Scene Categorization in Alzheimer’s Disease: A Saccadic Choice Task

Quentin Lenoble\textsuperscript{a}  Giovanna Bubbico\textsuperscript{a}  Sébastien Szaffarczyk\textsuperscript{a}  Florence Pasquier\textsuperscript{b}  Muriel Boucart\textsuperscript{a}

\textsuperscript{a}Laboratoire de Neurosciences Fonctionnelles et Pathologies, Université Lille Nord de France, CNRS, and \textsuperscript{b}Centre Mémoire, CHU, Université Lille Nord de France, Lille, France

Key Words
Alzheimer’s disease · Vision · Eye movements · Saccade · Scene perception · Aging

Abstract
Aims: We investigated the performance in scene categorization of patients with Alzheimer’s disease (AD) using a saccadic choice task. Method: 24 patients with mild AD, 28 age-matched controls and 26 young people participated in the study. The participants were presented pairs of coloured photographs and were asked to make a saccadic eye movement to the picture corresponding to the target scene (natural vs. urban, indoor vs. outdoor). Results: The patients’ performance did not differ from chance for natural scenes. Differences between young and older controls and patients with AD were found in accuracy but not saccadic latency. Conclusions: The results are interpreted in terms of cerebral reorganization in the prefrontal and temporo-occipital cortex of patients with AD, but also in terms of impaired processing of visual global properties of scenes.

Introduction
Characterized by cognitive impairment, memory deficits and dementia, Alzheimer’s disease (AD) is a progressive neurodegenerative disease that evolves over many years \[1, 2\] and affects more than 36 million people in the world \[3\]. In addition to memory and attentional deficits, patients with AD also show alterations in sensory perception, particularly in visual processing \[4–7\]. Patients with AD exhibit deficits in only some types of visual processing, while other aspects are relatively spared \[8\]. For instance, they show deficits in colour discrimination \[9\], visual motion processing \[10\] and other visual field deficits \[4, 11, 12\]. Patients also exhibit reduced contrast sensitivity \[4, 6, 13, 14\] which affects more specif-
ically the middle and low spatial frequency ranges [15]. It has been suggested [16, 17] that such a loss could result from a selective and early deficit of the magnocellular pathway. Indeed, patients with AD experience degeneration of the visual pathway including the central areas of the magnocellular pathway in the lateral geniculate nucleus and in the visual cortex, and a loss of retinal ganglion cells within the retinal nerve fibre layer [18–21].

Neuropsychological studies on object perception in patients with AD have also reported deficits at intermediate levels (perceptual organization, figure/ground segregation [22, 23]) as well as in higher-order visual functions such as picture naming, face discrimination, spatial localization, visuomotor coordination and semantic representations [24]. Tippett et al. [25] designed a study to address some of the methodological issues affecting the investigation of semantic category effects in AD. They concluded that the main determinant of whether or not a categorical impairment of either sort is found in AD is which stimulus properties are controlled during stimulus selection. Nevertheless, recent studies on AD have often shown category-specific difficulties that typically include relative impairment for natural scenes compared to manufactured objects [26–30]. Since this deficit has been shown in studies using well-matched category-specific materials, the impairment cannot be reduced to non-semantic deficits such as limited attention or impaired mental imagery [26]. Consistent with impaired categorization performance in figure/ground discrimination, Boucart et al. [23] showed that patients with mild AD were significantly less accurate than age-matched observers in detecting an animal within a scene in a saccadic choice task. In that study, each trial consisted of the presentation of two lateral (left/right) photographs of natural scenes, and the participants were asked to move their eyes as quickly as possible to the scene containing an animal. The authors suggested that the lower accuracy observed in patients with AD may result from impairment in processing diagnostic coarse features that would differentiate an animal from another object in peripheral presentation. Few studies have investigated scene perception in AD. Usually, studies in which scenes were used as stimuli involved detection or search of an object within a scene [23, 31]. Lee et al. [32] reported impaired performance for photographs of outdoor scenes in patients with AD, but they used a complex discrimination task of superimposed images (morphs).

