Neural Correlates of Stroop Performance in Alzheimer’s Disease: A FDG-PET Study

Je-Yeon Yun\textsuperscript{a} Dong Young Lee\textsuperscript{a, b} Eun Hyun Seo\textsuperscript{a, b}
Il Han Choo\textsuperscript{a} Shin Young Park\textsuperscript{c} Shin Gyeom Kim\textsuperscript{e}
Jong In Woo\textsuperscript{a, b, d}

\textsuperscript{a}Department of Neuropsychiatry, Seoul National University Hospital, \textsuperscript{b}Interdisciplinary Program of Cognitive Science, Seoul National University, \textsuperscript{c}Department of Neuropsychiatry, Daelim Saint Mary’s Hospital and \textsuperscript{d}Neuroscience Research Institute, Medical Research Center, Seoul National University, Seoul, and \textsuperscript{e}Department of Neuropsychiatry, Soonchunhyang University Bucheon Hospital, Bucheon, Korea

Key Words
Alzheimer’s disease · Neural correlates · PET · Prefrontal cortex · Stroop effect

Abstract

\textbf{Background/Aims:} The Stroop test is commonly applied in elderly subjects for the evaluation of cognitive impairment related to Alzheimer’s disease (AD) and related disorders. This study aimed to investigate the functional neural correlates of the Stroop performance in AD. \textbf{Methods:} In 136 probable AD patients and 54 cognitively normal elderly, a [\textsuperscript{18}F]-fluorodeoxyglucose positron emission tomography scan and Stroop Color Word Test (SCWT) were performed. The correlations between the Stroop effect, which was measured by 6 different scoring methods, and regional cerebral glucose metabolism (rCMglc) were explored using a region-of-interest (ROI) approach and voxel-based analysis. \textbf{Results:} Among 6 Stroop interference measures, only 2 scores, including the SCWT color-word (CW) score, were significantly correlated with rCMglc of the dorsolateral prefrontal and anterior cingulate ROIs. Voxel-based analysis revealed significant positive correlations between SCWT CW scores and rCMglc in the inferior parietal lobe, middle temporal gyrus and middle frontal gyrus. Such correlations remained significant only in the less severe AD group. \textbf{Conclusion:} In AD patients, the Stroop effect depends on the functional integrity of the prefrontal cortices. Some parietotemporal regions also appear to be responsible for the Stroop effect in AD individuals.
Introduction

The Stroop task was originally developed by early experimental psychologists who found the so-called ‘Stroop effect’, which shows that reading the names of colors printed using non-matching colored ink is always slower than simply naming colors or reading words alone [1]. The paradigm is known to measure selective attention or executive function, in terms of an individual’s ability to suppress habitual responses in favor of unusual responses [2, 3]. The Stroop test is widely used in clinical settings to differentiate, monitor, and describe various neuropsychiatric disorders [4, 5]. In particular, it is very commonly applied in elderly subjects for the evaluation of cognitive impairment related to Alzheimer’s disease (AD) and related disorders. It has also been found to sensitively differentiate patients with mild AD and mild cognitive impairment, a preclinical stage of AD, from cognitively normal (CN) elderly individuals [6–9].

Findings from brain activation studies suggested that the prefrontal cortical areas, especially the dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC), are closely associated with the Stroop effect [5, 9, 10]. It is, however, not clear whether or not the prefrontal cortical regions are the main pathoanatomical source for the impaired Stroop performance in AD. Main AD pathologies, including neurofibrillary tangles and neuronal or synaptic loss, initially appear in the medial temporal cortex [11, 12] and then progress to posterior cortical areas, such as the lateral temporal and parietal cortices [13], while the frontal cortex is not involved until advanced stages of AD.

In order to identify the functional neuroanatomical correlates of Stroop performances in patients with AD, we investigated the relationship between 6 kinds of Stroop interference (SI) scores reflecting the Stroop effect and regional cerebral glucose metabolism (rCMglc) measured by positron emission tomography (PET). We first tested whether or not the prefrontal dysfunction was related to SI scores in AD using a region-of-interest (ROI) approach. Secondly, we tried to explore overall neuroanatomical correlates of the SI scores through a voxel-based approach (VBA) without an a priori hypothesis.

