Reproducibility of Activation Maps for Longitudinal Studies of Visual Function by Functional Magnetic Resonance Imaging

Jing Ming,1,2 Keith R. Thulborn,2,3 and Janet P. Szlyk1,2,4,5

PURPOSE. To test the intra- and intersubject reproducibility of brain activation patterns that underlie visually guided saccades and word recognition in normally sighted subjects and patients with macular degeneration using functional magnetic resonance imaging (fMRI).

METHODS. Ten normally sighted subjects and five patients with macular degeneration were asked to perform two visually guided saccade tasks and two word-recognition tasks during fMRI with behavioral monitoring. The fMRI measurements were repeated three times at intervals of at least 4 weeks between sessions. The intersubject reproducibility of the brain activation patterns was examined in a model-independent manner by comparing the distributions of activation across the frontal, parietal, temporal, and occipital brain lobes using Intraclass Correlation Coefficients (ICCs). Intersubject reproducibility was examined by repeated-measure ANOVA.

RESULTS. Control subjects showed overall higher intersubject reproducibility of brain activation patterns (75% ICCs > 0.5) than that of patients with macular degeneration (56% ICCs > 0.5). The intersubject reproducibility for the patients improved when the target location was fixed, as in the word-recognition tasks (75% ICCs > 0.5), compared with the visually saccade tasks (37% ICCs > 0.5). Intersubject variability of brain activation patterns was strikingly high for both the control and patient groups.

CONCLUSIONS. The fMRI method can serve as a reliable within-subjects measure of brain activation that has potential for measuring longitudinal changes in brain networks associated with rehabilitation training. Striking intersubject variability reflected at the level of lobes of the brain among control subjects with similar behavioral performance, suggests individual analysis is necessary when implementing longitudinal brain activation studies. (Invest Ophtalmol Vis Sci. 2012;53:6153–6163) DOI:10.1167/iovs.11-8375

Macular degeneration is the most common visual impairment in persons over 50 years of age in the United States.1,2 The deficits in visual function as the result of macular degeneration are debilitating, because individuals lose their abilities to carry out many daily activities that require fine spatial detail recognition, such as reading.3–4 When the central visual acuity becomes progressively poorer, patients will naturally adopt one or multiple peripheral retinal loci (PRL) to substitute for the diseased fovea.5,6 The use of a PRL is often effortful and fatiguing, because it involves unnatural oculomotor control.7 Normal visual signals that correspond to a significant portion of the visual cortex associated with the macular area are disrupted. The success of some patients with macular degeneration in the successful negotiation of activities of daily life indicates that the brain is capable of compensation for this disruption.8–10 If we can understand this natural process, we may be able to enhance this process for other patients during visual rehabilitation.

Our long-term goal is to use blood oxygenation level-dependent (BOLD) functional magnetic resonance imaging (fMRI) to investigate the biological basis of such compensation mechanisms, and ultimately use this information in the design and to measure the effectiveness of visual rehabilitation programs. BOLD fMRI, characterized by high spatial resolution and noninvasiveness, is a powerful tool to investigate human brain function across distributed networks. In particular, fMRI is used to assess the cortical reorganization and brain plasticity of patients with different neural diseases.11–15 However, great variability occurs with repetition of the same fMRI studies on the same subject.16 There are a number of potential sources of this variability. The essential one may be the nature of BOLD contrast that arises from the local cerebral hemodynamic response to the increased neuronal activity associated with performing a task.17–19 BOLD contrast is small (1–5% signal change), delayed from the stimulus, and requires signal averaging with statistical methods for detection. In addition, BOLD contrast is also subject to psychologic factors, including mood, attention, strategy, and learning, as well as to non-psychologic factors, including differences in scanner equipment, head motion, and physiologic condition. Therefore, measuring and controlling these factors are necessary for effective design and accurate interpretation of fMRI studies.

When using fMRI in patients with macular degeneration, we designed a set of paradigms to measure the changes in brain activation initially due to the disease, and then rehabilitation. The paradigms emphasized the rehabilitation strategies. Most current fMRI studies of patients with macular degeneration are focused on visual cortex reorganization.20–26 Simple visual...
patients with age-related macular degeneration (AMD). These studies have shown that patients with AMD tend to recruit more cortical regions generally implicated in attention and effort when performing visual-saccade and word-recognition tasks.

It is essential to measure longitudinal intra- and intersubject reproducibility of brain activation patterns for these paradigms if they are to be used to investigate adaptations of brain networks in patients with macular degeneration during visual rehabilitation. Current fMRI reliability studies tend to use different indices as measures of reproducibility. These parameters include: (1) absolute activated voxel number within specific regions of interest (ROIs); (2) location of activated voxels; (3) the statistical significance or contrast intensity of activated voxels; and (4) proportion of the activated voxels within different ROIs. Proportion of the activated voxels within specific ROIs is considered to be a reliable index, for the sources of variability affect the whole brain in a similar fashion. Prior research with patients affected by stroke and patients with schizophrenia has shown greater variability in fMRI data for patients than for control subjects. However, moderate or even good reproducibility has been achieved among the patients with stroke when they perform a drawing task using the less affected hand.

