Individual differences in three-dimensional object recognition: An event-related optical topography study.

Takeo Kondou Hiroshima University bstorm@hiroshima-u.ac.jp

In this study, we examined whether individual differences in the capacity of spatial information processing (spatial span) relate to view-dependence / invariance in 3-D object recognition in the two experiments. In Experiment 1, according to the scoring of a spatial span task (Shah & Miyake, 1996), ten participants with a high spatial span score and ten participants with a low spatial span score were assigned to high and low spatial span groups, respectively. The results indicated that only the low spatial span group showed a view-dependence in the high cognitive load condition (number of targets) of the 3-D object discrimination task. In Experiment 2, we measured the homodynamic changes in the bilateral occipital-temporal region of human visual cortex in a 3-D object discrimination task with high cognitive load. The oxy- and deoxyhemoglobin change indicated that the high span group increased activity in the inferior parietal and temporal region with increasing the angle of viewpoint change as opposed to the low span group. In addition, the amount of relative increase in oxy-hemoglobin of high-span group was significantly larger than the contralateral region in the left inferior occipital-temporal region. This region considered as include the fusiform gyrus, which maintain view-invariant representation in the left hemisphere and view-dependent representation in the right hemisphere (Vuilleumier, Henson, Driver, & Dolan, 2002). These results indicated that there are behaviorally and neuro physiologically distinct individual differences in 3-D object recognition between groups with the high and low spatial span. This suggested that individual differences in spatial span could be one of the factors which affect the 3-D object recognition.

Keywords: individual differences, working memory, three-dimensional object recognition, optical topography, near-infrared spectroscopy, functional imaging, hemodynamics.

Introduction

There has been a controversy about “whether our 3-D object recognition is view-dependent or view-invariant”. It has been reported that many properties of task and object affect this dependent / invariant divergence (e.g., , categorical or exemplar specific: Tarr & Bülthoff, 1995; familiar or novel: Tarr, 1995; detectable or not with non non-accidental properties: Biederman & Bar, 1999). What factor affects this divergence is still a research interest.

Working memory (WM) is a limited resource that processes actively stored information to perform goal-oriented tasks, and there are individual and developmental differences. Although the stimulus object was same, simultaneously increasing in the number of target objects (memory set size) and the similarity between target and distracter objects, makes the recognition of 3-D objects view-dependent (Newell, 1998). This means that the increased load on processing and storing the 3-D objects produces a view-dependent performance. Individual differences also occur in visual-spatial WM, and a span test to assess the capacity of an individual's WM has been developed (Spatial span task; Shah & Miyake, 1996).

The factor of individual differences is regarded as an internal factor and the factor of task or object are regarded as external factor. Therefore these factors are different from each other and the individual differences in 3-D object recognition have not been investigated. The present study examined whether the behavioral performance (Experiment 1) and neuronal activity (Experiment 2) of 3-D object recognition would change with individual differences in a WM capacity.

Experiment 1

First, we investigated whether the behavioral performance will change with the individual difference in the visuo-spatial WM capacity. It predicted that performing a same task, the high and low capacity groups’ performance will be view-invariant and view-dependent, respectively.

Methods

Participants Thirty-two naive undergraduate and graduate students were volunteered to participate.

Materials The stimuli were twenty objects from Hayward & Tarr (2000). The set of target images consisted of a canonical view which was decided by the participants and another two views which were incremented from the canonical view by 40° and 80° respectively. The distracter objects had a main body of the same shape as the target objects, and had appendages at approximately the same location and with the same inter-relationship. The set of distracter images were made from the distracter objects and from the same viewpoint as the target image set.

Experimental Design & Procedure The design of the present study used the spatial span task (SST) score group (HIGH, LOW) as a between-participants variable, and the memory set size (LARGE, SMALL) and viewpoint (0º, 40º, 80º) as the within-participant variables.

First, a SST was used to assess the ability of the participant's WM capacity to simultaneously process and store spatial information.

In the following delayed discrimination task, the participants were initially asked to remember the shape of targets which were rotated five times on the display of a CRT. According to the memory set size condition, the participants were shown three or one target object. The participants then were shown all the views from thirty-six viewpoints of the target object, and asked to choose the best (canonical view: 0°) from these views. After that, the participants were asked to discriminate the target images from the distracter images by pressing a computer mouse button (regardless of the changes in viewpoint). Each trial was composed of a fixation cross for 500 ms, then a display of a test image (the target or distracter). The test image disappeared when the participant responded and the next trial (which was followed by blank for 500 ms) started.

Results & Discussion

Ten participants with a high SST score (median: 3.50, range: 3.25-5.00) and ten with a low SST score (median: 1.00, range: 1.00-1.50) were assigned to the HIGH and LOW score groups, respectively. The remainder of the participants were excluded from our analysis.

