Machine Learning Techniques and the Existence of Variant Processes in Humans Declarative Memory

Alex Frid[1], Hananel Hazan[2], Ester Koilis[3], Larry M. Manevitz[3,5], Maayan Merhav[4] and Gal Star[3]
[1]*Edmond J. Safra Brain Research Center, University of Haifa, Haifa, Israel*
[2]*Network Biology Research Laboratory, Technion, Haifa, Israel*
[3]*Computer Science Department, University of Haifa, Haifa, Israel*
[4]*German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany*
[5]*Center of Information and Neural Networks, National Institute of Information and Communications Technology, Suita, Osaka, Japan*

Keywords: Machine Learning, Classification, functional Magnetic Resonance Imaging (fMRI), Feature Selection, Support Vector Machines, Radial Basis Function Kernel, Declarative Memory, Information Biomarkers.

Abstract: This work uses supervised machine learning methods over fMRI brain scans to establish the existence of two different encoding procedures for human declarative memory. Declarative knowledge refers to the memory for facts and events and initially depends on the hippocampus. Recent studies which used patients with hippocampal lesions and neuroimaging data, suggested the existence of an alternative process to form declarative memories. This process is triggered by learning mechanism called "Fast Mapping (FM)", as opposed to the 'standard' "Explicit Encoding (EE)" learning procedure. The present work gives a clear biomarker on the existence of two distinct encoding procedures as we can accurately predict which of the processes is being used directly from voxel activity in fMRI scans. The scans are taken during *retrieval* of information wherein the tasks are identical regardless of which procedure was used for acquisition and by that reflect conclusive prediction. This is an identification of a more subtle cognitive task than direct perceptual cognitive tasks as it requires some encoding and processing in the brain.

1 INTRODUCTION

Human declarative memory is defined as the conscious information recollection of facts and events (Squire, 1992). Under the "standard model" theory for adult declarative memory systems, novel information is encoded explicitly into the memory using, amongst other brain parts, the hippocampus (McClelland et al., 1995). This standard, hippocampal dependant memory is acquired through intentional "Explicit Encoding (EE)" procedure. The encoded information is then slowly transferred from the hippocampus to the neo-cortex where it becomes permanently stored (Squire and Alvarez, 1995; Frankland and Bontempi, 2005). Overtime, the initially hippocampal dependant memories become independent of the hippocampus. It has been suggested that this re-organization process is done during sleep (Gais et al., 2007).

Amongst toddlers, the process of rapid language acquisition occurs prior to the full development of the hippocampus (Bauer, 2008; Uematsu et al., 2012). Moreover, some evidence from hippocampal injured subjects demonstrated an ability to acquire information which seems to have declarative-like characteristics despite severe damages in the hippocampus (Sharon et al., 2011; Merhav et al., 2014) and so must involve a different brain network than the one engaged by "EE". This alternative learning mechanism is called "Fast Mapping (FM)". It is unknown if the memory representations following FM undergo consolidation processes, similar to memories gained through EE. However, since it was shown that patients with hippocampal damages as well as healthy controls could learn and store information acquired via FM for a week (Sharon et al., 2011; Merhav et al., 2014), the scheme used to explain memory consolidation of other declarative memories cannot be applied for FM in a straightforward manner.

It remains somewhat controversial as to whether the FM is available for acquisition of words among amnesic patients (Warren and Duff, 2014). Proving that FM methods are mostly based on brain

Frid, A., Hazan, H., Koilis, E., Manevitz, L., Merhav, M. and Star, G..
Machine Learning Techniques and the Existence of Variant Processes in Humans Declarative Memory.
In *Proceedings of the 7th International Joint Conference on Computational Intelligence (IJCCI 2015) - Volume 3: NCTA*, pages 114-121
ISBN: 978-989-758-157-1

structures outside of the hippocampus area opens possibility for therapeutic approach for people with damages in these areas.