In the present study, we were interested in scene gist recognition, i.e. the ability of patients with AD to detect diagnostic information to categorize a scene. To our knowledge, no study has explored the semantic category-specific effect of visual scenes in patients with AD. Indeed, so far category-specific impairments had only been assessed with words or pictures in naming tasks [25, 33, 34]. We compared the performance of two spatial properties corresponding to different levels of scene analysis: categorization based on naturalness (natural vs. urban scenes) and categorization requiring the processing of more local information (to discriminate indoor from outdoor scenes). Studies on healthy adults have shown that the visual system is particularly sensitive to the global structure or overall spatial layout of a scene [35, 36]. Although these two properties (naturalness and indoor/outdoor) are considered as holistic or global (i.e. the categorization can be based on the overall layout), studies with young normally sighted observers [35, 37] have reported longer categorization times (around 470 ms) for indoor versus outdoor scenes than for naturalness (around 390 ms), likely because a more local object analysis is required to discriminate between indoor and outdoor scenes, whereas a coarse description based on orientation and colour is sufficient to decide if a scene is natural or urban [38, 39]. Consistent with this account, Boucart et al. [40] found that categorization of naturalness resisted the low resolution of large eccentricity (above 50°), whilst performance was at chance for the categorization of indoor versus outdoor scenes at the same eccentricities.

Based on these previous results and studies supporting impairments in visual recognition of coarse information [15–17, 27, 41] associated with a category-specific semantic deficit in
the recognition of natural/living objects observed in patients with AD [26, 28], we expected the patients to be more impaired in broad-level categorization (naturalness) than in categorization involving the processing of local information in scenes. A second objective of the study was to assess whether the category-specific deficit in the recognition of natural/living objects reported in patients with AD [25, 26] extends to natural versus man-made scenes.

Method

Participants

A total of 24 patients with probable AD at a mild stage (MMSE score >22) according to the NIA/AA criteria [42] were recruited in the Memory Clinic at the University Hospital of Lille, France. All the patients fulfilled the research criteria of the International Working Group [43] after a comprehensive workshop including neuropsychological assessment: MRI (n = 11), CSF biomarkers (n = 4), SPECT (n = 5) or FDG-TEP (n = 4). The control groups were 28 healthy elderly (relatives or friends of the patients) and 26 healthy young subjects (students of medicine or neuroscience). Young subjects were included as controls to dissociate the effect of aging from the effect of pathology as several studies have reported that oculomotor behaviour changes with normal aging [44–47] including intentional suppression of attentional capture by distracters [47, 48]. The patients were excluded if they had evidence of vascular lesions, major depression or ophthalmologic impairments (cataract, macular degeneration or glaucoma). Patients taking cholinesterase inhibitors were included if the dose had remained unchanged for at least 6 weeks prior to inclusion. Ophthalmologic screening included a detailed history of current or past visual disturbances, the assessment of visual acuity, the Amsler grid (for macular degeneration), the Pelli-Robson Contrast Sensitivity Chart, the visual field and items of the Ishihara colour vision test. Clinical and demographic data are presented in Table 1. The study was approved by the Ethics Committee of Lille (CPP N°05/79 Nord-Ouest IV). In accordance with the tenets of the Declaration of Helsinki, written informed consent was obtained from all participants.