In the present study, we report on the reproducibility of performance and brain activation patterns of normally sighted adults and patients with macular degeneration during the previously described visuospatial tasks. Unlike most other fMRI studies, we did not use specific ROI analysis based on a known network model. Such an approach may obscure different strategies used by different subjects to achieve success on any given visual task or changing strategies by a single subject on a repeated task. Instead, a model-independent analysis using the lobes of the brain was used to characterize the variability of global activation patterns. By separating the brain at the lobe level, any variability on this scale can be thought of as reflecting distinctly different cognitive patterns across sessions and across subjects. If present, it would reflect widely changing cognitive strategies. High variability in individual subjects would suggest a less consistent cognitive strategy. High variability among subjects would suggest different cognitive strategies.

**METHODS**

**Participants**

The demographic data of the control subjects (n = 10) and patients (n = 5) are found in Table 1. The control subjects consisted of younger normally sighted control subjects (n = 5, subjects 100–104, mean visual acuity (VA) = 20/18.4 ± 5.6) and older normally sighted control subjects (n = 5, subjects 200–204, VA = 20/21.6 ± 2.2). The patients consisted of those with the juvenile-onset macular degeneration (Stargardt disease, n = 2, subjects 300, 301) and those with AMD (n = 3, subjects 400–402). For all patients, the VA (for the better eye) was equal to 20/100. For each patient, the currently used PRL was assessed by our acuity microperimetry system developed in our laboratory with eye-tracking control capabilities. The PRL sizes and locations are also listed in Table 1. Signed informed consent, approved by Institutional Review Board of The University of Illinois, was obtained from each subject prior to the study.

**fMRI Paradigms**

The four tasks are illustrated in Figure 1. All paradigms used the same block design consisting of six cycles; each cycle was comprised of a rest period of 30 seconds followed by a stimulus period of 30 seconds.

### Table 1. Demographic Information for the Participants

<table>
<thead>
<tr>
<th>Case Number</th>
<th>Age (y)</th>
<th>Sex</th>
<th>Visual Acuity (VA)</th>
<th>PRL Size* (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>27</td>
<td>M</td>
<td>20/18</td>
<td>20/15</td>
</tr>
<tr>
<td>101</td>
<td>22</td>
<td>M</td>
<td>20/17</td>
<td>20/16</td>
</tr>
<tr>
<td>102</td>
<td>33</td>
<td>F</td>
<td>20/20</td>
<td>20/17</td>
</tr>
<tr>
<td>103</td>
<td>23</td>
<td>F</td>
<td>20/24</td>
<td>20/24</td>
</tr>
<tr>
<td>104</td>
<td>48</td>
<td>M</td>
<td>20/24</td>
<td>NA</td>
</tr>
<tr>
<td>200</td>
<td>70</td>
<td>M</td>
<td>20/21</td>
<td>20/20</td>
</tr>
<tr>
<td>201</td>
<td>63</td>
<td>M</td>
<td>20/25</td>
<td>20/24</td>
</tr>
<tr>
<td>202</td>
<td>61</td>
<td>M</td>
<td>20/22</td>
<td>20/20</td>
</tr>
<tr>
<td>203</td>
<td>54</td>
<td>F</td>
<td>20/24</td>
<td>20/24</td>
</tr>
<tr>
<td>204</td>
<td>65</td>
<td>F</td>
<td>20/20</td>
<td>20/20</td>
</tr>
<tr>
<td>Patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>44</td>
<td>F</td>
<td>20/60</td>
<td>20/32</td>
</tr>
<tr>
<td>301</td>
<td>55</td>
<td>F</td>
<td>20/400</td>
<td>20/450</td>
</tr>
<tr>
<td>400</td>
<td>64</td>
<td>M</td>
<td>20/76</td>
<td>20/91</td>
</tr>
<tr>
<td>401</td>
<td>77</td>
<td>F</td>
<td>20/121</td>
<td>20/166</td>
</tr>
<tr>
<td>402</td>
<td>82</td>
<td>F</td>
<td>20/520</td>
<td>20/560</td>
</tr>
</tbody>
</table>

* The PRL size for the better eye is shown.

stimuli (flashing checkerboard and pictures of natural scenes or objects) have been used in these previous studies to stimulate activation in the primary visual cortex. Beyond the shared interests of alterations in cortical activation relative to control subjects, the goal of this research has not been to focus on retinotopic mapping of the visual cortex, but instead, to investigate the brain networks that are associated with the coordination of eye movement (oculomotor control) and visual attention and recognition. These functions are highly congruent in normal-sighted people with the fovea being effortlessly placed on the point of visual attention automatically. In patients with macular degeneration, this process must be rebalanced to use a retinal location other than the fovea for successful viewing. Thus the paradigms have been designed to examine the oculomotor networks responsible for saccading the eyes to a simple moving target (dot) and visual language recognition (stationary target of increasing size from a single letter to short words and then longer words).

