Different influences of visual associative memory on the neural activity in the human ventral occipital cortex

Hyo Woon Yoon[1]†   Hyomin Park[2]

[1]Department of Art Psychotherapy, Daegu Cyber University
[2]Center for Governance Studies at Sungkyunkwan University

Specific responses in the fusiform face area (FFA) and the parahippocampal place area (PPA) presumably underlie visual object recognition, with the relative strength of neural activity coding for the object class. However, the issue of whether these areas rest on fundamentally distinct computational processes is an unexplored area. In this study, we examined whether object-class specificity is modifiable by the long-term learning of tasks that involve faces or buildings. The experiment consisted of four fMRI sessions over eight weeks. Eleven participants were involved in two novel tasks in which faces were arbitrarily associated with two-dimensional coordinates, and buildings were associated with architectural styles. As the learning progressed, BOLD responses at the FFA significantly decreased in the associative learning tasks regardless of the object class, while no such change was observed for brain activity at the PPA. These findings suggest that the FFA plays a role in visual associative memory. As the associations grow more efficient as the result of repeated experience, the neural activity required for retrieval may decrease in strength and/or in duration, leading to less BOLD activation.

Key words : FFA, PPA, long-term learning, visual associative memory, neural circuitry changes

† Corresponding Author (hyowoonyoon@dcu.ac.kr) : Department of Art Psychotherapy, Daegu Cyber University, Kyungsan, Kyungbuk, Korea

The ability to perceive objects and sort them into categories is one of the most important features of the human visual system. In particular, the ventrotemporal cortex in conjunction with other brain regions is thought to be strongly involved in category-specific representations (Aguirre et al., 1998; Kanwisher et al., 1997). The anterior portion of the fusiform gyrus (fusiform face area: FFA) is known as the face-selective area, and it is here that face-like objects are predominantly processed (Kanwisher et al., 1997; Kanwisher & Yovel, 2006). Another well-known category-specific area near the FFA is the place-stimulus-selective parahippocampal place area (PPA) (Epstein, 2008; Epstein et al., 1998). The PPA responds preferentially to photographs of landscapes and rooms rather than to faces or face-like objects (Epstein, 2008). This categorization process may result from learning and/or may be dependent upon experience. For example, a person who has experience in a particular field, such as a bird expert, can distinguish two different types of birds easily; while the same task is more challenging for a less experienced person (Gauthier et al., 1999; McGugin et al., 2014; Van den Linden et al., 2008; Van Turennout et al., 2000). Repetition suppression refes to an attenuation of neural activation caused by repeated presentation of a stimulus (Henson, 2003). Whereas a number of studies (Grill-Spector et al., 2006) postulate that repetition suppression underlies the bottom-up process (e.g. neural facilitation due to repeated exposure of stimuli), the predictive coding model emphasizes the top-down role of repetition suppression (Egner et al., 2010). According to this approach, the expectation (correct prediction) of upcoming stimulus influences a modulation process prior to the top-down-wise cognition. Thus, the repeated stimulus presentation corresponds to correct or low error prediction, which causes the attenuated neural activity (Mayrhauser et al., 2014). Moreover, Tsakiris and Apps (2013) provide evidence that the neural activity in the fusiform face area varies with the prediction error parameter that updated facial familiarity, which leads to the conclusion that facial familiarity in a computational model derived from predictive-coding principles. However, little is known about the exact neural mechanisms underlying such category-specific function as well as learning-dependency in the human ventral stream.

Here, we aimed to investigate the neural mechanism underlying learning / experience dependency of category-specific areas in the ventrotemporal cortex. We hypothesized that the FFA and PPA might respond differently when exposed to a repeated visual learning task. More specifically, we suggested that two ventral stream areas demonstrating similar competency

with categorization may have differential computational properties with regards to their experience-dependent adaptability.

In the present study, we used real images of faces and buildings as visual stimuli. Participants in the experiment were exposed to two different learning processes. First, the participants were given approximately eight weeks to learn specific information about several well-known faces. The information they learned was not directly related to the faces' unique characteristics such as identity, expression, viewpoint etc. In the other case, participants were instructed to learn the shapes of buildings by referring to predefined architectural types, which also includes non-direct relation with the buildings' own unique characteristics. Our approach to investigate the learning effect using non-manipulated real images was aimed at observing the learning effect utilizing objects which had already been categorized before the study began. This data acquisition of four times was aimed to observe the neural response changes dependent on the learning period.