Patients and Methods

Subjects

The study subjects were recruited from patients with AD who visited the Dementia and Age-Associated Cognitive Decline Clinic of the Seoul National University Hospital in Seoul. The study included 136 patients with AD who met the criteria for dementia of the Diagnostic and Statistical Manual of Mental Disorders [14] and criteria of probable AD of the National Institute of Neurological and Communicative Disorders and Stroke and the AD and Related Disorder Association [15]. Fifty-four CN elderly subjects were also selected from a pool of elderly volunteers with a normal neurologic and psychiatric history and examination, and a normal brain magnetic resonance imaging (MRI) scan. Exclusion criteria for this study were presence of any serious medical, psychiatric and/or neurological disorder that could affect mental function; evidence of focal brain lesions on MRI; the presence of severe behavioral or communication problems that would make a clinical or PET examination difficult; both- or left-handedness, and absence of a reliable informant. The Institutional Review Board of the Seoul National University Hospital, Korea, approved the study protocol, and informed consent was obtained from all subjects and their relatives.
Clinical Assessments and Neuropsychological Tests

All of the subjects were examined by psychiatrists who had advanced training in neuropsychiatry and dementia research according to the protocol of the Korean version of the Consortium to Establish a Registry for AD (CERAD-K) Assessment Packet [16, 17]. Psychiatric, general physical and neurological examinations were performed along with routine laboratory tests and MRI of the brain. Reliable informants were interviewed to acquire accurate information regarding the cognitive, emotional and functional changes as well as the medical history of the subjects. Seven neuropsychological tests in the CERAD neuropsychological battery, including the Verbal Fluency – ‘Animal Category’, 15-item Boston Naming Test, Word List Memory, Word List Recall, Word List Recognition, Constructional Praxis and Constructional Recall [16, 17], were applied.

A panel consisting of four psychiatrists with expertise in dementia research made clinical decisions, including the clinical diagnosis and clinical dementia rating (CDR) [18], after reviewing all the available raw data. All clinical assessments were carried out within 3 weeks of the PET examination. None of the subjects was receiving any antidepressant or other psychotropic medication cholinesterase inhibitors.

Stroop Color Word Test and SI Scores

The standardized version of the Stroop Color Word Test (SCWT) [19] was used as the basis for the test used in the present study, but stimulus pages were modified for Korean elderly subjects [20]. The SCWT consists of 3 pages, each containing 100 items presented in 5 columns of 20 items. On the Word (W) page, the words ‘red’, ‘yellow’, and ‘blue’ translated into Korean were printed in black ink, whereas on the Color (C) page, stimuli were printed as 2 Xs (XX) in red, yellow or blue, because in Korean the color words used had two letters on average (thus controlling for the size of colored regions on the page). The CW page contained the same color words printed in non-matching colors (e.g. ‘red’ printed in blue ink). When tested with the W page or the CW page, subjects were required to read the presented words as rapidly as possible in 45 s. When using the C page, subjects were instructed to name the color of each XX [20].

Experienced clinical psychologists administered the SCWT test. The W, C and CW subscores refer to the number of items read or named correctly for each test page during the test performance in 45 s.

The Stroop effect is defined as the extent of delay in naming the color of an incongruent color word relative to naming the color of a congruent color word or of a neutral non-color word [21]. The Stroop effect was measured by 6 SI scores calculated by 6 kinds of scoring methods. For SI scores, the scoring methods used in this study are described below.

- **SI-1**: SI-1 is the raw score itself of the CW page of the SCWT (i.e. CW score) [19].
- **SI-2**: SI-2 is the difference between the C and CW scores. A lower SI-2 score means less interference from incongruent words when naming the colors on the CW page [21].
- **SI-3**: SI-3 is the ratio of (C – CW)/C. Higher SI-3 means less interference from incongruent words when naming the colors in the CW condition [21].
- **SI-4**: SI-4 is the difference between the predicted CW score of (W·C)/(W + C) and the CW score. This equation stems from the hypothesis that the time to name a CW item is equal to the time needed to suppress the reading of a word plus the time to identify a color [19].
- **SI-5**: SI-5 is computed by the equation CW – (W + C)/2 [22].
- **SI-6**: SI-6 is calculated using the equation (W + CW) – C [23].