Four visuospatial tasks of increasing complexity have been developed with this research. The tasks encompass the oculomotor function of finding a target of variable size, or at an unpredictable location, and then the visual detection and recognition function of recognizing that target. Our battery of fMRI paradigms uses visually guided saccade tasks, similar to the function used in reading, to various targets (e.g., a moving dot, stationary one-letter, three-letter words, and six-letter words) of increasing size. These fMRI paradigms were designed to probe aspects of a reading rehabilitation program designed in our laboratory to train patients to use their PRLs more efficiently. All visual targets were designed at an appropriate size based on the patients' PRL sizes and visual acuity levels. Horizontal and vertical guidelines were present in the visual stimuli to help patients locate the visual target. All paradigms were designed to require subject feedback to provide performance parameters including accuracy and response time that reflect the maintained attention of the subject to the task. Previous studies in our laboratory have demonstrated that these paradigms elicit measurable brain activation in...
All paradigms began with 12 seconds of discarded data acquisition to ensure equilibrium of the longitudinal magnetization and ended with a 30-second stimulus condition. Orthogonal vertical and horizontal guidelines with a width of 0.4° were provided to help patients locate the targets in each paradigm. The size of each single target (a dot or single-letter) was 2°, which was within the sizes of the PRL of all patients (Table 1). Thus, each patient had the potential to capture the complete single target in the dot and letter paradigms, and a single letter in the word-recognition paradigm. All subjects and patients were studied with all paradigms in three separate sessions, each separated by approximately 4 weeks. The total interval covered by three sessions was around 8 weeks. The 8-week interval reflected the length of our reading rehabilitation program.

**Visually Guided Dot Saccade Task (DS).** The DS task required the subject to locate a dot (subtending 2° of visual angle, equivalent to a 20/600 stimulus) that was presented in one of seven possible unpredictable locations (at 3°, 6°, or 9° on either side of a center 0° location). Every 1.5 seconds, the position of this dot changed to the position immediately 3° to either the left, or right, of the previous position. The subject was required to indicate by pressing a response button when the dot changed to a cross (also subtending 2°). The rest condition for the DS task required maintained fixation on a 2° dot located at the center of the screen. In addition, the subjects were asked to respond by pressing a button when the dot became a cross. The finger switch responses were designed to maintain and measure attention during the paradigm.

**Letter Saccade Task (LS).** The LS task was the same basic task as the DS task except a 2° letter, that changed unpredictably, was substituted for the dot and a finger switch response was required when the letter presented was a vowel. The rest condition for the LS task required maintained fixation on a 2° cross located at the center of the screen. The finger switch responses were designed to measure not only attention but also demonstrate recognition of the target indicating that functional retina and presumably the PRL was being used for the task.

**Three-Letter and Six-Letter Word-Recognition Tasks (3L and 6L).** For this task, either a three-letter or a six-letter word was presented at the center of the screen. Subjects were asked to press one switch if the word represented a living thing, and the other switch if the word represented a nonliving thing. During the stimulus period, the word presented was changed every 3 seconds. Each letter subtended 2°. The rest condition required maintained fixation on a 2° cross that was located at the center of the screen. The use of a word simplified the planning of the eye movements because, once the first letter was found, the eye movements were always rightward as in reading. The longer word potentially required more eye movements, although word recognition/perceptual filling-in also plays an important role without requiring additional eye movements.

**Data Collection**

**Eye Movement Data Collection in the Low Vision Laboratory.** Eight control subjects (100–102, 104, 200–203) and all patients were trained prior to scanning on the four tasks outside of the scanner in our behavioral laboratory, and their eye movements during these practice sessions were recorded using an eye tracker that tracks the pupil and corneal reflections (Model 504; Applied Sciences Laboratory, Bedford, MA). The sampling and output rate of the eye tracker system was 60 Hz. During training, the subjects were seated at a viewing distance of 40 cm from the high-resolution display monitor (1024 x 768 pixels). Their head position was maintained by a forehead and chin support.

**Imaging Data Collection.** A 3.0-Tesla whole body scanner (EXCITE 2.0; GE Healthcare, Waukesha, WI) using spiral gradient echo, echo-planar imaging (plane = axial, repetition time (TR) = 2999 ms, time to echo (TE) = 25 ms, flip angle [FA] = 90°, bandwidth = 62 kHz, voxel size = 3.125 mm x 3.125 mm x 5 mm; acquisition matrix = 64 x 64, field of view (FOV) = 20 x 20 cm², slice thickness/gap = 3/1 mm/mm, slices = 34) was used for all image acquisitions. The paradigms were presented in the scanner projected onto a visor and coordinated with behavioral and physiological measurements (MRiX Technologies, Bannockburn, IL). When the subjects performed the tasks in the scanner, their eye movements were monitored within the scanner to ensure that every subject was performing the appropriate eye movements for the task (MRiX Technologies). Head position was stabilized with a tightly fitting head pillow in the volume radiofrequency coil.

