

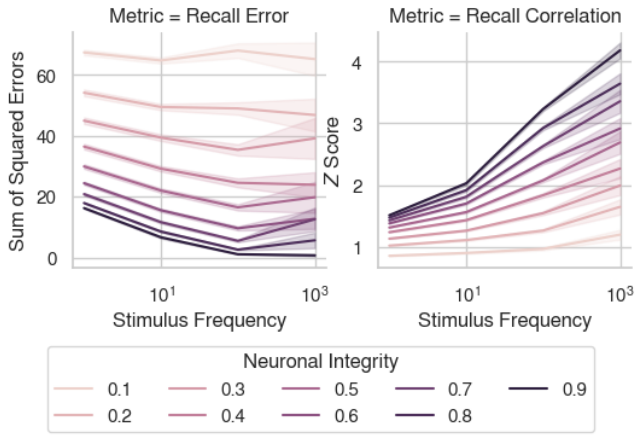

Figure 5: Effects of neural damage on recall accuracy. Lines and ribbons indicate means +/ SD.

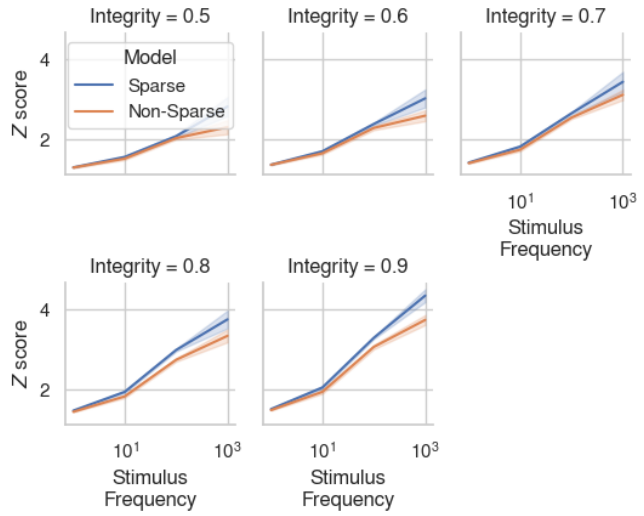

Figure 6. The *Z*-scored recall correlation coefficients of Sparse (blue) and Non-Sparse (red) models. Line and ribbons indicate means +/- SD.

Simulation 5: Interactions Between ASD and Neurodegenerative Disorders

Although ASD *per se* is not associated with notable changes in long-term memory function, it has been noted that dementia has a higher prevalence in individuals with ASD than in neurotypical controls (Fyfe, 2021). The results of Simulations 3 and 4 suggest that an additional advantage of efficient coding of memories is to buffer against neuronal death. Conversely, the less sparse memory representations in ASD might be more susceptible to damage from neuronal loss, thus explaining the greater prevalence of dementia in ASD. To test this hypothesis, we conducted a second series of lesion simulations, identical to the ones in Simulation 4 but with the model's λ parameter set to $\lambda = 0$. Because we are especially interested in the earlier stages of neurodegenerative disease, only the neuronal integrity values from 0.5 to 0.9 were examined. The results are summarized in Figure 6. As the figure shows, the Non-Sparse model is consistently more affected than the sparse model by damage, across all levels of stimulus frequency and neuronal integrity.

Discussion

This paper has dealt with the relationship between hippocampus size and memory functions across clinical and neurotypical populations. Specifically, it has shown that some puzzling findings in the literature can be reconciled when one analyzes the behavior of a neural network model of the hippocampus whose loss function includes a resource cost. The resource cost penalty induces sparseness in a form that is consistent with the principles of efficient coding and with the idea, first proposed by Smith et al. (2021), the hippocampus size reflects the information content of the stored memories. These contributions notwithstanding, a number of limitations must be acknowledged. First and foremost, the model uses an autoencoder architecture, while hippocampus models are more commonly implemented as autoassociators (e.g., Treves & Rolls, 1994).

As a consequence, the model requires error-driven methods to learn properly and is incapable of "one-shot" Hebbian learning. Autoencoders were chosen because they make it easier to capture the dynamics of encoding and recall and the relationship between cortical areas and the hippocampus. A proper model, however, should attempt to combine both architectures and include principles of auto-associative Hebbian learning with the hippocampus.

Second, a number of factors that affect the model's memory recall and performance are left unexplored. Among those, perhaps the most important is the role played by the number of epochs used in training. It is possible, for example, that sparse models would require longer epochs to achieve the same recall accuracy. The combined use of error-driven learning and multiple training epochs highlights another aspect of the model, namely, its need for multiple learning passes to discover efficient representations. As noted in the introduction, the hippocampus cannot assign efficient memory codes right away, as they require knowledge of an object's frequency. In the autoencoder, it is the presence of multiple learning passes and gradient descent that pushes for sparser and more efficient coding. It is possible that spontaneous brain activity, which is prominently displayed in the hippocampus at rest and during sleep (Pfeiffer, 2020), provides a biological surrogate for the necessary re-experience of memories that are needed for efficient coding.