In this work we aim to demonstrate the distinctiveness of brain systems, which support EE and FM memory process, by extracting activity patterns directly from brain data. Functional magnetic resonance imaging (fMRI) captures information from thousands of different localities (voxels) of the brain simultaneously. Multivariate pattern analysis approach (MVPA) (Norman et al., 2006) utilizes these activities by looking for changes in BOLD signal across different voxels. Different methods can be used for analysis on such complex data depending on the question of study (retrieval or decoding stimuli, mental states, behaviours and other variables of interest). A growing number of studies (Mitchell et al. 2008; Kriegeskorte et al., 2006; Nawa and Ando, 2014; Atir-Sharon et al., 2015) shows ability in using machine learning methods for analysis of neuroimaging data. Nevertheless, the feasibility to achieve successful results using machine learning on fMRI multivariate data is not trivial and relies on the sensitive choice of features to be considered in the analysis.

2 RELATED WORK

The mechanism of FM was examined among healthy individuals (Gilboa et al., 2011; Atir-Sharon et al., 2015). It was shown that two learning mechanisms, EE and FM, can be discriminated from fMRI data during memory acquisition using machine learning based classifier. In addition, memories acquired during scanning were tested for recollection success later, outside the fMRI machine. Successful accuracy results were achieved when identifying scans corresponding to successful and unsuccessful recollection within EE group and within FM group, for each participant separately and cross-participant.

However, the different nature of the procedures used for acquisition of information (EE and FM), does not allow for complete control over the task with regard to the behavioural experience. Therefore, the possibility remained that the successful classification obtained in the experiment is a result of differences in the acquisition procedures and not in the learning mechanisms.

To overcome this limitation, in another study (Merhav et al., 2015), the neural correlates of FM and EE were explored during a retrieval procedure, designed to be identical for both mechanisms. In addition, the study was focused on overnight re-organization of memory representations, following both EE and FM. Findings suggested that, despite the identical retrieval tasks, memories that were gained through FM induced distinct neural substrates from those involved EE (Merhav et al., 2015). While retrieval of data learned through EE engaged the expected hippocampal and vmPFC related network, retrieval of information acquired through FM immediately engaged an ATL related network, typically supporting well-established semantic knowledge. In addition, analysis of neuroimaging data associated with EE showed the expected overnight changes in network connectivity where for FM minimal overnight changes were presented. The analysis was performed by a multivariate technique of Spatiotemporal Partial Least Squares (PLS), helping to identify assemblies of brain regions that co-vary together.

3 CURRENT STUDY

In this study, fMRI brain data was captured during retrieval of memories, acquired through either EE or FM. The goal is to provide a biomarker directly from these fMRI scans using machine learning methods. Such classification ability based on the neural activity data gives strong evidence for the existence of distinct neural processes associated with EE and FM.

Multivariate classification is performed on fMRI features obtained during memory recollection, where tasks performed by the participants are identical for EE and FM. We also perform classification to explore re-organization processes following both learning mechanisms. Classification was performed over brain scans which were acquired either 30 minutes before scanning (recent memory) or a day before scanning (remote memory).

Regarding the distinction between the two memory processes during recollection, we address two questions:

1. Is it possible to distinguish between the two learning modes (i.e. EE and FM) based on neural activity information collected during the recollection of memories?
2. Is it possible to distinguish between items learned recently and remotely?

4 EXPERIMENT PROCEDURE

4.1 Participants

The experiment, full details of which can be seen in Merhav et al. (Merhav et al., 2015), was conducted in Rotman Research Institute at Baycrest, Canada. Here we mention the salient points.

32 participants (20 females) were recruited and randomly assigned to one of the two groups (EE or FM). All participants were English native speakers, right-handed and had no history of neurological or psychiatric disorders and no learning disabilities. A written informed consent was obtained according to Baycrest's Research Ethics Board's guidelines. Gender and age distributions (10 females in each group) were similar in the FM and in the EE groups, respectively. The two groups also did not differ on the number of years of education, I.Q. estimates and WMS-III Verbal Paired Associates retention.