Following the completion of each functional imaging session, a single 3D high-resolution anatomical scan was acquired (3D inversion recovery fast spoiled gradient recalled echo sequence, plane = axial, TR = 9ms, TE = 2.0 ms, FA = 25°, bandwidth = 15.6 kHz, acquisition matrix = 256 x 256, FOV = 22 x 16.5 cm², slice thickness/gap = 1.5/0.0 mm/mm, slices = 124).

**Data Analysis**

**Eye Movement Data Analysis.** The subjects’ eye movements, recorded outside the scanner, included x (horizontal) and y (vertical) coordinates of gaze and pupil diameter at each temporal sampling point. A blink was represented by a zero pupil diameter for a certain time period. The gaze positions from 0.04 second before a blink to 0.2 second after a blink were distorted, therefore were excluded from further quantification. In the present study, eye movement data within one representative cycle of each paradigm is presented. The eye movement pattern in the fixation condition was characterized by the gaze standard deviation (SD) during a complete block (30 seconds) of fixation in both the x (SD fixation x) and y (SD fixation y) directions. The eye movement pattern in the task condition was characterized by averaging the gaze SD during the late 2/3 time period of each stimulus trial in the x (Mean SD trial x) and the y (Mean SD trial y) directions. This is the period after the target has been found and is maintained presumably on the PRL.

The gaze locations within the beginning 1/3 time period of each trial were excluded because the saccadic eye movement used to relocate gaze on the new target during this period increased the gaze SD that was a measure of fixation ability. The Mann–Whitney test was used to compare the eye movement patterns between the control paradigms.
group and the patient group. All the analyses above were done using customized software (Matlab, The MathWorks, Inc., Natick, MA).

fMRI Data Analysis. A locally developed software package (NIVANA; MRix Technologies) and the software package AFNI\(^1^) were used to analyze the fMRI data. Data from any imaging session with head movement exceeding 1 mm (1/3 of a voxel dimension) were excluded from further analysis. This accounts for the data that are indicated as missing in the Results section. Overall, 16 of 150 (10.7\%) acquisitions have been excluded for control subjects, and 17 of 75 (22.7\%) acquisitions have been excluded for patients.

Brain activation maps were produced by comparing the voxel signal intensity between task condition and rest condition via a voxelwise two-tailed Student’s t test. Thresholds of different t-values were examined. Consistency does not increase with different thresholds, until a very high threshold is tested, when only the primary visual cortex is active and there are too few voxels to be meaningful. A lower threshold that includes the eye movement network known to operate in these tasks is appropriate for investigating consistency of responses across individuals. The data presented are for a uniform t-value equal to 3 for all subjects. Subsequently, individual functional activation maps for each paradigm were superimposed over the structural images from the same individual.

Activation patterns, characterized as the percentages of activated voxels in each lobe (frontal, parietal, temporal, and occipital) of the brain collapsed across hemispheres, were used to determine the fMRI reproducibility across sessions and subjects. Intraclass correlations were applied to test for the significant intrasubject variations of the activation patterns and the performance data for each paradigm across sessions. In addition, repeated-measure ANOVA was used to investigate the intersubject variability for each paradigm and session.

Percentage of Voxel Activation

The lobes of the cerebrum, (frontal, parietal, temporal, and occipital) were defined according to sulcal landmarks using each individual subject’s anatomic images, and defined by the neuroradiologist on our team (KRT). The frontal lobe was defined as the tissue anterior to the central sulcus and superior to the Sylvian fissure (lateral sulcus). The parietal lobe was defined as the tissue posterior to the central sulcus, medial to the Sylvian fissure and superior to parietooccipital fissure. The temporal lobe was defined as the brain parenchyma lateral to the Sylvian fissure and anterior to the lateral projection of the parietooccipital sulcus. The occipital lobe was defined as all tissue posterior and inferior to the parietooccipital fissure and above the tentorium.

The number of activated voxels for each lobe (\(V_i\)) was counted and calculated as a percentage (\(P_i\)) of the total number of activated voxels in all four lobes of the cerebrum (\(V_t\)):

\[
P_i = \frac{V_i}{V_t}
\]

The percentage activation in each lobe was then displayed in a pie chart to summarize the activation patterns of each paradigm for each subject for each session.

Intraclass Correlation Coefficient

The Intraclass Correlation Coefficient (ICC) is well known in psychometry as an index of reliability\(^6^) and has been applied to determine the reliability of the fMRI activation patterns. There are several ICC model alternatives, and this study uses the one-way ANOVA model (\(ICC_1\)) that is specifically suitable for this type of fMRI data.\(^3^,\(^1^)\)

In this study, only within-subject variance (\(\sigma_{\text{within}}\)) and between-subject variance (\(\sigma_{\text{between}}\)) are considered. ICC estimates the proportion of variance that is due to differences between the subjects rather than differences due to the measurements; therefore, it can be described by the following equations 2 and 3. Thus, smaller values of \(\sigma_{\text{within}}\) result in a higher ICC.