Investigating the training effect using non-manipulated real images is novel, since most of the previous studies have focused on investigations of neural plasticity effects due to intensive learning of novel / unknown objects, even using morphing procedures. Plasticity effects due to such categorization related training is observed in the occipitotemporal

region among others (Jiang et al. 2007; van den Linden et al. 2008). Using real images of faces and buildings for investigating training effect related neural mechanism has two main reasons. Firstly, the presentation of real object images allows to functionally localize the respective category specific areas such as FFA and PPA. Some literature evidences show that there are neural response changes due to more exposure of same objects or more learning of novel objects in occipitotemporal region (van den Linden et al., 2008; van Tourennout et al., 2000; Weisberg et al., 2007), but no one observed neural responses involved in learning mechanisms in exact localized category specific areas such as FFA and PPA. Secondly, it is meaningful to observe and compare the neural responses due to long term training in face selective and place selective areas. It is also valuable to observe the neural responses of non-preferred stimuli in these two distinct areas how they are behaving during the learning period. Simple working memory task for recognition of novel faces and buildings were done for four times to compare between intensive learning and simple repeated visual exposure.

Materials and Methods

Participants Fifteen right-handed participants (mean age: 27 years, S.D.: 2.5 years, six

females) participated in the study. They had normal or corrected-to-normal vision; they did not have any medical, neurological or psychiatric illness at past or present; and they did not take medication. They were all members (research fellows) of the institute and had worked in the institute for two years or more.

Each subject participated in the study four times during eight weeks for face and building judgment tasks (learning experiments). All subjects consented to the protocol approved by the Institutional Ethics and Radiation Safety Committee. Four participants were excluded from the analysis because they did not fulfill the expected learning sessions.

Stimuli   Stimuli used for the long-term learning experiment were pictures either of faces (frontal views with neutral expressions) of the participant's colleagues or of "western classical" buildings (e.g. the Duomo in Florence, Italy) surrounded by random natural backgrounds. The pictures were black and white. Pictures of 16 different identities of frontal faces of internal members and 16 western classical buildings were presented during the learning process. For the working memory task, the pictures of new faces and buildings with surrounding backgrounds were used as stimuli. The measurement of working memory task was aimed to localize each participants' function oriented FFA or PPA. The stimuli were presented using custom-made

software running on a PC and projected via an LCD projector onto a screen at the head of the participants (rear projection). The participants viewed the screen via a homemade reflection mirror attached on the head with an RF coil.

Experimental Design   There were two types of experiment; the long-term learning task and working memory task (as localizer scan). In the experiments, all the pictures of faces which were used as stimuli were very familiar to the participants because the participants had been working in the same institution with the persons in the pictures for a while. During the presentation, the participants were asked to recall spatial or specific information about each individual's face. In order to provide this information, each face was put into specific position in a 4 x 4 matrix (see figure 1 a), so that each of the 16 faces were assigned to different 2D coordinates. The subjects were asked to judge by pressing buttons on a key pad with four buttons. The location in the 4 x 4 matrix of one face was given by pressing two buttons in sequence (e.g. serial pressing of the first and third button corresponds to coordinates of 1 x 3). Subjects practiced this button pressing procedure using four fingers for 30 minutes before going into the scanner in order to familiarize themselves with this motor activity.
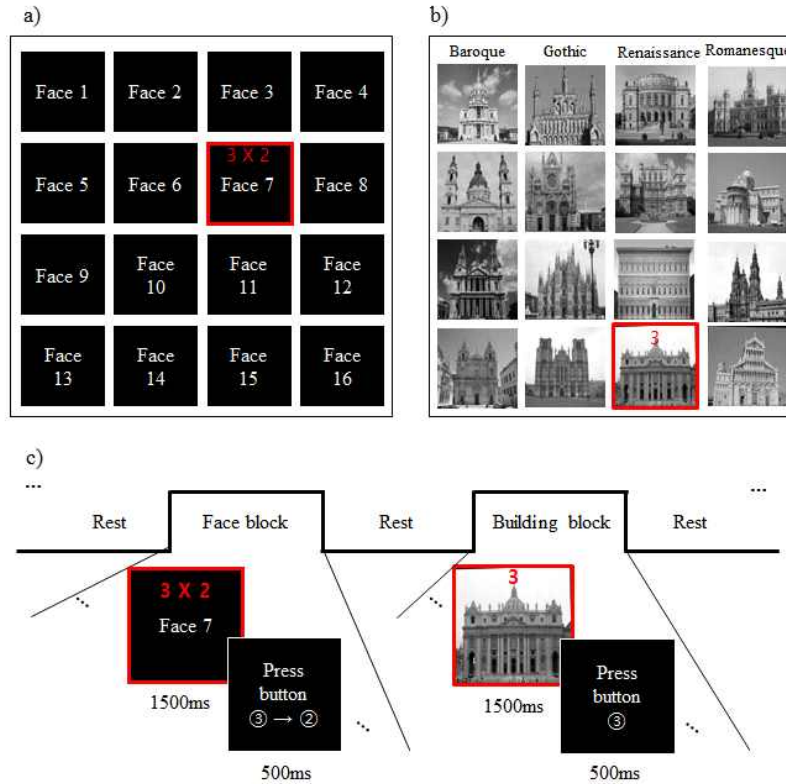
In the building judgment, the 16 different