PET Image Acquisition and Preprocessing

PET studies were performed using the ECAT EXACT 47 scanner (Siemens-CTI, Knoxville, Tenn., USA), which has an intrinsic resolution of 5.2-mm full width at half maximum
and images of 47 contiguous transverse planes with a 3.4-mm thickness for a longitudinal field of view of 16.2 cm. Before administering [18F]-fluorodeoxyglucose (FDG), transmission scanning was performed using 3 germanium-68 rod sources to correct the attenuation. Static emission scans began 30 min after the intravenous injection of 370 MBq (10 mCi) [18F]-FDG and were continued for 30 min. All of the [18F]-FDG PET scans were performed in a dimly lit room with minimal auditory stimulation during both the injection and PET scanning. Subjects were in a supine position with their eyes closed during the scan in order to minimize the confounding effects of any activity. The transaxial images were reconstructed using a filtered backprojection algorithm employing a Shepp-Logan filter with a cutoff frequency of 0.3 cycles/pixel as 128×128×47 matrices with a size of 2.1×2.1×3.4 mm.

Imaging data were analyzed using Statistical Parametric Mapping 2 (SPM2; Institute of Neurology, University College of London, UK) implemented in Matlab (Mathworks Inc., Natick, Mass., USA). Before statistical analysis, all images were spatially normalized to the Montreal Neurological Institute (MNI, McGill University, Montreal, Calif., USA) space to correct intersubject differences in brain morphology [24]. An affine transformation was performed to determine the 12 optimal parameters essential for registering the brain on the MNI template. Subtle differences between transformed images and the template were removed by a nonlinear registration method using the weighted sum of predefined smooth basis functions in a discrete cosine transformation. The glucose metabolism value of each voxel was normalized for the pontine value, which was extracted for each scan, because glucose metabolism in the pons tends to be relatively preserved in AD [25]. Normalized images were smoothed by convolution using an isotropic Gaussian kernel with 16-mm full width at half maximum to accommodate intersubject differences in brain morphology and to increase dataset signal-to-noise ratio.

**PET Image Analysis**

**CN Group versus Total AD Group**

For additional confirmation of clinical AD diagnosis in AD subjects, differences in glucose metabolism between the AD and the CN groups were measured on a voxel-by-voxel basis using the 2-sample t test of SPM2. The resulting set of t values constituted the SPM(t) map. The SPM(t) was then transformed into a normal distribution to give a SPM(Z) [family-wise error (FWE)-corrected p < 0.05, k > 50] [26, 27].

**Relationship between SI Scores and rCMglc**

To investigate the relationship between SI scores and rCMglc, we used two different approaches: an ROI-based approach with a hypothesis and VBA without a hypothesis.

**ROI-Based Approach.** The automatic anatomic labeling (AAL) algorithm and a region-combining method [28, 29] were applied to set ROIs to measure regional brain metabolism in the 6 prefrontal areas – bilateral DLPFC (AAL template Nos. 03–04 and 07–08), bilateral orbitofrontal cortex (OFC; AAL template Nos. 05–06, 09–10, 15–16, 21–22, 25–26 and 27–28) and bilateral ACC (AAL template Nos. 31–32; fig. 1).

Correlations between rCMglc in each ROI and 6 SI scores were tested by partial correlation analysis controlling for age, gender and education (SPSS, Cary, N.C., USA). The level of statistical significance for ROI-based analyses was set at 2-tailed p < 0.0083 (Bonferroni corrected).

**Voxel-Based Approach.** Correlations between the SI-1 score and rCMglc were additionally examined through VBA. VBA is a technique for assessing regional changes in brain tissue content over the entire brain on a voxel-by-voxel basis without the need for an a priori ROI. VBA was conducted using the 'Single Subject: Covariates Only' menu of SPM2, with 3 variables of age, gender and education included as nuisance variables within the AD group.
The resulting set of t values constituted the SPM(t) map. The SPM(t) was then transformed into a normal distribution to give a SPM(z) (FWE-corrected p < 0.05, k > 50) [26, 27].