4.2 Experiment Paradigm & Procedure

32 healthy adult participants (20 females) were randomly assigned of one to two groups (EE or FM). On day 1 the participants learned 50 new unfamiliar picture-word associations. On day 2 (24 hours later) they learned another set of 50 new picture-word associations. A retrieval memory test for all the 100 new picture-word associations took place 30 minutes after the acquisition of a second set of associations. During the retrieval, brain activity was scanned (Figure 1A). Therefore, the participants were tested on both recently and remotely encoded information. The two learning tasks (EE / FM) were designed differently due to different nature of both learning procedures (Figure 1B).

The retrieval task was designed as an event related fMRI experiment in which memory for all 100 items was assessed via an associative four-alternative forced choice recognition task. The retrieval procedure was identical for EE and FM as it was performed inside the scanner (Figure 1C). Each retrieval trial of an item was 12.5 seconds long and contained the following intervals: blank screen (1 sec), target label as text and auditory input (1.5 sec), 4 choice pictures appeared on screen, below the target label (2.5 sec), the word "choose" appeared onscreen and participants had to respond by selecting the appropriate key (5 sec), confidence rating (2.5 sec).

The experiment was designed intentionally to have participants perform either EE or FM, rather than perform both EE and FM tasks. It was important that learning through FM will be implicit and unintentional, so participants should not know that the task is a mnemonic task (i.e., requires memory). However, in EE, participants are explicitly asked to remember the name of the item.

4.3 Data Acquisition & Pre-processing

The participants were scanned using the Siemens Trio 3 T scanner, at Baycrest Institute. They acquired T2*-weighted images, covering the whole brain using an echo-planar imaging (EPI) sequence of 50 slices, with repetition time (TR) of 2500 ms, echo time (TE) of 27 ms, 64 × 64 matrix, slice thickness of 3.5 mm and a field of view (FOV) of 200 mm. The procedure was designed as an event related fMRI study.

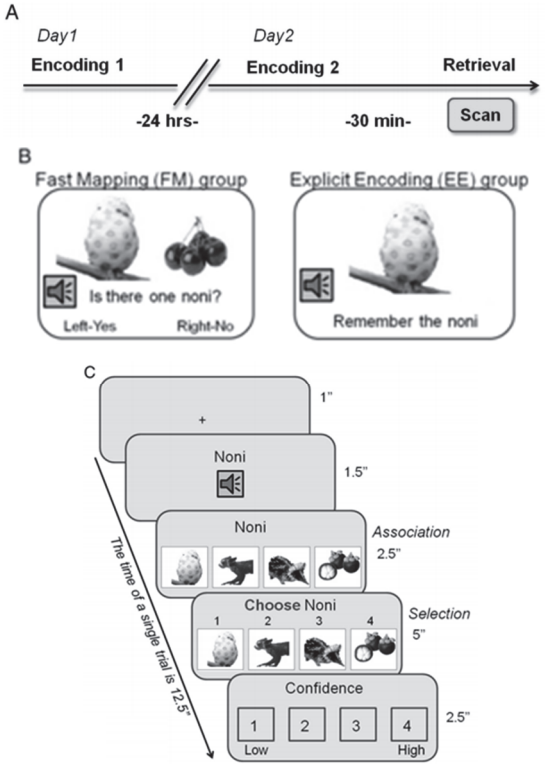

Figure 1: (A) The experiment structure. (B) Examples of acquisition through FM (left) and through EE (right). (C) Retrieval test design which took place inside the fMRI scanner.

The pre-processing steps included conversion to 4-dimensional AFNI format (Cox, 1996), slice timing correction using the first slice as a reference (Figure 2A), movement correction for unintended head motions and spatial smoothing with 6mm FWHM Gaussian kernel to increase signal-to-noise ratio (Figure 2B). Finally, individual participant's

data was converted to a standard coordinate system (Talairach) to allow data analysis across individuals.

The scanning of each participant was done during four consequent runs creating a joint dataset out of four time-series datasets with approximately 150 data volumes each of size 109x91x91, resulted as a dataset with approximately 600 data volumes. Therefore, each data volume (data point) contained 1490580 different voxels. We demonstrate the structure of the collected data in Figure 3.