Lastly, the model is silent about the nature of forgetting, another prominent feature of memory that might be connected to the spontaneous replay of memories at rest (Zhou et al., 2021). Future research will be needed to further explore the nature of these processes within the model.

References

Anderson, J. R. (2009). *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*(6), 396-408.

Bansal, R., Hao, X., Liu, F., Xu, D., Liu, J., & Peterson, B. S. (2013). The effects of changing water content, relaxation times, and tissue contrast on tissue segmentation and measures of cortical anatomy in MR images. *Magnetic Resonance Imaging, 31*(10), 1709-1730.

Botdorf, M., Canada, K. L., & Riggins, T. (2022). A meta‐analysis of the relation between hippocampal volume and memory ability in typically developing children and adolescents. *Hippocampus, 32*(5), 386-400.

Cellot, G., & Cherubini, E. (2014). GABAergic signaling as therapeutic target for autism spectrum disorders. *Frontiers in Pediatrics*, 2, 70.

Dickerson, B. C., Bakkour, A., Salat, D. H., Feczko, E., Pacheco, J., Greve, D. N., ... & Buckner, R. L. (2009). The cortical signature of Alzheimer's disease: regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. *Cerebral Cortex, 19*(3), 497-510.

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., ... & Kramer, A. F. (2011). Exercise training increases size of hippocampus and improves memory. *PNAS, 108*(7), 3017-3022.

Fyfe, I. (2021). Early-onset dementia in autism spectrum disorder. *Nature Reviews Neurology, 17*(10), 595-595.

Guderian, S., Dzieciol, A. M., Gadian, D. G., Jentschke, S., Doeller, C. F., Burgess, N., ... & Vargha-Khadem, F. (2015). Hippocampal volume reduction in humans predicts impaired allocentric spatial memory in virtual-reality navigation. *Journal of Neuroscience*, 35(42), 14123-14131.

Hardcastle, C., O'Shea, A., Kraft, J. N., Albizu, A., Evangelista, N. D., Hausman, H. K., ... & Woods, A. J. (2020). Contributions of hippocampal volume to cognition in healthy older adults. *Frontiers in Aging Neuroscience, 12*, 593833.

Huffman, D. A. (1952). A method for the construction of minimum-redundancy codes. *Proceedings of the IRE*, 40(9), 1098-1101.

Karl, A., Schaefer, M., Malta, L. S., Dörfel, D., Rohleder, N., & Werner, A. (2006). A meta-analysis of structural brain abnormalities in PTSD. *Neuroscience & Biobehavioral Reviews, 30*(7), 1004-1031.

Kim, E. J., Pellman, B., & Kim, J. J. (2015). Stress effects on the hippocampus: a critical review. *Learning & Memory, 22*(9), 411-416.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *ArXiv preprint* arXiv:1412.6980.

Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, 37(2), 233-243.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *PNAS, 97*(8), 4398-4403.

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus, 16*(12), 1091-1101.

Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., ... & Nadel, L. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. *Journal of Anatomy, 207*(1), 35-66.

Nobis, L., Manohar, S. G., Smith, S. M., Alfaro-Almagro, F., Jenkinson, M., Mackay, C. E., & Husain, M. (2019). Hippocampal volume across age: Nomograms derived from over 19,700 people in UK Biobank. *NeuroImage: Clinical, 23*, 101904.

Pfeiffer, B. E. (2020). The content of hippocampal "replay". *Hippocampus, 30*(1), 6-18.

Pohlack, S. T., Meyer, P., Cacciaglia, R., Liebscher, C., Ridder, S., & Flor, H. (2014). Bigger is better! Hippocampal volume and declarative memory performance in healthy young men. *Brain Structure and Function, 219*, 255-267.

Smith, B. M., Thomasson, M., Yang, Y. C., Sibert, C., & Stocco, A. (2021). When fear shrinks the brain: A computational model of the effects of posttraumatic stress on hippocampal volume. *Topics in Cognitive Science, 13*(3), 499-514.

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B* (Methodological), 58(1), 267-288.

Tonegawa, S., Pignatelli, M., Roy, D. S., & Ryan, T. J. (2015). Memory engram storage and retrieval. *Current Opinion in Neurobiology, 35*, 101-109.

Treves, A., & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus, 4*(3), 374-391.

Varghese, M., Keshav, N., Jacot-Descombes, S., Warda, T., Wicinski, B., Dickstein, D. L., ... & Hof, P. R. (2017). Autism spectrum disorder: neuropathology and animal models. *Acta Neuropathologica, 134*, 537-566.

Vivanti, G., Tao, S., Lyall, K., Robins, D. L., & Shea, L. L. (2021). The prevalence and incidence of early‐onset dementia among adults with autism spectrum disorder. *Autism Research, 14*(10), 2189-2199.

Xiao, H., Rasul, K., & Vollgraf, R. (2017). Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms. *ArXiv preprint* arXiv:1708.07747.

Zhou, P., Sense, F., van Rijn, H., & Stocco, A. (2021). Reflections of idiographic long-term memory characteristics in resting-state neuroimaging data. *Cognition, 212*, 104660.