Table 1. Demographic and clinical data of the study population

<table>
<thead>
<tr>
<th></th>
<th>Young controls</th>
<th>Old controls</th>
<th>Patients with AD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>26</td>
<td>28</td>
<td>24</td>
</tr>
<tr>
<td>Age, years</td>
<td>26.7 ± 2.3</td>
<td>69.1 ± 7.1</td>
<td>71.4 ± 5.8</td>
</tr>
<tr>
<td>MMSE score</td>
<td>29.1 ± 0.6</td>
<td>23.8 ± 1.1</td>
<td>128.1 ± 9.1</td>
</tr>
<tr>
<td>Bell test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dementia rating scale (/144)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education, years</td>
<td>12.9</td>
<td>12.1</td>
<td>11.9</td>
</tr>
<tr>
<td>Acuity (Armaignac acuity scale)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contrast perception</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Pelli-Robson)</td>
<td>2.1 normal</td>
<td>1.8 normal</td>
<td>1.6 normal</td>
</tr>
<tr>
<td>Ishihara colour vision test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laterality (R/L)</td>
<td>24R/2L</td>
<td>27R/1L</td>
<td>23R/1L</td>
</tr>
<tr>
<td>Gender (F/M)</td>
<td>13F/13M</td>
<td>18F/10M</td>
<td>14F/10M</td>
</tr>
</tbody>
</table>
of 120 Hz, and compensates for head movements. The manufacturers report a gaze position accuracy of <0.1°. Images of the eyes are analysed in real time by detecting the pupil, calculating the centre and eliminating artefacts. The same calibration was used for patients and controls. Eye movement data were collected binocularly. Analysis of the recorded eye movement data was conducted with the software BeGaze (Senso-Motoric Instruments).

The experiment was generated with a computer (Windows 7 professional 32-bit SP1 operating system) with a Intel Core i7-3540M (3.00 GHz) processor and 4 GB RAM, and presented on a 17'' screen (Dell E172FPT) using an NVIDIA NVS200M graphic (1,024 MB GDDR5) card (field rate of 60 Hz).

**Stimuli**

The stimuli were 64 photographs of scenes provided by A. Oliva (MIT). Two scene properties were selected: ‘naturalness’ in one block (natural vs. urban scenes) and indoor/outdoor scenes (man-made constructions) in another block. Natural scenes included no man-made objects (e.g. mountains, rivers, fields and forests). Urban scenes were mainly composed of skyscrapers. Examples are shown in figure 1. The angular size of the photographs was 13° × 13° at a fixed viewing distance of 61 cm. For each trial, two pictures were simultaneously presented, one on the left and the other on the right of the screen, with an eccentricity of 8° from the centre of the screen to the centre of each picture. The pictures were displayed on a black background.

**Procedure**

Participants were seated in a dimly illuminated room at a viewing distance of 61 cm in front of the computer screen. The eye tracker was first calibrated for each subject and then at the beginning of each block. The calibration was a pursuit of a red dot with 5 fixed points performed with the software Experiment Center 3 (SMI). The software accepts the 5 fixation points automatically. Using these reference points, the system creates a mapping function that relates all eye positions to points in the calibration area. The recording trials were initiated only if the eye tracker classified the calibration as ‘good’. Following the calibration, the participants performed the saccadic choice task. Neither speed nor accuracy was emphasized. The participants were advised to ask for breaks during the experiment whenever necessary. Each trial started with a central white fixation cross displayed for 500 ms on a black background. This was followed, after a gap of 200 ms, by a pair of photographs of scenes displayed for 1 s. The centre of each lateral picture was located 8° from fixation. In the block ‘naturalness’ (pictures composed of natural and urban scenes only), the participants were asked to make a saccadic eye movement to the picture containing a natural or urban scene (counterbalanced across participants). Each participant took part in only one of the conditions (target ‘natural scene’ or target ‘urban scene’). Each trial was composed of one natural and one urban scene, and a natural scene was always compared to the same urban scene and reciprocally. In another block (including indoor and outdoor scenes), participants were asked to make a saccadic eye movement to the picture containing an indoor or an outdoor scene (counterbalanced across participants). Each participant took part in only one of the conditions (indoor or outdoor scene). Each trial was composed of one indoor and one outdoor scene, and an indoor scene was always compared to the same outdoor scene and reciprocally. The left/right location of the target occurred randomly and equally. The trial order was selected randomly. One comparison between two scenes was always with the same pictures. Ten practice trials with scenes not used in the experimental set preceded the experiment which included 64 trials (32 with the natural vs. urban task and 32 with the outdoor vs. indoor task). The block order was randomized. The session lasted approximately 15 min.
Results