Figure 1. Types of stimuli for intensive learning task. Sixteen highly familiar faces a) and novel buildings of four architectural categories b) were put into a 4 x 4 matrix. Subjects had to intensively learn and memorize each facial identity's and building matrix positions during a period of eight weeks before conducting fMRI experiment. c) Subject were instructed to press two buttons which were serial matrix position numbers, sequentially in face blocks and press a single button which was same number among four architectural categories in building block.

building pictures, which were categorized according to architectural types (i.e., gothic, Renaissance, baroque and Romanesque), were used as stimuli. Subjects were asked to judge the architectural style of each presented building by pressing appropriate button (i.e. the first button on the key pad for gothic, the second for Renaissance, etc.). The accuracy and reaction times of each subject were recorded.

Each stimulus picture was presented for 1500 ms and then followed by 500 ms of a blank screen. Each block consisted of eight stimuli pictures. A total of eight blocks of presentation (four for faces, four for buildings) were carried out, each time alternating with a baseline picture (a scrambled image of building and face

together) for 16 seconds. Each stimulus picture was presented twice, so that the four-block presentation included 32 pictures in total. The order of the presentation of face and building blocks was randomized to prevent the subject from predicting next stimulus presentation.

Before the first experiment of both face and building judgments, the subjects were exposed to the 16 faces in the 4 x 4 matrix and the 16 building pictures of four different types in printed form for 30 minutes. During this phase, they were instructed to memorize specific information related to each face and the specific details of each building in order to recall them in the scanner. After completion of the first experiment, all participants were instructed to match this information with the printed pictures of faces in the 4 x 4 matrix and the buildings of each architectural type. This experimental procedure was performed every two weeks, and repeated four times over the course of eight weeks. In order to learn this information intensively, printed 4 x 4 matrix pictures of faces and four types of buildings were located in the work place of each subject. The pictures were simply hung on the walls around their desks. The subjects were exposed to these stimuli pictures at all times when they were seated at their desks. It is hard to control the learning effect of all subjects equally, but to avoid possible deviation of each subject the whole procedure was done within eight weeks.

A separate working memory experiment (one back matching task) was administered to localize the functionally defined areas (the FFA and PPA) in each subject. The stimulus images in the experiment were novel faces (also frontal views with neutral expressions) or buildings with natural backgrounds from our databank, also in black and white. This experiment consisted of eight blocks of stimulus presentation alternating with the scramble baseline. There were four blocks of face images as well as four blocks of building images. The same experiment was performed four times prior to each learning experiment.

Data Acquisition and Analysis

Images were acquired by using a 3 Tesla MRI scanner (VERIO, Siemens, Erlangen, Germany) with a quadrature head coil. Following a T1-weighted scout image, high-resolution anatomic images were acquired using an MPRAGE sequence with TE = 2.93 ms, TR = 1900 ms, flip angle = 9°, TI = 900 ms, and image size of 256 x 256. T2*-weighted functional data were acquired by using echo planar imaging (EPI) with TE = 30 ms, flip angle = 90°, TR = 2000 ms, and image size of 64 x 64. Field of view (FOV) was 192 x 192. We obtained 40-slices EPI images with slice thickness of 3 mm and no gaps between slices for the whole brain. The spatial resolution of each obtained voxel is 3 x 3 x 3 mm. 204 total volumes of images were

acquired in each experimental session of learning tasks and 109 volumes for the localizer scan.

