Identify of Spatial Similarity of Electroencephalography (EEG) during Working-Memory Maintenance

Yufan Song[1], ZhiHong Zhang[1], Te Hu[1], Xiaoliang Gong[1], Asoke K. Nandi[2,1]*

1. College of Electronic and Information Engineering, Tongji University, Shanghai, China
2. Department of Electronic and Computer Engineering, Brunel University London, Uxbridge, Middlesex, UK.

*Corresponding Author: Xiaoliang Gong (gxllshsh@tongji.edu.cn)

ABSTRACT

Working memory maintenance is one of the important procedures during working memory storage into long-term memory. This paper utilizes consensus clustering to analyze the spatial similarity amongst whole brain regions during working-memory maintenance processing with 128 channels of scalp Electroencephalography (EEG) records. This paper sets the methodology research to extract the similarity of spatial information processing on the larger brain system during working memory maintenance based on a data-driven method. Based on group analysis of 20 subjects, the EEG channels with similarities are extracted to illustrate the functional brain connectivity of material-specific memory maintenance. The power of alpha frequency band (8-12Hz) appears to provide the discriminative information of the material-specific memory maintenance.

***Index Terms*—**Consensus clustering, data-driven, spatial similarity, working memory maintenance, EEG.

1. INTRODUCTION

The observations about communications between brain regions provide new insights into the researches about memory processing [1-2]. At the same time, the memory maintenance procedure plays a crucial role in transferring the working memory contents to long-term memory. The scalp Electroencephalography (EEG) is one of the neuroimaging methods to demonstrate the brain connectivity among the entire brain cortex during working memory maintenance procedure. It is different from the Neuroscience community, which are focused on the micro-mechanism during working memory maintenance [3]. EEG measurements can illustrate the collaborative work within the whole brain cortex system. Previous works always focused on event-related potential (ERP) analysis [4-5], which are focused on the 500ms data points just after each stimulation. Continuous EEG signals may reveal the continuous processing of memory maintenance. Even though the individually firing of memory maintenance makes it challenging to explore the mechanism of working memory maintenance on EEG data, there is still a great need for the methodology research.

Unsupervised machine learning is also a data-driven method to illustrate the meaningful subgroups among the reality data sets, which are complex flexible and variable [6]. Supervised machine learning algorithms were frequently used for decoding cognitive concepts based on EEG records [7]. Clustering algorithm is also a useful tool to reveal quantitative EEG analysis [8]. Consensus clustering can include the consistency between different types of algorithms and different datasets. This method is suitable for data-driven analysis of neuroimaging data. In 2004, consensus clustering was first utilized in gene-expression data analysis [9]. Consensus clustering is frequently used in the data mining field, such as gene data [10-15], and functional brain networks [16]. In 2016, Liu et al. proposed a tunable consensus clustering paradigm named 'UNCLES' to analysis functional magnetic resonance imaging (fMRI) [17, 18] which used M-N plots to evaluate the best clustering groups automatically. However, this automatic evaluation is a time-consuming procedure.

This paper utilizes 'UNCLES' tool to generate the consensus clustering by exploring the similarity of spatial information processing during materials based working memory maintenance by EEG data. To reduce the computation cost and evaluate the stability of this method based on EEG data, systematic parameter optimization is carried out. Based on group analysis, the similarity of spatial information during the same memorized materials is worked out. Finally, the difference of spatial information between sub-condition between Face vs House and Digit vs Letter have been analysed.

2. DATASET

The continuous EEG data was acquired from Sternberg tasks working memory experiment. A group of 20 healthy Germany volunteers, aged between 18 and 30 years with no history of neurological or psychiatric disorders, were used. The participants needed to memorize two categories memory materials, Face/House pictures, and Digit/Letter-pictures. All the simulation screens showed for 100ms in random order one category followed by a 4s maintenance interval. After that, volunteers needed to retrieve the

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

materials that they memorized by the following ways: the experimenters showed a picture and let volunteers choose whether or not have seen the picture. 128-channel EEG records were acquired with a sampling rate of 1000Hz. The 4s maintenance interval time windows were used in this work. About 40 trials were carried out for each sub-condition (Face/House/Digital/Letter), and then a 4x128x4000x40 tensor for one participant was constructed. At last, the dataset consisted of 20 tensors. During the signal acquisition, those signals that were too noisy were marked as bad trail. The rest good trial (about 80%) raw time-domain EEG data were used.