Statistical Analysis of Demographic and Neuropsychological Data

The demographic and clinical data for AD and CN subjects were compared with an independent t test for continuous variables and χ² tests for categorical variables, using SPSS 18.0.

Results

Demographic and Clinical Characteristics of the Subjects

Demographic and clinical characteristics of the subjects are shown in table 1. There were no significant differences in age, education and gender between CN and AD group. The patients with AD showed a significantly lower mean Mini-Mental State Examination (MMSE) score than the CN group (t = 19.17, p < 0.001). The AD group included 66 patients with very mild (CDR 0.5), 55 with mild (CDR 1) and 15 with moderate AD (CDR 2). Twenty-two AD patients were taking cholinesterase inhibitors [donepezil (21 patients) and rivastigmine (1 patient)]. No subject took memantine.

Stroop Test and Other Neuropsychological Test Scores in AD

With respect to all 8 neuropsychological tests, there were significant differences between the AD and CN groups (p < 0.001). Regarding the Stroop performance, the CN group was significantly better than the AD group in all 3 subscores (C, W and CW; p < 0.001; table 1).

Comparison of rCMglc between AD and CN Groups

Figure 2 shows the brain areas with significantly lower rCMglc in the AD group than in the CN group (FWE-corrected p < 0.05, k > 50), documenting the expected hypometabolism in the bilateral superior frontal gyri [Brodmann’s area (BA) 8, BA 10], the bilateral middle frontal gyri (BA 8, BA 10, BA 11, BA 46), the bilateral inferior frontal gyri (BA 45, BA 47), the right medial frontal gyrus (BA 25), the bilateral rectal gyrus (BA 11), the right ACC (BA 25), the left posterior cingulate gyrus (BA 31), the right caudate nucleus, the bilateral uncus...
(BA 28), the bilateral inferior parietal lobule (BA 40), the right parahippocampal gyrus (BA 37) and in the bilateral inferior temporal gyri (BA 20). No voxel was observed with significantly increased rCMglc in AD.

**Correlations between SI Scores and rCMglc of the Prefrontal ROIs in AD**

Among the 6 SI scores used in this study, SI-1 and SI-6 scores were significantly correlated with rCMglc of the bilateral DLPFC and bilateral ACC in the overall AD group, while SI-2, SI-4 and SI-5 scores were not correlated with the glucose metabolism of any prefrontal
region. In contrast, the SI-3 score was only correlated with bilateral DLPFC metabolism (table 2). Referring to these results and the simplicity of scoring, we selected the SI-1 score as a representative score of SI effects for further analyses.

To explore the effect of global AD severity on the relationship between SI effect and rCMglc, we divided the patients with AD into two severity subgroups according to the CDR Sum of Boxes score (SOB): the less severe AD group (CDR SOB ≤ 4, which equals the median CDR SOB score of all AD patients) and the more severe AD group (CDR SOB > 4). Positive correlations between the SI-1 score and rCMglc were found in all 6 prefrontal regions [partial correlation coefficient (r_p) = 0.408, p = 0.001 for the left DLPFC; r_p = 0.454, p < 0.001 for the right DLPFC; r_p = 0.385, p = 0.001 for the left ACC; r_p = 0.400, p = 0.001 for the right ACC; r_p = 0.336, p = 0.006 for the left OFC; r_p = 0.363, p = 0.003 for the right OFC] for the less severe subgroup, whereas no such correlations were found for the more severe subgroup.

Voxel-Based Correlations between SI Scores and rCMglc of the Whole Brain in AD

Using VBA (fig. 3; table 3), we found significant correlations between SI-1 scores and rCMglc of the left inferior parietal lobule, left middle temporal gyrus and left middle frontal gyrus in the total AD group. We also performed a subgroup analysis for each subgroup. In the less severe group (CDR SOB ≤ 4), there were significant correlations between SI-1 scores and rCMglc in very similar but wider regions compared with regions found to be significant in the total AD group. In contrast, no significant correlations existed at the threshold level of p < 0.05 (FWE corrected).