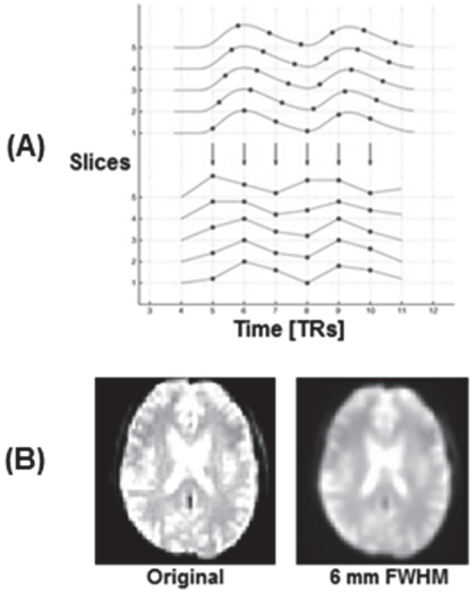

Figure 2: Examples for pre-processing steps on fMRI data. (A) Correction of individual's hemodynamic responses slices acquired aligned to the exact same time (Sladky et al., 2011). (B) Performance of spatial smoothing on fMRI volume taken from single participant.

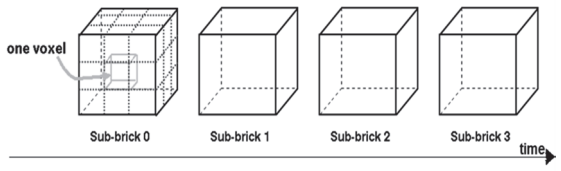

Figure 3: 4-Dimensional structure of AFNI format BRIK (Cox, 1996) file including 3-dimensional dataset over time sequence.

5 METHODS

The data points used for analysis were constructed using scan data obtained for TR=2. This temporal cut was selected after performing pre-test classification as suggested in Atir-Sharon et al. work (Atir-Sharon et al., 2015), taking into consideration the accordance to the expected HRF response.

We performed further pre-processing over the time-series data. At first, all non-brain voxels were removed using a mask. This was done by selecting voxels from the fMRI dataset that correspond to non-zero elements in the mask (creating data points of approximately 200,000 voxels). Afterwards, linear detrending was performed on each participant's data set and for each run separately.

Then, normalization over all scans was conducted. The normalization was done voxel-wise using z-score for each participant separately. In our case, the combined dataset involved scans from different groups and participants taken from different distributions. Therefore, transformation of features from different scales to a single scale, with consideration to the original distributions, was needed. The z-score method considers the different distribution characteristics of every group (Wiesen, 2006), hence, it was chosen as the normalization procedure. The z-score formula is presented in (1), where z-val is the new z-scored value, f-val is the original feature value and (μ, σ) are the mean and standard deviation values:

$$z\text{-val} = (f\text{-val} - \mu) / \sigma \quad (1)$$

For the mean and standard deviation computation in the z-score equation, several assignments were tested: (i) from all scans in the dataset; (ii) from individual participants' scans and (iii) from the distribution of scans marked as control (baseline) in the training set. Best classification results were achieved by using the mean and standard deviation computed from the distribution of baseline scans (option (iii)).

Each volume was represented as an individual data point in the dataset (i.e. each voxel was considered as a feature). Since the amount of scans from EE and FM groups was not equal, counter-balancing of the dataset was performed. This was done by randomly sampling data points from the smaller group. This method was applied only on the training set. Otherwise, more weight would have been given to prediction accuracies of duplicated data points against weight of accuracies for data points that were not duplicated. Therefore, testing set was left untouched.

Machine learning classification techniques were used for data analysis. Considering the high dimensionality of data used in the current study, feature selection procedure was performed in order to reduce the number of features used for multivariate classification analysis. There are several generic methods for selecting informative features.

We aimed to select the features that best discriminate between conditions based on their activation values. It was achieved by ranking the importance of each feature according to the ANOVA F-score value obtained for between-group (EE vs. FM) comparisons.