3. METHODS

3.1. Consensus Clustering

The procedure of this work is as the following precedures.

3.1.1. Preprocessing

Authors use the fast Fourier transform (FFT) algorithm to calculate Discrete Fourier Transform (DFT) and only the magnitude information will take into account. Transform the time domain data (of size 4*128*4000*40, meaning 4 sub-conditions, 128 channels, 4000ms, 40 trails) into frequency domain data (4*128*400*80); frequency band is ranged from 0-100Hz.

To enhance the representation of the useful information and reduce the impact of any noise, the raw EEG data is averaged from trials under the same channel. One processed sample data is a 2D matrix $D_{C\times F}$ with C rows representing channels and F columns representing frequency point.

Finally, a P×C×F 3D list is built to form the final data with P-value representing the number of the volunteers. When we use 20 volunteers' data, the value of P is 20, and get P×C partitions. "UNCLES" can extract the similarity in different clustering results after the applications of averaging and Fuzzy consensus partition matrix (CoPaM) operation, which will be introduced in 3.1.4.

3.1.2. Partition generation

Three clustering methods are used in this work to generate partition results, such as K-means [19], self-organizing (SOM) [20, 21], and hierarchical clustering (HC) [22]. R partitions are obtained through R clustering experiments. Each resulting partition P_i, for i = 1, ..., R, is a 2D matrix $U_{i:}$ (K×M), with K rows representing the cluster numbers and M columns representing the data point. The value "0/1" means whether the point belongs to this cluster or not. Using some apriori domain knowledge, some parameters are adjusted in the algorithm to control the range of values of K to obtain a more reasonable result.

3.1.3. Relabeling

Due to the stochastic nature of some clustering methods and the unsupervised nature of any clustering process, the clusters' label of partition produced during different experiments over the same dataset may not match each other, i.e., the first cluster in partition P1 may correspond to any cluster in the partition P2. This is an NP-complete combination problem named labeling correspondence problem. The min-max relabeling approach can help to solve this problem. Although it cannot make a complete match between cluster results, it can find the most consistent cluster member among results.

3.1.4. Fuzzy Consensus Partition Matrix Generation

After relabeling, authors average the value in every partition on the corresponding row and column to generate a final CoPaM where each data has its fuzzy membership value coming from the partitions.

If the input data is only one sample data, we will do only one CoPaM operation between the partition under the same parameters except for different clustering methods. If the input data has multi-sample data or pseudo-multi-sample data (one sample data repeated several times), we could get many Fuzzy consensus partitions, the number of which is equal to the number of samples. So there needs another average operation to make several fuzzy consensus partition matrixes to only one matrix. In the second CoPaM operation, it will regard the different sample's corresponding CoPaM as the new cluster results and do the same CoPaM operation as the first time.

3.1.5. Fuzzy stretch Partition Matrix Generation

After generating CoPaM, the matrix will apply fuzzy stretch operation to make it easier to binarization later. This operation can map [0,1] data nonlinearly into wider data space. The new data space is calculated by the equations:

$$x_1 = \begin{cases} \frac{\pi \times x}{2 \times x_0} - \frac{\pi}{2}, \ if\ x < x_0 \\ \frac{(x - x_0) \times \pi}{2 \times (1 - x_0)}, \ otherwise \end{cases} \quad (1)$$

$$y = \begin{cases} x_0 + x_0 \times \sin(x_1), \ if\ x_1 < x_0 \\ x_0 + (1 - x_0) \times \sin(x_1), \ otherwise \end{cases} \quad (2)$$

x is the data which can get after 3.1.4 step. Parameter x0 is the threshold. The data bigger than x0 will be bigger, and the date smaller than x0 will be smaller, which will make the fuzzy membership more separable. The value of x0 has two choices. We can fix a value through experience artificially or set x0 equals the mean of the corresponding row's value of x. x_1 can be got from equation (1). Use x_1, x0, and equation (2), y can be got. In the final matrix, the number of columns is equal to the number of data point x, the number of rows is equal to the number of clusters. Every value in every column vector means the probability of the point belongs to the corresponding cluster.