\[
ICC = \frac{\sigma_{\text{within}}}{\sigma_{\text{total}}}
\]

\[
\sigma_{\text{total}} = \sigma_{\text{within}} + \sigma_{\text{between}}
\]

From a computational view, ICC, \(\sigma_{\text{within}}\) and \(\sigma_{\text{between}}\) are represented by the mean square between subjects (MSB) and mean square within subjects (MSW). In equations 4 and 5, index \(i\) represents \(i\)th subject, and index \(j\) represents the \(j\)th subject. The symbol \(k\) is the number of total sessions, which is equal to 3. The total subject number, \(n\), is different in different paradigms due to data that were excluded due to head movement. We applied the ICC analysis on the activation percentage for each lobe of the brain for each paradigm across the three sessions.

\[
MSB = k \sum_{j=1}^{n} (\bar{X}_j - \bar{X})^2 / (n - 1)
\]

\[
MSW = \sum_{i=1}^{k} \sum_{j=1}^{n} (X_{ij} - \bar{X}_j)^2 / (n - 1)
\]

\[
ICC_1 = \frac{MSB - MSW}{MSB + (k - 1)MSW}
\]

RESULTS

Eye Movement Data

Figure 2 shows the eye movement patterns collected in the behavioral laboratory for the eight control subjects and the five patients during fixation (upper row in each pair) and the active condition (lower row in each pair) for each of the four paradigms. Table 2 is a quantitative summary of these eye movement patterns for these same subjects and patients. The eye movement pattern was quantified by the previously defined variables: SD_fixation_x, SD_fixation_y, Mean_SD_trial_x, Mean_SD_trial_y. The results of the Mann–Whitney test used to compare these variables between the eight control subjects and the five patients is indicated in Table 2. From Figure 2 and Table 2, the patients show greater variation in fixation and an increased variation in saccades to all types of targets compared with the control subjects. However, it is also evident that all of the participants were able to perform the tasks of moving their eyes to targets across the visual field.

Performance Data

Figure 3 illustrates the performance data during fMRI for all paradigms, including task accuracy and response time. The error bars represent the SD values of accuracy and response time across the three sessions. The patients have significantly lower mean accuracy and longer mean response time across sessions than the control subjects (\(P < 0.05\)) for all four paradigms.

Table 3 is a numerical summary of the mean accuracy and response times across the three sessions, as well as the SD of accuracy and response time across the three sessions for both the control and the patient groups in each paradigm. From both Figure 3 and Table 3, we observed that all control subjects achieved a similar performance level, whereas there was great intersubject variability for performance within the patient group. Variance ratio test (F test) results confirmed our observation as indicated in Table 3. Furthermore, the control subjects tended to achieve more consistent intrasubject behavioral patterns compared with those of the patients. This
Figure 2. Eye movement patterns for (a) control subjects ($n = 8$) and (b) patients ($n = 5$) in all 4 paradigms. The upper and lower rows in (a) and (b) show the eye movement patterns in the fixation condition and during the active condition of each paradigm for the controls and patients, respectively. All eye movement data are ordered by task; dot saccade (DS), letter saccade (LS), three-letter word recognition (3L), and six-letter word recognition (6L) from left to right. The colors indicate the patterns for individual subjects. Control subjects: 100 = red, 101 = green, 102 = cyan, 104 = magenta, 200 = yellow, 201 = black, 202 = red in small circle, 203 = blue. Patients: 300 = red, 301 = green, 400 = magenta, 401 = yellow, 402 = blue.

Table 2. Quantification of Eye Movement Patterns for Four Paradigms

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>DS</th>
<th>LS</th>
<th>3L</th>
<th>6L</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gaze SD during fixation condition in degrees</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD fixation x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>8</td>
<td>$0.42 \pm 0.17$</td>
<td>$0.59 \pm 0.16$</td>
<td>$0.62 \pm 0.38$</td>
<td>$0.42 \pm 0.19$</td>
</tr>
<tr>
<td>Patient</td>
<td>5</td>
<td>$1.37 \pm 1.02^*$</td>
<td>$1.51 \pm 0.99^*$</td>
<td>$1.66 \pm 1.28$</td>
<td>$1.77 \pm 0.64^{**}$</td>
</tr>
<tr>
<td>SD fixation y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>8</td>
<td>$0.84 \pm 0.52$</td>
<td>$1.25 \pm 0.70$</td>
<td>$1.26 \pm 0.70$</td>
<td>$0.93 \pm 0.34$</td>
</tr>
<tr>
<td>Patient</td>
<td>5</td>
<td>$1.70 \pm 1.32$</td>
<td>$2.05 \pm 1.16$</td>
<td>$2.52 \pm 1.36^*$</td>
<td>$2.68 \pm 1.19^{**}$</td>
</tr>
<tr>
<td><strong>Averaged gaze SD during each stimulus trial in task condition in degrees</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean SD trial x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>8</td>
<td>$0.35 \pm 0.09$</td>
<td>$0.35 \pm 0.16$</td>
<td>$0.44 \pm 0.22$</td>
<td>$0.75 \pm 0.51$</td>
</tr>
<tr>
<td>Patient</td>
<td>5</td>
<td>$0.77 \pm 0.34^{**}$</td>
<td>$1.02 \pm 0.50^{**}$</td>
<td>$1.26 \pm 0.82^*$</td>
<td>$1.60 \pm 0.83^*$</td>
</tr>
<tr>
<td>Mean SD trial y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>8</td>
<td>$0.20 \pm 0.07$</td>
<td>$0.32 \pm 0.20$</td>
<td>$0.71 \pm 0.57$</td>
<td>$0.63 \pm 0.37$</td>
</tr>
<tr>
<td>Patient</td>
<td>5</td>
<td>$0.82 \pm 0.61^{**}$</td>
<td>$1.09 \pm 0.98$</td>
<td>$1.61 \pm 1.01$</td>
<td>$1.20 \pm 0.61$</td>
</tr>
</tbody>
</table>