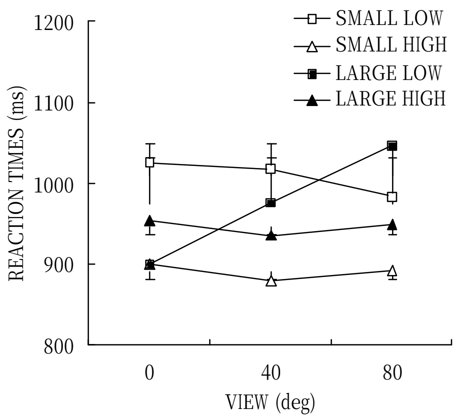

Figure 2. Mean reaction times as a function of the angles of the stimulus object's viewpoint, with SST score (LOW, HIGH) group and memory set size (SMALL, LARGE) as a parameter.

The median reaction times for the correct responses as a function of viewpoint with the SST score group and memory set size as a parameter are shown in Figure 2.

A three-way analysis of variance was calculated with one between-participants factor (SST group; HIGH, LOW) and two within-participant factors (memory set size: LARGE, SMALL; viewpoint: 0°, 40°, 80°). The analysis revealed a significant interaction of memory set size and view [$F(2, 36) = 3.751$, $p < .05$]. The simple main effect of view was significant between 0o and 80o in the condition of the large set size. This result means view-dependence and replicates the results of Newell (1998). Also, the analysis revealed a significant interaction of all of the three factors [$F(2, 36) = 3.457$, $p < .05$]. In a second analysis, only for the LARGE memory set size condition, the LOW group's reaction time for 80° was significantly longer than for 0°. Further analysis revealed that there was not a significant effect of viewpoint for the SMALL memory set size and the HIGH group. Thus, on the contrary to the result of HIGH group, view-dependence was observed only when the LOW groups memorized and recognized the multiple target objects simultaneously. That is, the factor of individual differences is one of the factors that explain the disparity of 3-D object recognition performance.

Experiment 2

In Experiment 2, we measured homodynamic changes during the 3-D object discrimination task with the high cognitive load similar to Experiment 1. Measuring hemodynamic changes, we investigated neuronal basis of the individual differences in the 3-D object recognition observed in Experiment 1. To measure the hemodynamic changes, we used an optical topography system which based on the technique of near-infrared stereoscopy (NIRS). NIRS assess the oxyhemoglobin and deoxyhemoglobin changes in brain-tissues by detecting the difference of light absorption spectra of those hemoglobin.

It is reported that the fusiform gyrus is responsible for depth rotated 3-D object recognition (Gauthier, Hayward, Tarr, Anderson, & Skudlarski, 2002). The fusiform gyrus has laterality. The left and right fusiform gyrus were related to the process of view-invariant and view-dependent manner, respectively (Vuilleumier, Henson, Driver, & Dolan, 2002). Thus we measured the bilateral occipital-temporal region of human visual cortex including the fusiform gyrus.

It predicted that there will be the differences in the hemodynamic patterns for the viewpoint changes. This difference might be the activation of the LOW group in the region that correspond to the fusiform gyrus will be lower than the activation of the HIGH group.

Methods

Participants Twenty-two naive undergraduate and graduate students were volunteered to participate. Eight out of twenty-two participants were selected to participate to an optical topography experiment according to the SST score. They showed right-handedness.

Materials The stimuli were almost same as Experiment 1, but a procedure was different. Due to procedure that the participants arbitrarily decided the 0o views, the 80o views were often a mirror-image of the 0o views by rotating objects clockwise. To avoid this, we decided rotating orientation (clockwise / anti-clockwise) not to produce the mirror-image. An optical topography system ETG-100 (Hitachi Medico, Tokyo, Japan) was used to measure and analyze the hemodynamic changes. Twenty-four channels in both hemispheres were simultaneously recorded every 500 ms to obtain the transmittance data $\ln T(\lambda, t)$ as a function of wavelength (λ) and a measurement time (t). Twelve channels assigned to the left hemisphere and the rest of the channels were assigned to right hemisphere. These channels covered an 8 ×8 cm^2 area in each hemisphere (Figure. 3). We used T5 (according to the international 10-20 method) to decide the location of the channel.

Experimental Design & Procedure The design of the present study used the spatial span task (SST) score group (HIGH, LOW) as a between-participants variable, and the hemisphere (LEFT, RIGHT) and viewpoint (0o, 40o, 80o) as the within-participant variables.

The 3-D object discrimination task was similar to the condition of the LARGE memory set size. To obtain event-related hemodynamic response, a intertrial interval of 30 sec was inserted. This procedure makes single trials well-separated events.