3.1.6. Binarization

The fuzzy CoPaM data is not binarized. They are the value b between [0,1]. In this way, one data point may belong to many clusters. In order to solve this problem and make sure the data points are only included in one cluster. We applied the Difference Threshold Binarization (DTB) to our data. In this method, every data point can be assigned a cluster only if it meets two conditions. The first condition: a point can only be assigned to the cluster with maximum membership value. The second condition is that the maximum membership value is bigger than the closest competitor value at the threshold δ.

$$if\ x \in Cluster\ X \begin{cases} \max(M_x)\ is\ X \\ a - b > \delta \end{cases} \tag{3}$$

(a: maximum membership value of x, b: the second maximum value of x, δ: a threshold that we give to control the strict degree of the judgment, M_x: the membership value of x, X: a cluster, x: a data point/channel)

With this determine requirements, there are some points not assigned to any cluster. The number of these points are decided by the threshold δ.

3.1.7. Clusters Extraction

After the above procedure, the Bi-CoPaM is achieved. The M-N scatter plot technique is trying to select the best final cluster results [13]. All the resulting clusters can be plotted in a 2D plotted with the horizontal axis (M) representing the average mean square error (MSE) [17] and the vertical axis (N) representing the logarithm of the number of the data point.

$$\mathrm{MSE}_{\mathrm{cluster}(k)} = \frac{1}{LN_k} \sum_{l=1}^{L} \sum_{n \in C_k} \|\mathbf{x}_n^l - \mathbf{z}_k^l\|^2 \tag{4}$$

x^l_n is the normalized signal vector of nth data point in the lth sample in the cluster, z^l_k is the average normalized signal vector of data point in the kth cluster from the lth sample, L is the number of samples, N_k is the number if data point in the kth cluster, C_k is the kth cluster.

First, the cluster which is closest to the top left corner meaning the largest cluster with the smallest MSE. In this way, we can find the cluster which not only has tight with high correlation (high horizontal axis value, the data points in the cluster are more similar), but also has a larger number of the data point (high vertical axis value). Then we remove the other clusters, which are overlapped with it. Until now, we have finished an integral process. We should do this process again to the second, the third cluster until the plot is empty. Finally, the best clusters are obtained.

3.1.8. Parameters optimization

To save the calculation cost and to achieve a stable result, based on one sample (subject NO.1), the following parameters are optimized: repeating calculated times {1, 10, 20, 100}, binarization parameter {0, 0.5, 0.7, 1}, stretch {0.1, 0.4,0.5, 0.6, 0.9}, and cluster number {4, 8, 12, 16} and compared.

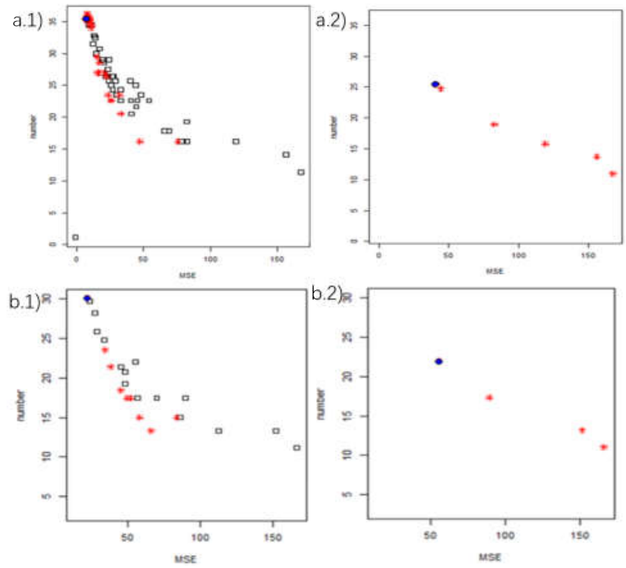

Fig. 1. The M-N plots based on one sample (subject No.1) under different parameters. a) default value (repeating calculated times = 1, binarization parameter = (0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1), stretch = average, and cluster number = (4,8,12,16); b) optimized parameters optimization (repeating calculated times = 1, binarization parameter = 0.7, stretch = 0.5, and cluster number = 4).

3.2. The Extraction of spatial information

Using the combination of optimized parameters in the UNCLES algorithm, we find the clusters which have the high likelihood of spatial similarity among the frequency signals of 128 channels of 4 sub-conditions (Face/ House/ Digit / Letter) based on 20 samples.