SD fixation x and SD fixation y represent the SD of the gaze fixation position in the x (horizontal) direction and the y (vertical) direction, respectively, during the 30-second fixation condition. Mean SD trial x and Mean SD trial y represent the SD of the average gaze position in the x direction and y direction, respectively, during the late 2/3 time period of each stimulus trial in each paradigm. Statistical significance for comparisons of control and patient groups is indicated by asterisks.

* $P < 0.05$.

** $P < 0.01$. 

Downloaded From: http://iovs.arvojournals.org/pdfforaccess.axd?url=/data/Journals/IOVS/932977/ on 12/04/2015
tendency is reflected by the lower SD values of accuracy and response time across sessions in all paradigms for the control group. However, the Mann–Whitney test that was used to compare the SD of accuracy and response times across sessions between the control and patient groups did not reach general significance (the only statistically significant result was for the DS paradigm), probably because of the small sample size. The intrasubject variability of the patients on performance appeared to be closely related to their visual acuity and PRL size. Patient 300 and Patient 400 who achieved relatively consistent performance across sessions had better visual acuity (VA Patient 300 = 20/32, VA Patient 400 = 20/76) and larger PRL sizes (12°) compared with that of the other three patients.

Pie Chart Matrix and ICC Summary

Figures 4 and 5 contain the pie charts illustrating the distribution of activation across the brain for each control subject and patient, respectively, for each task paradigm, accompanied by the ICC results. Visual inspection of these pie chart matrices indicates two major points. First, there is relative consistency of distribution patterns across sessions for both the control subjects and patients, with the control subjects having higher consistency than that of patients. Second, a review of the patterns across subjects and patients (along rows), in Figures 4 and 5, shows striking variability among members of both the control and patient groups.

For control subjects, 12 of 16 (75%) ICCs > 0.5 and 15 of 16 (94%) correlation coefficients were significant, indicating strong intrasession consistency. For patients, 9 of 16 (53%) ICCs > 0.5 and 7 of 16 (44%) correlation coefficients were significant. The intrasubject consistency for patients improved when the target location was fixed, as in the word-recognition tasks. The patients showed 6 of 8 (75%) ICCs > 0.5 and 5 of 8 (63%) significant ICCs in 3L and 6L paradigms. Considering the statistical significance of ICCs is skewed by a smaller sample size of the patient group, we use the absolute ICC value to represent intrasubject reproducibility of the brain activation patterns for control subjects and the patients. Overall, patients showed reasonably good reproducibility in word-recognition tasks. The reproducibility of activation within each lobe across sessions for each paradigm was demonstrated for the control subjects and patients. This lobar region of interest analysis permitted patterns to be compared without reference to specific networks that appear to have large variations among the control subjects and patients, with the control subjects having higher consistency than that of patients. Second, a review of the patterns across subjects and patients (along rows), in Figures 4 and 5, shows striking variability among members of both the control and patient groups.

### Table 3. Numerical Summary of fMRI Performance Data

<table>
<thead>
<tr>
<th></th>
<th>DS</th>
<th>LS</th>
<th>3L</th>
<th>6L</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Patient</td>
<td>Control</td>
<td>Patient</td>
</tr>
<tr>
<td>Mean Acc across sessions (%)</td>
<td>99.3 ± 1.7</td>
<td>75.4 ± 27.8**</td>
<td>99.0 ± 1.5</td>
<td>91.1 ± 10.6**</td>
</tr>
<tr>
<td>SD of Acc across sessions (%)</td>
<td>1.2 ± 2.9</td>
<td>8.4 ± 7.6</td>
<td>1.3 ± 18</td>
<td>5 ± 4.7</td>
</tr>
<tr>
<td>Mean RT across sessions (s)</td>
<td>0.5 ± 0.05</td>
<td>0.9 ± 0.33**</td>
<td>0.5 ± 0.07</td>
<td>1 ± 0.38**</td>
</tr>
<tr>
<td>SD of RT across sessions (s)</td>
<td>0 ± 0.02</td>
<td>0.1 ± 0.13*</td>
<td>0 ± 0.02</td>
<td>0.1 ± 0.07</td>
</tr>
</tbody>
</table>

A variance ratio test (F test) was used to compare the variance of mean Acc and mean RT across sessions between control group and patient group. A Mann–Whitney test was used to compare the SD of accuracy and response time across sessions between the control and patient groups. Statistical significance is indicated by asterisks. Acc = accuracy; RT = response time; DS = dot saccade; LS = letter saccade; 3L = 3-letter word recognition; 6L = 6-letter word recognition.