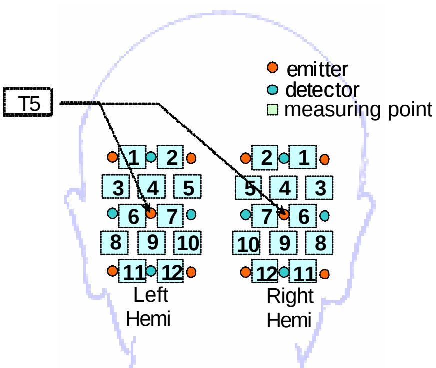

Figure 3. The location of channels.

Results

Four participants with a high SST score [median: 4.50, range: 4.25-4.50] and four with a low SST score [median: 1.25, range: 1.0-1.5] were assigned to the HIGH and LOW score groups. The remainder of the participants were excluded from our analysis.

Behavioral data An ANOVA conducted to investigate whether the performances of each SST score group (HIGH / LOW) were view-invariant or view-dependent. Error trials were excluded from the analysis of reaction time. In terms of the reaction times, the main effect of view was significant [$F(2,6) = 5.806$, $P < .05$] only for LOW group. Further analysis revealed the significant difference between 0° and 40°. Thus, the behavioral performance of the LOW group was view-dependent and the HIGH group was view-invariant.

Table 1. Behavioral data*

	Group	0°	40°	80°
Reaction time (ms)	HIGH	1207 ± 15.7	1301 ± 120.7	1309 ± 59.6
	LOW	1492 ± 153.3	1808 ± 216.4	1644 ± 145.6
Error rate (%)	HIGH	8.3 ± 8.3	10.4 ± 7.8	10.4 ± 5.2
	LOW	14.6 ± 8.6	18.8 ± 10.9	37.5 ± 10.5

* Values are expressed as mean ± SE (n = 8).

Hemodynamic responses Excluding error trials, signals of 3-9 sec from the stimulus onset were used for analysis. We conducted a three-way ANOVA (SST × Hemisphere × View) for the oxyhemoglobin (C_{oxy}) and deoxyhemoglobin change (C_{deoxy}) of each channel.

In terms of C_{oxy}, there were significant differences at the channel 2 and 3 (Figure 4). At the channel 2, there were significant main effects of View [$F(2,12) = 6.297$, $P < .05$] and SST [$F(1,6) = 9.381$, $P < .05$]. In a second analysis, the C_{oxy} for 0° view was significantly smaller than 40° ($P < .05$) and 80° ($P < .01$). For SST, the C_{oxy} of the HIGH group was significantly larger than the LOW group ($P < .05$). Also, a significant interaction between SST and View were found at the channel 2 [$F(2,12) = 6.318$, $P < .05$] and 3 [$F(2,12) = 5.869$, $P < .05$]. Further analysis revealed that only for the HIGH group, the C_{oxy} for 0° view was significantly smaller than 40° ($P < .001$) at the channel 2, and, smaller than 40° ($P < .05$) and 80° ($P < .01$) at the channel 3.

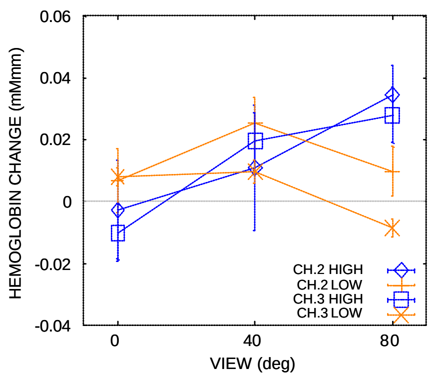

Figure 4. Mean C_{oxy} at the channel 2 and 3 as a function of the angles of the stimulus object's viewpoint, with SST score (LOW, HIGH) group as a parameter. Error bars expressed SE.

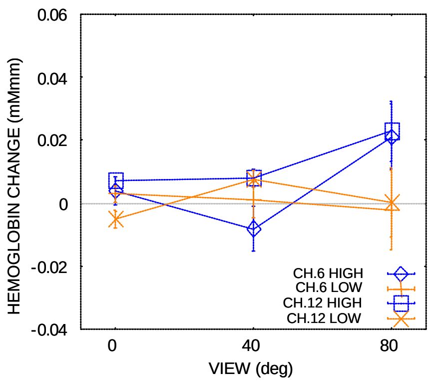

Figure 5. Mean C_{deoxy} at the channel 6 and 12 as a function of the angles of the stimulus object's viewpoint, with SST score (LOW, HIGH) group as a parameter. Error bars expressed SE.