Discussion

This is the first study to demonstrate functional neuroanatomical correlates in Stroop performance in AD patients. Our ROI-based approach confirmed that the Stroop effect shows a good correlation with rCMglc of the prefrontal cortices, including the bilateral DLPFCs and ACCs in AD. Applying VBA with no a priori hypothesis, we also found that rCMglc in the posterior cortical regions, including the left inferior parietal lobule, temporal

<table>
<thead>
<tr>
<th>SI scores</th>
<th>DLPFC_L</th>
<th>DLPFC_R</th>
<th>OFC_L</th>
<th>OFC_R</th>
<th>ACC_L</th>
<th>ACC_R</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI-1</td>
<td>Correlation</td>
<td>0.231</td>
<td>0.304</td>
<td>0.129</td>
<td>0.199</td>
<td>0.252</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>&lt;0.001</td>
<td>0.022</td>
<td>0.004</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>SI-2</td>
<td>Correlation</td>
<td>-0.015</td>
<td>0.019</td>
<td>0.020</td>
<td>0.013</td>
<td>-0.055</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>0.863</td>
<td>0.825</td>
<td>0.818</td>
<td>0.884</td>
<td>0.528</td>
</tr>
<tr>
<td>SI-3</td>
<td>Correlation</td>
<td>-0.299</td>
<td>-0.270</td>
<td>0.005</td>
<td>-0.065</td>
<td>-0.171</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>&lt;0.001</td>
<td>0.002</td>
<td>0.461</td>
<td>0.050</td>
<td>0.075</td>
</tr>
<tr>
<td>SI-4</td>
<td>Correlation</td>
<td>0.133</td>
<td>0.162</td>
<td>0.051</td>
<td>0.103</td>
<td>0.172</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>0.128</td>
<td>0.063</td>
<td>0.560</td>
<td>0.238</td>
<td>0.049</td>
</tr>
<tr>
<td>SI-5</td>
<td>Correlation</td>
<td>-0.014</td>
<td>-0.032</td>
<td>-0.042</td>
<td>-0.025</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>0.876</td>
<td>0.719</td>
<td>0.634</td>
<td>0.777</td>
<td>0.787</td>
</tr>
<tr>
<td>SI-6</td>
<td>Correlation</td>
<td>0.261</td>
<td>0.301</td>
<td>0.160</td>
<td>0.205</td>
<td>0.279</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>&lt;0.001</td>
<td>0.001</td>
<td>0.019</td>
<td>0.003</td>
<td>0.003</td>
</tr>
</tbody>
</table>

* See the text for the definition of each SI score.

Significant correlations (p < 0.0083, Bonferroni corrected, two tailed) are indicated with bold characters. _L = Left; _R = right.
cortex, posterior cingulate cortex, precuneus and the prefrontal regions, is associated with the Stroop effect in AD.

Since the Stroop test has been regarded as the 'Frontal Function Test' [30], which mostly involves the DLPFC [31–38] and ACC [34, 35, 37–46], we used an ROI-based approach to confirm the correlation between the degree of the Stroop effect and rCMglc of the prefrontal cortices including bilateral DLPFCs and bilateral ACCs in the whole AD group. In this study, the previously proposed 6 scoring methods for SI reflecting the Stroop effect demonstrated different correlation patterns with prefrontal rCMglc. SI-1 [19] and SI-6 [23] scores were significantly correlated with bilateral DLPFC and ACC rCMglc, and the SI-3 score [21] showed significant correlation mainly with bilateral DLPFC rCMglc. In contrast, the other SI scores, including SI-2, SI-4 and SI-5, demonstrated no significant correlation with prefrontal area rCMglc. Our results support the importance of the prefrontal function for the Stroop effect in AD patients, which could be revealed only with the help of the proper SCWT SI scoring method.