To find the optimal subset of features for analysis, we performed exhaustive search for different sizes of features sets starting from 10 features to full brain features in exponential manner. Finally, the top 1000 features with highest F-scores were selected. This relatively large number of features was chosen to take advantage of inclusion of weakly informative voxels which can contribute to an increase in classification rates (Gonzalez-Castillo et al., 2012). In Figure 4, we illustrate the extracted features in the form of a brain map. In this example, we display in red selected subset of features for recollection (correct vs. Incorrect) classification. This was performed on individual's fMRI data that belongs to the FM group. Note, that not all the selected features can be depicted in a single brain map, but it can be seen that they concentrated in a specific areas.

A cross-validation classification scheme using Support Vector Machine classifier (Vapnik, 1998) with RBF (Radial Basis Function) kernel (Vert, Tsuda and Schölkopf, 2004) was applied to the selected features.

Parameters that are not learnt directly within estimators can be set by searching a parameter space for the best cross-validation score. Grid search for C and gamma parameters was performed in the ranges of 2^{-5} to 2^{15} and 2^{-15} to 2^{3} respectively. Grid search was executed before training on a training portion of the dataset to achieve increase in accuracy rates. A pseudo-code for the performed grid search is presented in Figure 5. In all runs parameters C and gamma were set to 1 and 2^{-3} respectively.

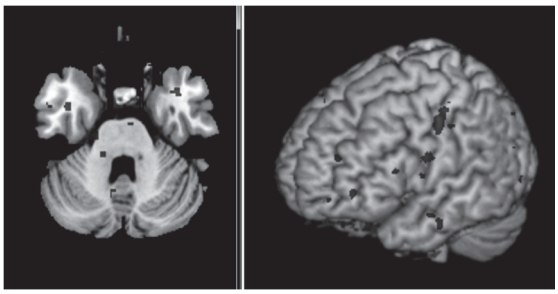

Figure 4: Brain map displaying features selected for classification analysis of FM recollection (correct vs. Incorrect).

In cases where the testing set consisted of scans that were taken from one group only (i.e. all scans were EE or all scans were FM), a decision making function was applied. We used majority voting method as a decision making function, defined as follows: if the majority of the scans were rated correctly per participant, the accuracy was set to 1, otherwise, the accuracy was set to 0.

```
for c in [2^-5, 2^-3,...,2^15]:
 for g in [2^-15, 2^-13, ..., 2^5]:
  for train, test in partition:
    model = svm_train(train, c, g)
    score = svm_predict(test, model)
    cv_list.insert (score)
  scores_list.insert(mean(cv_list),c,g)
print max(scores_list)
```

Figure 5: Pseudo-code for grid search procedure.

The software used for the classification was developed using Python programming language and based on LibSVM (Chang and Lin, 2011) and PyMVPA software packages (Hanke et al., 2009). In Figure 6 we present a complete analysis flow diagram including all the relevant pre-processing and processing stages.

6 RESULTS

6.1 Memory Performance

In the information retrieval test, correct response rates for the recent and for the remote associations were significantly above chance (binomial tests, $p < 0.0001$, for both times-of-acquisition, in both learning groups). Overall, participants from the FM group were less successful in retrieval, compared to those from the EE group, in both the recent and the remote conditions ($F(1,30) = 12.2$, $p < 0.005$).

In both groups, recent items were better recognized than remotely presented items ($F(1) = 9.12$, $p = 0.005$) with no significant interaction between the time of acquisition and the learning mode ($F(1,15) = 0.334$, $p = 0.565$).

6.2 Classification

First, we addressed the question of classifying scans obtained during correct and incorrect recollection. Using the proposed classification scheme, we performed 4-fold (leave one run out) cross-validation within participants. The mean values of classification accuracy were close to the chance level for both groups (EE and FM). We theorized the