In terms of C_{deoxy}, there were significant differences at the channel 2, 3, 6, 8, 11 and 12. At the both channel 2 and

3, there were a significant main effect of SST [$F(1,6) = 7.199$, $P < .05$ at the channel 2, $F(1,6) = 6.287$, $P < .05$ at the channel 3]. The C_{deoxy} of the HIGH group was significantly smaller than the LOW group at the both channels. These decreases at the channel 2 and 3 might be the counter effect of the increase in the C_{oxy} at the same channels. A significant interaction between SST and View was found at the channel 6 [$F(2,12) = 4.256$, $P < .05$] and 12 [$F(2,12) = 3.933$, $P < .05$] (Figure 5). In a second analysis, a simple main effect of SST for 80º was significant at the channel 6 [$F(1, 18) = 8.213$, $P < .05$] and 12 [$F(1, 18) = 12.797$, $P < .001$]. At the both channels, the C_{deoxy} of the HIGH group was significantly larger than the LOW group ($P < .05$). Also, a significant interaction between SST and Hemisphere was found at the channel 8 [$F(2,12) = 13.650$, $P < .05$] and 11 [$F(2,12) = 6.139$, $P < .05$] (Figure 6). A significant simple main effect of View was also found in the HIGH group at the channel 6 [$F(2,12) = 6.565$, $P < .05$] and 12 [$F(2,12) = 3.933$, $P < .05$]. The Cdeoxy of 40º was smaller than 80º ($P < .005$) at the channel 6and, the 0º was smaller than 40º ($P < .05$) and 80º ($P < .05$) at the channel 12. There was a simple main effect of Hemisphere in the HIGH group at the channel 8 ($P < .01$) and 11 ($P < .05$).

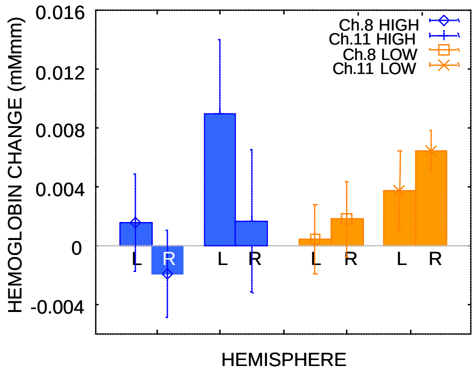

Figure 6. Mean C_{deoxy} at the channel 8 and 11 as a function of the hemisphere (LEFT, RIGHT) with SST score (LOW, HIGH) group as a parameter. Error bars expressed SE.

Discussion

The behavioral data indicated the divergence of view-dependent/invariant with the SST score groups. The HIGH group showed view-invariance and the LOW group showed view-dependence and this confirmed the results of Experiment 1. Thus the relationship between individual differences and 3-D object recognition was observed in both experiments.

This relationship was compensated by the fact that the hemodynamic responses of the HIGH and LOW group were different from each other. There were significant effects of View, SST and Hemisphere. The C_{oxy} at the channel 2 (inferior parietal region) and 3 (temporal region), and the C_{deoxy} at the channel 6 and 12 (occipital region) were increased in view-dependent manner especially in the HIGH group, which were linearly increased with viewpoint changes. These hemodynamic patterns were different from those of the LOW group, which showed little linear increase with viewpoint changes (Figure 4, 5).

The results indicated that, at the channels of 8 and 11 (left inferior occipital-temporal region), the C_{deoxy} of the HIGH group was significantly larger than the contralateral region (Figure. 6). This region considered to include the fusiform gyrus that maintain distinct view-dependent and view-invariant object representation (Vuilleumier, Henson, Driver, & Dolan, 2002). There might be a difference in processing strategies of 3-D object representation between the HIGH and LOW group.

Conclusion

In this study, we found a new viewpoint of individual differences in spatial processing capacity for investigating 3-D object recognition. Through this viewpoint, it will be exposed distinct processes in performances seem like view-dependent or view-invariant at a glance. We suggest this difference should be taken into consideration when we investigate 3-D object recognition.

References

Biederman, I. & Bar, M. 1999 One-shot viewpoint invariance in matching novel objects. *Vision Research*, **39**, 2885-2899.

Gauthier, I., Hayward, W. G., Tarr, M. J., Anderson, A. W., & Skudlarski, P. BOLD activity during mental rotation and viewpoint-dependent object recognition. *Neuron*, **34**, 161-171.

Hayward, W. G., & Williams, P. 2000 Viewpoint dependence and object discriminability. *Psychological Science*, **11**, 7-12.

Newell, F. N. 1998 Stimulus context and view dependence in object recognition. *Perception*, **27**, 47-68.

Shah, P., & Miyake, A. 1996 The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, **125**, 4-27.

Tarr, M. J. 1995 Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review*, **2**, 55-82.

Tarr & Bülthoff, 1995 Is human object recognition better described by geon-structural-descriptions or by multiple-views? Comment on Biederman and Gerhardstein 1993). *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 1494-1505.

Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J., 2002 Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, **5**, 491-49.