To figure out the different brain regions between Face and House or Digit and Letter, the channels in each cluster are compared one by one, and the different channels are recorded and plotted.

3.3. Computing resource

The consensus clustering method used in this work is implemented in the R package "UNCLES" available on http://cran.rproject.org/web/packages/UNCLES/index.html. R.3.4.4 platform is used in this work.

All the plots are produced in MATLAB 2016a. All the work in this study is calculated on Windows10 with CPU 2.80 GHz, RAM 8.0GB.

4. RESULTS AND DISCUSSION

4.1. Systematic parameters optimization of Consensus Clustering based on one sample.

This study is preliminary research to explore the similarity of spatial information processing on the whole brain system during working-memory maintenance based on a data-driven method. To save the computation cost and test the repeatability and stability of this unsupervised machine learning method, based on one sample (subject NO.1).

Optimized the following parameters: repeating calculated times, binarization parameter, stretch, and cluster number. The comparison M-N plot between default and optimized parameters are shown in Fig. 1. One can find out with the optimized parameters, the number in each consensus cluster at different stages coarse tuning (Fig. 1. a1 / b1) and fine tuning (Fig. 1. a2 / b2) are smaller which may also prove that with the optimized parameters, one can remove those isolate channels during consensus clustering. After these systematically parameters optimized of consensus clustering, a combination of parameters is achieved and used in the following 20 samples group analysis.

4.2 The interpretations of frequency information of the working memory maintenance processing based on 20 samples.

The authors superimpose the frequency spectra of channels belonging to the same cluster to try to analyze the frequency band signal characteristics of material memory in the working memory maintenance stage. The result shows in Fig. 2. For each cluster, select all channels belong to it in 20 people's data, average all the frequency signals of these channels, and plot final average spectrum of each cluster.

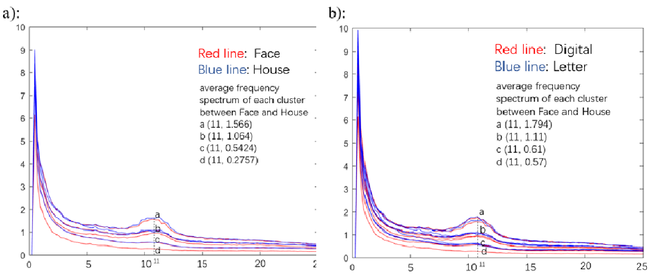

Fig. 2. The picture of the channels' clustering result on frequency domain based on 20 samples, the pictures are the average frequency spectrum of 20 people of each channel: a) Face (red line) vs House (blue line) b) Digital (red line) vs Letter (blue line)

From current results, it is easy to find the following conclusion. Firstly, there is a high degree of similarity between the clusters. For most clusters, the red line has its "corresponding" blue line which has a similar waveform, especially in alpha band. Secondly, no matter for House (Fig. 2a. Red) /Face (Fig. 2a. blue) or Digital (Fig. 2b. Red) /Letter (Fig. 2b. blue), the strength of red lines are lower than the strength of blue lines. In other words, when brain deal with the Face pictures, the low-frequency alpha signals intension is higher than when dealing with House. Similarly, the signals of Letter are higher than the signals of Digital. Thirdly, there are differences between brain area during alpha band signals: the intension of different clusters is different.

To sum up, the alpha frequency stage is important in working memory maintenance, especially for the intension of the alpha band (8-12Hz) signal which is the key to find the difference when people memory various materials. Moreover, there are regional differences when the brain deals with specific materials. Because the clusters' order of magnitude is obvious in the picture.

With the optimized parameters (repeating calculated times=1, binarization parameter=0.7, stretch=0.5, and cluster number=4.), based on 20 samples dataset, we explored the spatial information of the working memory maintenance. The scalp topography is shown in Fig.3. It clearly shows that the brain region located at the prefrontal lobe, parietal lobe, and occipital lobe are with high likelihood among 20 subjects during 4-sub-condition memory maintenance. It is interesting to find out that only the channels PPO9h, POO2, and P8 show memory maintenance of the pictures of 'House'; and only PO9, FC1, FCC2h, FCC4h, FC2, AFp1, Af4, and Fp2 shown when memorized of 'Face'. At the same time, it is found that only C1 channel is shown when memorized of 'Letter'; and only POO9h, PO7, O1, PO3, POz, PO4, FFC1h, FFC4h, FC4, and FCC4h are shown for 'Digit'. Except for the PFC and posterior regions [23], more regions are found in this work providing a complex, higher-dimensional form [7, 24].