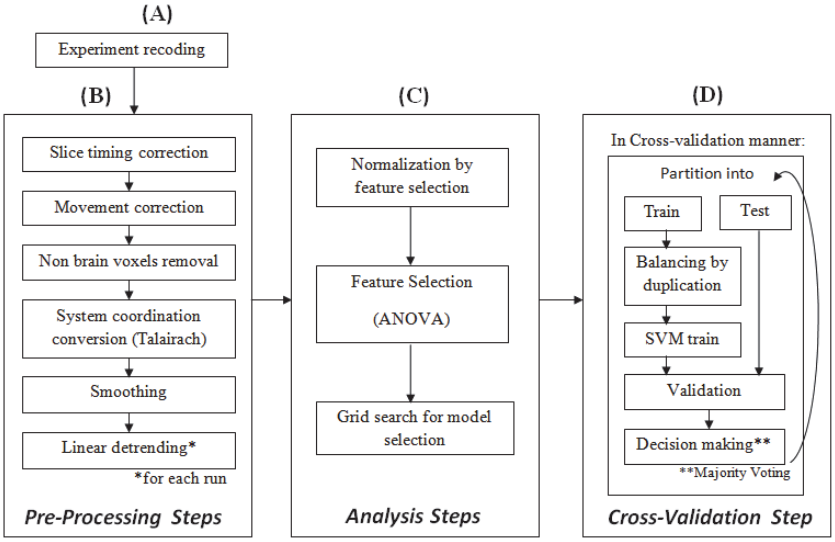

Figure 6: Schematic diagram of the steps performed for whole brain analysis procedure. It consist the following stages: (A) The initial stage representing the neuroimaging data delivery. (B) The pre-processing stage. (C) Data reduction stage: reducing data variability efficiently by feature selection. (D) Learning stage: performing multiple times by cross validation procedure.

reason for that is the existence of two additional different sub-groups, recent and remote word acquisition, within each of the initial groups. Therefore, we classified correct and incorrect scans within each possibility: EE recent, EE remote, FM recent, FM remote. For each possibility we chose 10% of all data points randomly as a testing set. The rest of the data points were used for training. Then, 10-fold cross validation was performed. We report the values for mean and standard deviation of classification accuracy over 10 cross-validation folds for EE in Table 1 and for FM in Table 2.

These results show that a trained classifier was able to distinguish scans obtained during correct and incorrect word recollection within each group. The accuracy is higher for classification of scans for words learned recently, rather than for words learned remotely. Furthermore, the discriminating ability is better within EE group rather than within FM group.

Next, we classified whether the process used for information acquisition was EE or FM using only scans from the successful recollection attempts in the behavioural experiment. We chose randomly 10% from all the scans of all participants as a testing set. The rest of the scans were used as a training set. The values and standard deviations for classification accuracy are presented in Table 3. The results show that using the neuroimaging data from each one of the participants for training, we could distinguish between EE and FM scans very well.

Table 1: Correct vs. Incorrect classification within Explicit Encoding (EE) using 10-fold cross validation.

	Mean Accuracy	**Standard Deviation**
Recent	0.708	0.09
Remote	0.584	0.067

Table 2: Correct vs. Incorrect classification within Fast Mapping (FM) using 10-fold cross validation.

	Mean Accuracy	**Standard Deviation**
Recent	0.599	0.063
Remote	0.55	0.068

Table 3: EE vs. FM (using only correct recollection scans) across participants using 16-fold cross-validation.

Testing set selection method	**Mean Accuracy**	**Standard deviation**
Random selection	0.937	0.069
Leave one participant out	0.638	0.07

These results raise the question of whether the representation of all the participants in the training set is crucial to the classification success. That is, can a machine learning classifier, trained over the collected data, can successfully distinguish which label to assign to a new person scan, despite the fact that the classifier has never seen data from this participant. To answer this question, we performed a leave-one-participant-out classification. This was

done across all 16 participants in a cross-validation manner (leave one participant out). Note that per iteration, the scans in the testing set are all EE or all FM. Therefore, we were able to use the majority voting method for this analysis. The results averaged across all participants presented in Table 3.