The functional data were analyzed with the software BrainVoyager TMQX (Brain Innovation, Maastricht, The Netherlands). A standard sequence of preprocessing steps, such as 3D-motion correction, linear trend removal, and filtering with a high-pass of three cycles per scan was performed. The functional data set was projected to the IR-EPI-images, co-registered with the 3D-data set, and then transformed to Talairach-space (Talairach & Turnoux, 1988). Data was smoothed with a 6 mm FWHM Gaussian kernel. The EPI data were analyzed using multiple regression models (General Linear Model; GLM) consisting of predictors, which corresponded to the particular experimental conditions of each experiment. The predictor time courses used were computed on the basis of a linear model of the relation between neural activity and hemodynamic response, assuming a rectangular neural response during phases of visual stimulation (Boynton et al., 1996). In order to identify each subject's ROIs (FFA and PPA) the conjunction of the contrast (face vs. buildings) in the working memory task was computed. Changes of neural responses of the subjects were tested using repeated-measure ANOVA (mostly referred to as a random-effects analysis) on the averaged signals over each ROI, task type, and session.

Results

For the analysis of imaging data, the ROI of the FFA and the PPA were identified in each participant and in each session using localizer task data. We identified the area of the ROI independently for each session to avoid potential errors that might emerge from mis-registration of brain volumes acquired across the sessions separated in time by weeks.

The experimental data (BOLD signal changes) were analyzed by repeated-measures ANOVA, taking into account cross-subject variability. This was done using three factors: the session (from 1st to 4th measurement), the ROI (FFA and PPA), and the task (face and building judgment). The analysis revealed no significant effect (interaction) among the three factors $[F(3,28) = 1.2, p = 0.33]$. However, 'session' showed a significant interaction with 'ROI' $[F(3,28) = 4.32, p < 0.01]$. That is, session (the learning effect) has a different effect on FFA activations than on PPA activations.

In a separate ROI analysis, the FFA showed a significant effect for the session $[F(1,42) = 7.55, p < 0.001]$, whereas it showed no significant interaction with the task $[F(3,42) = 2.00, p = 0.13]$. The PPA showed no significant effect for the session $[F(3,42) = 0.55, p = 0.65]$ and no interaction between session and task $[F(3,42) = 1.71, p = 0.18]$. This means that the FFA showed a clear
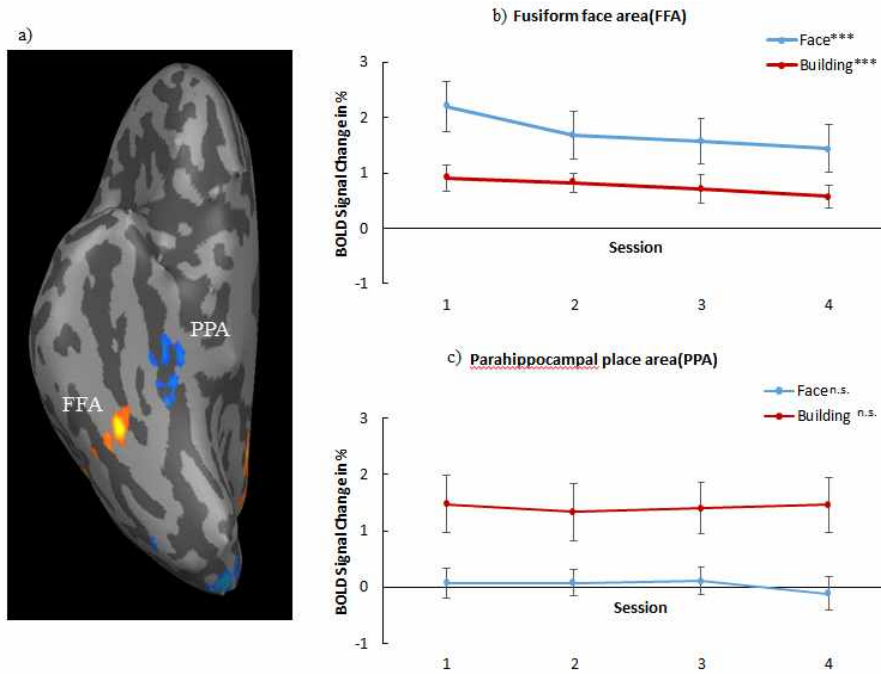
Figure 2. Activation map of learning experiment. The left figure is an activation map from a subject. The FFA is colored by yellow and the PPA is by blue. Right plots are averaged BOLD responses from all participants of face and building judgment tasks separately from the first experiment to fourth by long term learning task in FFA and PPA. The BOLD response were extracted and averaged from each subject's identified FFA and PPA. Significant reduction of BOLD responses was observed in FFA and no significant in PPA across all sessions. *** $p < .001$, n.s for no significant.

learning effect regardless of the stimulus type, while no evidence for such a learning effect was observed for the PPA. It is important to note that no interaction between the learning effect and the task itself was observed. This suggests that the PPA has no learning effect.