Considering the ROI-based analysis results and the simplicity of the scoring method, we selected the SI-1 score as a representative of the Stroop effect for further VBA. For the total AD group, significant relationships between the Stroop effect and rCMglc of the left DLPFC, the inferior parietal lobule and the middle temporal cortex were found. This is in accordance with previous functional neuroimaging studies, which indicated that the Stroop effect is re-

<table>
<thead>
<tr>
<th>Table 3. Brain areas showing significant correlations between the SI-1 score and rCMglc in the AD group¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain region</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Whole AD group²</td>
</tr>
<tr>
<td>Lt inferior parietal lobule</td>
</tr>
<tr>
<td>Lt middle temporal gyrus</td>
</tr>
<tr>
<td>Lt middle frontal gyrus</td>
</tr>
<tr>
<td>Less severe group</td>
</tr>
<tr>
<td>Lt supramarginal gyrus</td>
</tr>
<tr>
<td>Lt middle temporal gyrus</td>
</tr>
<tr>
<td>Lt superior temporal gyrus</td>
</tr>
<tr>
<td>Lt superior temporal gyrus</td>
</tr>
<tr>
<td>Lt insula</td>
</tr>
<tr>
<td>Lt cingulate gyrus</td>
</tr>
<tr>
<td>Lt precuneus</td>
</tr>
<tr>
<td>Lt inferior parietal lobule</td>
</tr>
<tr>
<td>Lt inferior temporal gyrus</td>
</tr>
<tr>
<td>Lt posterior cingulate</td>
</tr>
<tr>
<td>Lt fusiform gyrus</td>
</tr>
<tr>
<td>Lt superior frontal gyrus</td>
</tr>
<tr>
<td>Lt middle frontal gyrus</td>
</tr>
<tr>
<td>Lt putamen</td>
</tr>
<tr>
<td>Rt middle frontal gyrus</td>
</tr>
<tr>
<td>Rt angular gyrus</td>
</tr>
<tr>
<td>Rt superior frontal gyrus</td>
</tr>
</tbody>
</table>

¹ Coordinates (x, y and z) refer to a standard stereotactical space. Each coordinate indicates the voxel location with the highest z score within each brain region. Lt = Left, Rt = right.
² FWE-corrected p <0.05, k > 50.
lated to the brain areas of the posterior parietal cortex [47], inferior parietal lobule [37, 40] and inferior temporal cortex [37]. Likewise, a postmortem brain study of AD patients showed that the degree of the Stroop effect was correlated with the increase in neurofibrillary tangles in the midfrontal area and the other diffuse cortical areas of the hippocampus, inferior parietal lobule and superior temporal cortex [8]. The results of the current study, together with those from previous ones, indicated that the Stroop effect was not only closely related with prefrontal activity but also with parietotemporal activity, supporting the ‘parallel-distributed processing models’ for the Stroop effect [48].

Because the brain areas involved in AD pathologies progressively expand as the clinical severity of AD increases [49], the neuroanatomical substrates related to the Stroop effect are expected to change from earlier to later stages of AD. Therefore, we divided the AD patients into two subgroups based on the severity of SOB, and explored the relationship between the
Stroop effect and regional brain function separately within each subgroup. For the less severe group, the Stroop effect was closely associated with hypometabolism in brain regions similar to the regions with significant relationships between the Stroop effect and glucose metabolism in the whole AD group. Since the prefrontal cortex is typically affected in later stages of AD [11, 12], the association between the Stroop effect and frontal function may be related with compensatory neuronal activation in the region to cover up the impairment in the medial temporal lobe [50, 51]. However, this association was not found for more severe AD patients. The lack of an association between the Stroop effect and metabolism suggests that the Stroop task is not a valid measure for regional brain function in more severe AD. The paucity of association might be the result of a ‘floor effect’ of SI score, rCMglc or both in the relatively advanced stages of AD.

In conclusion, this study showed functional neuroanatomical substrates of the Stroop effect in AD patients. Our results support the idea that the Stroop effect in AD patients depends on the functional state of prefrontal cortices, similar to the findings from brain activation studies. In addition, some parts of the parietotemporal region also appear to be responsible for the Stroop effect in AD individuals. In terms of the influence of disease severity, the paucity of a relevant neuroanatomical correlation with the Stroop effect in the more severe AD group suggests that SCWT might not be a proper neuropsychological measure of executive function in advanced AD.

Acknowledgments

This study was supported by a grant from the Korea Healthcare Technology R & D Project, Ministry of Health, Welfare & Family Affairs, Republic of Korea (grant No. A092145).

References