7 DISCUSSION & CONCLUSIONS

In this work, we showed that it is possible to identify correct and incorrect recollection of memories acquired through two learning mechanisms: either Explicit Encoding (EE) or Fast Mapping (FM) directly from neuroimaging data using machine learning techniques. The findings suggest that it is easier to identify recollection success and failure for information acquired recently rather than for information after a period of time through EE mechanism. It may indicate that the newly gained information, acquired through EE, has started to take part in consolidation process. At the same time, no significant change between recollection results of recent and remote acquisition was seen within the FM mechanism. This may indicate that FM does not engage consolidation processes. Further classification experiments are required to reach a more general conclusion.

The current results provide additional evidence for the existence of two memory formation processes by successfully classifying scans of correct retrievals following EE and FM. Note that the classification results for scans taken from an individual's data, which were not used previously for training, were still significant (although less accurate when training data from a subject were included). These findings suggest that associative learning through FM employs alternative neural pathways to acquire declarative knowledge, which bypasses the dominant hippocampal-vmPFC axis. This also indicates that the FM process is eligible for therapeutic approach for people with hippocampal brain injuries.

8 FUTURE WORK

Future work should include mapping of the brain regions and extraction of functional networks associated with all four group combinations, EE recent, EE remote, FM recent and FM remote. A list of possible implementation approaches includes constructing brain maps using "searchlight" techniques (Kriegeskorte et al., 2006).

In addition, future work should include brain regions correlations tests during the retrieval of memory through EE and through FM in recent and in remote modes. Those correlations would provide information regarding the involvement of the hippocampus and vmPFC regions in the consolidation processes. To achieve that, one may use causality analysis techniques (Hu & Liang, 2012) to reveal the causality influences the brain regions, which are involved with each learning procedure, have on each other. This could help reveal new information regarding the mechanism involved in memory consolidation processes of FM.

ACKNOWLEDGEMENTS

This work is part of the M.Sc thesis of Ms. Gal Star at University of Haifa under the supervision of Prof. Larry Manevitz at the Neuro-Computation Laboratory at Caesarea Rothschild Institute (CRI), Haifa, Israel.

The research is based on data gathered by Rotman Research Institute at Baycrest, Toronto, Canada. The examining of this data was suggested by Dr. A. Gilboa and complements the work of Merhav, Karni and Gilboa (Merhav et al., 2015). The authors are listed in alphabetical order.

REFERENCES

Atir-Sharon, T., Gilboa, A., Hazan, H., Koilis, E. & Manevitz, L. M. 2015. "Decoding the formation of new semantics: MVPA investigation of rapid neocortical plasticity during associative encoding through Fast Mapping.". *Neural Plasticity*, vol. 2015, Article ID 804385, 17 pages.

Bauer, P. J., 2008. "Toward a neuro-developmental account of the development of declarative memory". *Dev Psychobiol*, vol. 50, no. 1, pp. 19-31.

Chang, C. C. & Lin, C. J., 2011. "LIBSVM: a library for support vector machines". *ACM Transactions on Intelligent Systems and Technology*, available from: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

Cox, C., 1996. "AFNI: software for analysis and visualization of functional magnetic resonance images", *Computers and Biomedial Research*, vol. 29, pp. 126-173.

Frankland, P. W., and Bontempi, B., 2005. "The organization of recent and remote memories". *Nature Review: Neuroscience*, vol. 6, pp. 119-130.

Gais, S., Albouy, G., Boly, M., Dang-Vu, T.T., Darsaud, A., Desseilles, M., Rauchs, G., Schabus, M., Sterpenich, V., Vandewalle, G., Maquet, P., Peigneux, P., 2007. "Sleep transforms the cerebral trace of

declarative memories". *Proceedings of the National Academy of Sciences of the USA,* vol. 104, no. 47, pp. 18778-18783.

Gilboa, A., Hazan, H., Koilis, E., Manevitz, L. and Sharon, T, 2011. "Two memory systems: identifying human memory encoding mechanisms from psychological fMRI data via machine learning techniques". *Proceedings of the International Joint Conference on Neural Networks (IJCNN),* pp. 54.

Gonzalez-Castillo, J., Saad, Z.S., Handwerker, D.A., Inati, S.J., Brenowitz, N., Bandettini, P.A., 2012. "Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis". *Proceedings of the National Academy of* Sciences, vol. 109, no. 14, pp. 5487-5492.