* P < 0.05.  
** P < 0.01.
individuals, although being reproducible across sessions for individuals. Representative brain activations for a control subject (203) and a patient with Stargardt disease (300) across three sessions for each paradigm are shown in Figures 6 and 7, respectively. The maps display consistent areas of activation across the sessions in regions of the brain associated with eye movement, right hand sensorimotor movement (finger switch), and reading, including the frontal eye fields, supplementary eye fields, prefrontal cortex, intraparietal sulcus, and the visual cortex (V1, V2/V3, MT/V5).

**Intersubject Variability**

The repeated-measure ANOVA, used to quantify the intersubject variability for the activation patterns for the control subjects and the patients, is summarized in Table 4. For the control subjects, statistical significance was achieved for 16 of 16 F values for a between-subjects effect, indicating substantial intersubject variability. By comparison, none of the F values for the within-subjects effect (Session factor) showed statistical significance, indicating little variability across sessions. This is consistent with the ICC results for the control subjects. For the patients, statistical significance was achieved for 11 of 16 F values for the between-subjects effect. Only 1 of the 16 F values for the within-subjects effect showed statistical significance. This result indicates substantially larger brain activation variability for the Subject factor than for the Session factor among the patients.

**Intrasubject Reproducibility**

Although data from both normally sighted control subjects and visually compromised patients show statistically significant intrasubject consistency, the control subjects show higher...
consistency than that of patients with macular degeneration, as would be expected given their better performance. This result is consistent with previous studies on fMRI reliability cited earlier that were involved with patients affected by stroke and schizophrenia.

Given the criterion that ICC > 0.5 indicates good reproducibility, patients with macular degeneration show reasonably good reproducibility in word-recognition tasks when the target location was fixed (6 of 8 ICCs > 0.5, ranging from 0.61–0.93). However, with moving targets in the visually guided saccade paradigms, the patients showed greater variability in brain activation pattern (3 of 8 ICCs > 0.5, ranging from 0.51–0.70). This may reflect reduced skill or variable strategies being used to find the target. Across four paradigms, the intrasubject reproducibility was highest for the occipital lobe (4 of 4 ICCs > 0.5, ranging from 0.51–0.86). Because the activation in the occipital lobe relates to visual processing, rather than the oculomotor pathways of the frontal and parietal lobes used to direct the eyes and presumably the PRL of the retina to the target, these results suggest the major

Figure 6. Representative activation patterns for an older control subject (203) across each of 3 scan sessions (top, first session; middle, second session; bottom, third session) for each of the 4 paradigms: (a) dot saccade paradigm, (b) single-letter saccade paradigm, (c) three-letter word paradigm, and (d) six-letter word paradigm. Head motion was similar in each session and less than 33% of a voxel dimension (3 mm isotropic). Performances on recognition tasks were better than 90% (Fig. 3).

Figure 7. Representative activation patterns for a patient (300) across each of 3 scan sessions (top, first session; middle, second session; bottom, third session) for each of the 4 paradigms: (a) dot saccade paradigm, (b) single-letter saccade paradigm, (c) three-letter word paradigm, and (d) six-letter word paradigm. Head motion was similar in each session and less than 33% of a voxel dimension (3 mm isotropic). Performances on recognition tasks were better than 90% (Fig. 3).
challenge for patients with macular degeneration may be oculomotor control. Unlike normally sighted individuals with congruent gaze and attention at the fovea, the patients with AMD must retrain the oculomotor control network to place the PRL rather than the destroyed fovea on the target. This increased difficulty of eccentric viewing may explain why the patients achieved more stable brain activation patterns across sessions for paradigms with a fixed target with fewer eye movements as in the 3L paradigm. As demonstrated by our previous studies, patients tended to recruit more higher-order cortical regions, such as prefrontal cortex in the frontal lobe, and intraparietal sulcus in the parietal lobe, to compensate for their compromised visual system.27,28 However, it may be difficult for them to naturally develop a stable strategy to efficiently control eye movement for eccentric viewing. This unstable high-order compensation strategy may lead to the larger intersubject variability for the frontal lobe and the parietal lobe compared with the occipital lobe for patients with macular degeneration. Rehabilitation training may help patients to establish more efficient eccentric viewing strategies for reading by achieving better oculomotor control using a congruent gaze and attention at the fovea, the patients with a neural deficit, the general activation index extracted from a larger amount of voxel data may have higher reproducibility. Care must be taken when comparing activation patterns across different patient populations where the disease affects the brain directly (e.g., stroke, schizophrenia) and where the disease affects the input into the brain (e.g., AMD). The compensation mechanisms and the rehabilitation potential may be quite different.