Discussion

This study investigates the functional differences between the FFA and the PPA. The main finding of the present study is that a differential pattern between FFA and PPA exists for BOLD signal changes over the learning period: The FFA was activated in the early phases but then the level of activation decreased, while the responses in the PPA showed no such activation changes. This was observed not only for faces, which are the preferred stimuli for FFA, but also for buildings, the less preferred stimuli for the area. This underscores functional differences between the

FFA and the PPA. Previous studies contrasted both of the human ventrotemporal areas in terms of their object class specifying ability, and implied that they are positioned at a similar level in the hierarchy of visual information processing (McKeeff et al., 2007). One might expect the object-class specificity to become sharper, leading to a greater difference in neural activities between FFA and PPA to a visual object class, as subjects became more proficient in tasks involving the object class, the relative activities enhanced, which neural correlates underlying the behavioral improvement, in the form of enhanced differences. This suggests that two equally well-categorizing ventral stream areas may have differential computational properties with regards to their experience-dependent adaptability.

For the eight weeks, the subjects were instructed to memorize specific information associated with several well-known faces. As they were learning, the meaning of the original character of each stimulus, such as the face, was attenuated; and the meaning of facial identity was decreased.

Associative information based on facial identity becomes more meaningful rather than the recognition of the facial stimulus / identity per se. Recalling the associate memory leads to changes in the neural circuitry. Recognizing faces is one obligatory option for accessing other information, e.g. position on the matrix. In other words, the neural circuitry, which preferentially responds to facial stimuli, must be changed / reorganized effectively in order to successfully achieve the judgment task. It is probable that the continuous reduction of neural responses reflects this fact.

Despite the fact that the number of neuronal cells in the FFA which respond to buildings is much smaller than those for faces, a similar change in neural signal occurs (smaller amplitude of BOLD signal changes). On this basis, we postulate that the neuronal circuitry in the FFA is not only dominant in responding to facial stimuli, but also responsive to other objects such as buildings. However, the FFA is sensitive to such learning effects, and will be activated if exposed to further learning processes.

Our observation of the plasticity in the FFA is consistent with previous studies on the influence of learning / training on neural activity (Aguirre et al., 1998; Epstein, 2008; Epstein et al., 1998; Kanwisher et al., 1997; Kanwisher & Yovel, 2006). Reduced responses were observed in the same tasks using identical visual stimuli repeatedly, a process known as "repetition suppression" (Gauthier et al., 1999; McGugin et al., 2014).

It has been proposed that the attenuation of neural activity during repeated exposure reflects the optimization of stimulus representation in the ventral stream. It hypothesizes that repeated exposure to the same stimuli leads to a smaller

prediction error for expected stimuli, which in turn causes a decrease in activation for repeated stimuli (Mayrhauser et al., 2014).

This model, however, is at odds with our observation that FFA responses decreased for both preferred and less preferred object classes. Within the FFA, object specifying function remains intact, whereas an optimization process for less preferred objects was also induced by repeated exposure to stimuli.

We propose that the signal reduction in the FFA during the learning process results from the optimization of memory retrieval. As an associative connection between a face and spatial coordinates is formed, neuronal circuitry in the FFA is reorganized so that less time is required for the neurons responsible for coding the face to trigger a connection.

Another important aspect of the present study is that the PPA does not show neural response changes caused by long-term learning (Figure 2). The neuronal clusters in this area respond dominantly to building-like stimuli, but not to face-like stimuli. The response to faces in this area was almost zero, a tendency which was consistent across the experimental procedures. The PPA responses to the building and facial stimuli were not sensitive to any repetition. On this basis, we suggest that the PPA and the FFA probably differ in terms of their involvement in visual associative memory, and this distinction should be incorporated in modeling the functional architecture of the human ventral stream.