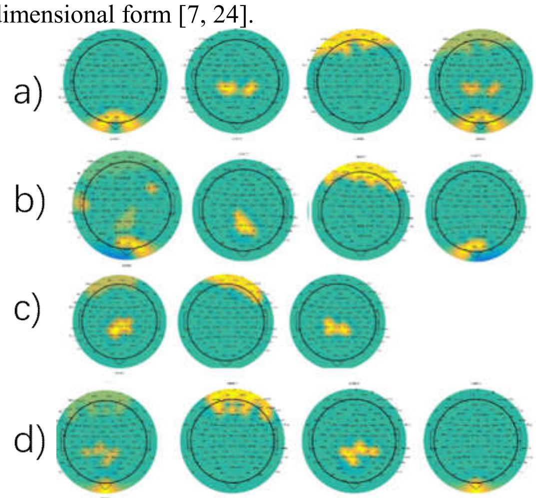

Fig.3. The similarity spatial channels calculated based on 20 samples during the working memory maintenance processing of a) Face; b) House; c) Letter; d) Digital.

5. CONCLUSION

This study used a data-driven method to extract the similarity of spatial information processing on the larger brain system during working-memory maintenance by scalp EEG records. Systematic parameter optimization was carried out and we figured out a combination of parameters suitable for this kind of EEG data analysis. Group analysis illustrated there were similar processing brain regions under the same material-specific memory maintenance. Furthermore, the results show the importance of alpha frequency band in working memory maintenance and their regional differences.

6. ACKNOWLEDGEMENTS

Professor Asoke K. Nandi is a Distinguished Visiting Professor at Tongji University, Shanghai, China. Professor Steffen Gais from Medical Psychology and Behavioral Neurobiology department of Eberhard Karls of University of Tubingen, Germany, has provided the EEG data in this work. This work was partly supported by the National Science Foundation of Shanghai grant number 16JC1401300 and the Fundamental Research Funds for the Central Universities of Tongji University (22120170043, 22120180542).

7. REFERENCES

[1] H. Eichenbaum, “Memory: Organization and Control”, *Annual Review of Psychology,* 68(1), pp. 19-45, 2016.

[2] N. Axmacher, M. M. Henseler, O. Jensen, I. Weinreich, C. E. Elger, and J. Fell, “Cross-frequency coupling supports multi-item working memory in the human hippocampus”, *Proceedings of the National Academy of Sciences of the United States of America*, 107(7), pp. 3228-3233, 2010.

[3] I. B. Zovkic, M. C. Guzman-Karlsson, and J. D. Sweatt, "Epigenetic regulation of memory formation and maintenance", *Learning & Memory,* 20(2), pp. 61-74, 2013.

[4] Y. Pan, Z. Tan, Z. Gao, Y. Li, and L. Wang, “Neural Activity Is Dynamically Modulated by Memory Load During the Maintenance of Spatial Objects”, *Frontiers in psychology,* 9, pp. 61-74, 2018.

[5] S. Ludyga, C. Herrmann, M. Mücke, C. Andrä, S. Brand, U. Pühse, and G. Markus, “Contingent negative variation and working memory maintenance in adolescents with low and high motor competencies”, *Neural Plasticity*, 2018, pp.9, (Article ID 9628787), 2018.

[6] J. H. Bolin, J. M. Edwards, W.H.Finch, and J.C.Cassady, "Applications of cluster analysis to the creation of perfectionism profiles: a comparison of two clustering approaches", *Frontiers in Psychology,* 5, Article 243, 9 pages, 2014.

[7] S. Alizadeh, H. Jamalabadi, M. Schonauer, C. Leibold, S. Gais, " Decoding cognitive concepts from neuroimaging data using multivariate pattern analysis", *NeuroImage,* 159, 449–458, 2017.

[8] C. B. Miller, D. J. Bartlett, A. E. Mullins, K. L. Dodds, C. J. Gordon, S. D. Kyle, J. W. Kim, A. L. D’Rozario, R. S. Lee, M. Comas, N. S. Marshall, B. J. Yee, C. A. Espie, and R. R. Grunstein, "Clusters of Insomnia Disorder: An Exploratory Cluster Analysis of Objective Sleep Parameters Reveals Differences in Neurocognitive Functioning, Quantitative EEG, and Heart Rate Variability. " *Sleep,* 39(1), pp. 1993-2004, 2016.