### Intersubject Reproducibility

The striking intersubject variability of brain activation pattern is found in both the control and patient groups. This is consistent with the findings of previous reproducibility studies.57-59 This intersubject brain activation pattern variability at the lobe level must be considered when we try to generate multisubject level inferences for these paradigms. Currently, multisubject analysis must be considered when we try to generate multisubject level inferences for these paradigms. Currently, multisubject analysis falls into two types. One is a fixed-effects model, which assumes that the experimental stimulus has the same effect on the BOLD signal for every subject; the other one is a random-effects model, which assumes that the experimental stimulus could have a different effect on each subject, and the effect across subjects fits to a specific distribution, typically a normal distribution. Our experimental results do not fit the fixed-effects analysis. The intersubject variability in the patient group may also be due to a combination of factors such as age, lesion size, and acuity, as well as differences in cognitive strategies.51 Even among control subjects who achieve very similar performance levels in each paradigm, and are of similar age and acuities (as in the case of control subjects 101 and 102), the activation patterns analyzed across the lobes of the brain show differences that suggest different networking strategies for processing visual tasks.

Control for this intersubject variability is required to perform group analyses. A recent study conducted by

### Table 4. Intersubject and Intersession Variability from a Two-Way ANOVA of Brain Activation Patterns for Four Paradigms

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>Lobe</th>
<th>Control Group</th>
<th>Patient Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Within-Subject F Value</td>
<td>Between-Subject F Value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Dot saccade paradigm</td>
<td>Frontal</td>
<td>1.36 0.29</td>
<td>28.42 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>0.95 0.41</td>
<td>34.54 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Temporal</td>
<td>1.48 0.26</td>
<td>10.49 &lt;0.05*</td>
</tr>
<tr>
<td></td>
<td>Occipital</td>
<td>0.09 0.91</td>
<td>214.16 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.77 0.21</td>
<td>106.95 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>1.77 0.21</td>
<td>18.47 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Temporal</td>
<td>1.67 0.22</td>
<td>24.79 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Occipital</td>
<td>0.95 0.41</td>
<td>193.15 &lt;0.01**</td>
</tr>
<tr>
<td>Letter saccade paradigm</td>
<td>Frontal</td>
<td>1.08 0.36</td>
<td>60.79 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>1.39 0.28</td>
<td>51.26 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Temporal</td>
<td>0.06 0.94</td>
<td>41.65 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Occipital</td>
<td>0.87 0.44</td>
<td>42.90 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.44 0.97</td>
<td>23.89 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24.23 &lt;0.01**</td>
<td>0.44 0.67</td>
</tr>
<tr>
<td>3-Letter word paradigm</td>
<td>Frontal</td>
<td>0.84 0.45</td>
<td>37.78 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>0.80 0.47</td>
<td>23.89 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Temporal</td>
<td>3.65 0.053</td>
<td>24.23 &lt;0.01**</td>
</tr>
<tr>
<td></td>
<td>Occipital</td>
<td>0.11 0.90</td>
<td>100.23 &lt;0.01**</td>
</tr>
</tbody>
</table>

Session and subject are treated as two random factors. Significant correlations are indicated by asterisks.

* P < 0.05.
** P < 0.01.
temporal SNR is represented by the average signal-to-noise ratio (SNR) for each subject. The global activation clustering algorithms, which can separate the zonation method that can be applied when SNR is reproducible degeneration. Our approach of using regional activation strategies, their method provides a possible way to create a general brain activation pattern in patients with macular degeneration. Our approach of using regional activation proportions of the total brain activation is a simple normalization method that can be applied when SNR is reproducible with stable scanner performance and subjects who are well trained with the paradigms and imaging environment. Brain activation clustering algorithms, which can separate the subjects into subgroups with more homogeneous brain activation patterns, provides another way to perform more accurate group analysis.

In conclusion, the significant repeatability of the fMRI brain activation patterns from our paradigms within individuals shows that fMRI can possibly serve as a reliable measure of cortical function and has potential for measuring outcomes of vision rehabilitation. Furthermore, this study has suggested that the activation proportion method could serve as a reliable index to measure reproducibility. The poorer intrasubject reproducibility of patients with macular degeneration in visually guided saccade tasks suggests that these patients may lack a successful oculomotor processing strategy for coordinating eye movement for finding and detecting a target. The implication of these findings is that the success of training visually compromised patients to use specific eccentric viewing strategies to maximize use of their PRL may be reflected by improved performance and a reduction in the variability of brain activation patterns. Considering the striking intersubject variability, even in control subjects, individual analysis is necessary when conducting longitudinal brain activation studies for the patients using fMRI, and group mapping should be applied only after considering the intersubject variability.

Acknowledgments
The authors thank Marlos Viana, PhD, for his advice on the methods of statistical analysis and Gerald A. Fishman, MD, for his referral of the patients with Stargardt disease.

References