Hanke, M., Sederberg, P. B., Hanson, S. J., Haxby, J. V., and Pollmann, S., 2009. "PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data". *Neuroinformatics*, vol. 7, no. 1, pp. 37-53.

Hu, S. & Liang, H., 2012. "Causality analysis of neural connectivity: New tool and limitations of spectral granger causality". *Neurocomputing*, vol. 76, no. 1, pp. 44-47.

Kriegeskorte, N., Goebel, R., and Bandettini, P., 2006. "Information-based functional brain mapping". *Proceedings of National Academy of Science USA*, vol. 103, no. 10, pp. 3863-3868.

McClelland, L., McNaughton, B. L., and O'Reilly, R. C., 1995. "Why there are complementary learning system in the hippocampus and neo-cortex: insights from the successes and failure of connectionist models of learning and memory". *Psychological Review,* vol. 102, no. 3, pp. 419-457.

Merhav, M., Karni, A. and Gilboa, A., 2014. "Neocortical catastrophic interference in healthy and amnesic adults: A paradoxical matter of time". *Hippocampus,* vol. 24, no. 12, pp. 1653-1662.

Merhav, M., Karni, A. and Gilboa A., 2015. "Not all declarative memories are created equal: fast mapping as a direct route to cortical declarative representations". *Neuroimage*, vol. 117, pp. 80-92.

Mitchell, T., Shinkareva, S., Carlson, A., Chang, K. M., Malave, V. L., Mason, R. and Just M. A., 2008. "Predicting human brain activity associated with the meanings of nouns". *Science*, vol. 320, no. 5880, pp. 1191-1195.

Nawa, N. E. & Ando H., 2014. "Classification of self-driven mental tasks from whole-brain activity patterns". *PLos One*, vol. 9, no. 5, e97296.

Norman, K. A., Polyn, S. M., Detre, G. J. & Haxby, J. V., 2006. "Beyond mind-reading: multi-voxel pattern analysis of fMRI data". *Trends in cognitive science*, vol. 10, no. 9, pp. 424-430.

Sharon. T., Moscovitch, M., and Gilboa, A., 2011. "Rapid neocortical acquisition of long-tem arbitrary associations independent of the hippocampus". *Proceedings of the National Academy of Science of the USA,* vol. 108, no. 3, pp. 1146-1151.

Sladky, R., Friston, K. J., Tröstl, J., Cunnington, R., Moser, E. & Windischberger, C., 2011. "Slice-timing effects and their correction in functional MRI". *Neuroimage*, vol. 58, no. 2, pp. 588-594.

Squire, L. R., 1992. "Declarative and non-declarative memory: multiple brain systems supporting learning and memory", *Journal of Cognitive Neuroscience,* vol. 4, no. 3, pp. 232-243.

Squire, L. R., and Alvarez, P., 1995. "Retrograde amnesia and memory consolidation: a neurobiological perspective". *Current Opinion in Neurobiology*, vol. 5, no. 2, pp. 169-177.

Uematsu, A., Matsui, M., Tanaka, C., Takahashi, T., Noguchi, K., Suzuki, M. and Nishijo, H., 2012. "Developmental trajectories of amygdale and hippocampus from infancy to early adulthood in healthy individuals". *PLos One*, vol. 7, no. 10, e46970.

Vapnik, V., 1998. "Statistical learning theory". New York, NY: Wiley.

Vert, J. P, Tsuda, K. and Schölkopf, B., 2004. "A primer on kernel methods". *Kernel Methods in Computational Biology.*

Warren, D. E. and Duff, M. C., 2014. "Not so fast: Hippocampal amnesia slow word learning despite successful fast mapping". *Hippocampus*, vol. 24, no. 8, pp. 920-933.

Wiesen J. P., 2006. "Benefits, Drawbacks, and Pitfalls of z-Score Weighting". *30th Annual IPMAAC Conference.* Available at: "http://annex.ipacweb.org/library/conf/06/wiesen.pdf" (27 Jun 2006).