[9] S. Swift, A. Tucker, V. Vinciotti, N. Martin, C. Orengo, X. Liu, and P. Kellam, "Consensus clustering and functional interpretation of gene-expression data", *Genome Biology,* 5(11), pp. R94-R94, 2004.

[10] B. Abu Jamous, R. Fa, D. J. Roberts, and A. K. Nandi, "Paradigm of tunable clustering using Binarization of Consensus Partition Matrices (Bi-CoPaM) for gene discovery", *Plos One,* 8(2), pp. e56432, 2013.

[11] B. Abu Jamous, F. Rui, D. J. Roberts, and A. K. Nandi, “Application of the bi-copam method to five Escherichia coli datasets generated under various biological conditions”, *Journal of Signal Processing Systems,* 79(2), pp. 159-166, 2015.

[12] B. Abu Jamous, R. Fa, D. J. Roberts, and A. K. Nandi, “Yeast gene CMR1/YDL156W is consistently co-expressed with genes participating in DNA-metabolic processes in a variety of stringent clustering experiments", *J. R. Soc. Interface,* vol. 10, no. 81, 2013.

[13] B. Abu Jamous, R. Fa, D. J. Roberts, and A. K. Nandi, “Comprehensive analysis of forty yeast microarray datasets reveals a novel subset of genes (APha-RiB) consistently negatively associated with ribosome biogenesis", *BMC Bioinformatics*, vol. 15, no. 322, 2014.

[14] B. Abu Jamous, R. Fa, D. J. Roberts, and A. K. Nandi, “UNCLES: method for the identification of genes differentially consistently co-expressed in a specific subset of datasets", *BMC Bioinformatics*, vol. 16, no. 184, 2015.

[15] B. Abu Jamous, F. M. Buffa, A. L. Harris, and A. K. Nandi, “In vitro downregulated hypoxia transcriptome is associated with poor prognosis in breast cancer", *Molecular Cancer*, vol. 16, no. 105, (19 pages), 2017.

[16] A. Ozdemir, M. Bolanos, E. Bernat, and S. Aviyente, "Hierarchical Spectral Consensus Clustering for Group Analysis of Functional Brain Networks", *IEEE Transactions on Biomedical Engineering,* 62 (9), pp. 2158-2169, 2015.

[17] C. Liu, B. Abu Jamous, E. Brattico, and A. K. Nandi, "Towards Tunable Consensus Clustering for Studying Functional Brain Connectivity During Affective Processing", *International Journal of Neural Systems,* 27(02), 2016.

[18] C. Liu, E. Brattico, B. Abu Jamous, C. Pereira, T. Jacobsen, and A. K. Nandi, “Effect of explicit evaluation on neural connectivity related to listening to unfamiliar music", *Frontiers in Human Neuroscience*, DOI: 10.3389/fnhum.2017.00611, vol. 11, (13 pages), 2017.

[19] J. M. Pena, J. A. Lozano, P. L, “An empirical comparison of four initialization methods for the K-Means algorithm”, *Pattern Recognition Letters*, 20, pp. 1027-1040, 1999.

[20] M. W. Seefeldt, and J. J. Cassano, "A description of the Ross Ice Shelf air stream (RAS) through the use of self-organizing maps (SOMs)", *Journal of Geophysical Research Atmospheres,* 117, D9, 2012.

[21] T. Kohonen, “Data Management by Self-Organizing Maps”, *IEEE World Conference on Computational Intelligence: Research Frontiers*, 5050, pp.309-332, 2008.

[22] C. Goutte, P. Toft, E. Rostrup, F. A. Nielsen, L. K. Hansen, "On clustering fMRI time series", *Neuroimage,* 9(3), pp. 298-310, 1999.

[23] J. Eriksson, E. K. Vogel, A. Lansner, F. Bergstrom, and L. Nyberg, “Neurocognitive architecture of working memory”, *Neuron,* 88(1), pp. 33-46, 2015.

[24] L Nyberg, and J. Eriksson, “Working memory: maintenance, updating, and the realization of intentions”, *Cold Spring Harbor Perspectives in Biology*, 8(2), a021816, 2016.