The limitation of the study is that different training tasks (matrix location/category associations) were used for the two categories (faces vs. buildings). If the same task were used for both categories, it should have been much easier to evaluate the training effects and interpret the results. Further study is needed to support such interpretation of the present results.

Acknowledgement

References

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron, 21*, 373–383.

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience, 16*, 4207-4221.

Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *Journal of Neuroscience, 30*,

16601-16608.

Epstein, R. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Neuroscience*, *12*, 388-396.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999), The parahippocampal Place Area: recognition, navigation or encoding?. *Neuron, 23*, 115- 125.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2*, 568-573.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Neuroscience, 10*, 14-23.

Henson, R. N. (2003). Neuroimaging studies of priming. *Progress in Neurobiology, 70*, 53−81.

Jiang, X., Bradley, E., Rini, R. A., Zeffiro, T., VanMeter, J., & Riesenhuber, M. (2007). Categorization training results in shape and category-selective human neural plasticity. *Neuron, 53*, 891-903.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302-4311.

Kanwisher, N., & Yovel, G. (2006), The fusiform face area: a cortical region specialized for the perception of faces. *Philosophy Transaction of Royal Society B, 361*, 2109-2128.

Mayrhauser, L., Crone, J. J., & Kronbichler, M.

(2014). Neural repetition suppression: evidence for perceptual expectation in object-selective regions. *Frontiers in Human Neuroscience, 8*, 225.

McGugin, R. W., Newton, A. T., Gore, J. C., & Gauthier, I. (2014). Robust expertise effects in right FFA. *Neuropsychologia, 63*, 135-144.

McKeeff, T. J., Remus, D. A., & Tong, F. (2007). Temporal limitations in object processing across the human ventral visual pathway. *Journal of Neurophysiology, 98*, 382−393.

Talairach, J., & Turnoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-Dimensional proportional system.* New York, NY: Thieme.

Tsakiris, M., & Apps, M. A. (2013). Predictive codes of familiarity and context during the perceptual learning of facial identities. *Nature Communications, 4*, 2698.

Van den Linden, M., Murre, J. M. J., van & Turennout, M. (2008). Birds of a feather flock together: experience-driven formation of visual object categories in human ventral temporal cortex. *PLoS ONE, 3*, 1-11.

Van Turennout, M., Ellmore, T., & Martin, M. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience, 3*, 1329-1334.

Weisberg, J., van Turennout, M., & Martin, A. (2007). A neural system for learning about object function. *Cerebral. Cortex, 17*, 513-521.

복측두엽 영역에서의 시각연합기억의 다양한 영향에 관한 신경기전 변화 연구

윤 효 운   박 효 민

대구사이버대학교 미술치료학과   성균관대학교 거버넌스연구센터

방추상회 및 해마 주변의 특수한 신경반응은 재인된 자극의 종류에 따라 구분된다고 할 수 있다. 그러나, 두 영역의 기능에 대한 근본적인 설명은 아직도 이루어지지 않고 있다. 본 연구는 범주적 특성을 가진 두 영역에서 얼굴 및 건물 자극에 대해 집중적인 학습을 하는 과제를 제시해 학습효과를 차이를 살펴보고자 한다. 동일한 피험자에게 4회의 실험이 실시되었다. 실험은 2주 간격으로 총 8주 동안 진행되었다. 제시되는 얼굴자극은 피험자들에게 친숙한 얼굴들로 구성되고 피험자들은 그 얼굴의 정체성(identity)이 아닌 새로운 정보(16명의 얼굴사진을 4 x 4의 행렬에 배치하여 각 한 개의 얼굴자극이 특정한 열과 행의 번호(좌표)를 가졌다. 예를 면 A라는 얼굴은 1x1, B라는 얼굴은 2x1이라는 정보)를 학습하여 하나의 얼굴이 자극으로 제시될 경우, 정확한 응답을 하도록 학습해야 한다. 실험결과를 보면 FFA 영역에서의 통계적으로 유의미한 신호감소 효과를 관찰할 수 있었다. 반면, PPA 영역에서는 그렇지 않은 결과를 관찰되었다. FFA 영역의 BOLD 신호감소는 자극에 대한 반복적 학습으로 효과라고 여겨지며, 자극의 재인시 신경세포 활동의 감소와 연관된다.

주제어 : Fusiform Face Area, Parahippocampal Place Area, 장기기억, 시각연합기억, 신경기전 변화