Analyses of variance using the software Statistica (7.1; Tulsa, Okla., USA) were conducted on the latency of the first saccade toward the targets and on accuracy. Participants whose global performance differed by 2 SD values away from the mean were discarded. Two patients with AD, 1 healthy elderly and 1 young were excluded. Responses shorter than 100 ms were also excluded, as were trials with incorrect responses. The factors were the group of participants (young and elderly controls, and AD patients). Analyses were conducted separately for

Fig. 1. Examples of pairs of scenes used in the experiment. a Pictures in the indoor/outdoor task. b Pictures in the natural/urban task.
natural/urban and for indoor/outdoor categorization tasks. Accuracy is presented in figure 2 and latencies of the first saccade in figure 3. No significant main effect of the target’s location (left/right) was observed, neither for accuracy nor for duration, in the 3 groups of participants.

**Natural versus Urban Categorization Task**

A main effect of the task was observed on accuracy \([F(1, 69) = 9.6, p < 0.01]\). Performance was better when the target was an urban scene (68.5%) than when the target was a natural scene (58.8%). There was no main effect of the group on accuracy \([F(2, 69) = 1.9, p = 0.15]\; young controls: 67.9%; elderly controls: 62.2%; AD patients: 60.8%). The interaction between group and condition was significant for accuracy \([F(2, 69) = 3.2, p < 0.05]\). Statistical planned contrast comparisons were conducted to clarify the effects. The patients with AD were less accurate when the target was a natural scene (50.2%) than when the target was an urban scene \([71.3%; F(1, 21) = 13.2, p < 0.001]\), whereas no significant difference was observed between natural and urban scenes for age-matched controls \([urban: 65.2\% vs. natural: 59.3\%]\; young controls: 69.4%; elderly controls: 63.8%; AD patients: 58.8%).

**Fig. 2.** Percentage of correct responses (accuracy with standard errors) for the 3 groups (young controls, old controls, and patients with AD) as a function of the categorization task (natural/urban and indoor/outdoor scenes).

**Fig. 3.** Mean latency of the first saccade (with standard errors) for the 3 groups (young controls, old controls, and patients with AD) as a function of the categorization task (natural/urban and indoor/outdoor scenes).
F(1, 27) = 1.6, p = 0.23] and for young participants [urban: 69.1% vs. natural: 66.9%; F(1, 24) = 0.2, p = 0.60]. Categorization of natural scenes did not significantly differ from chance in AD patients [(t(20) = 0.22, p = 0.96]. Young participants categorized natural scenes with higher accuracy (>16%) than the AD group [F(1,21) = 10.2, p < 0.01]. There was no significant difference between young and older participants [F(1, 24) = 1.2, p = 0.15] or between AD patients and age-matched controls [trend effect F(1, 23) = 1.5, p = 0.07]. No significant difference between the 3 groups was observed for urban scenes. Saccade latencies did not differ significantly between the 3 groups [young controls (242.1 ms; SD: 15 ms), elderly controls (237.2 ms; SD: 17 ms) and patients with AD (237.8 ms; SD: 19 ms)]. There was no difference between saccade latencies for natural pictures (243.9 ms; SD: 21 ms) and urban pictures (234.2 ms; SD: 19 ms). Moreover, there was no significant interaction between group and task on saccade latencies [F(2, 69) = 0.3, p = 0.74].

**Indoor versus Outdoor Categorization Task**

Neither age nor pathology significantly affected accuracy [F(2, 70) = 0.3, p = 0.76; young controls: 58.0%; elderly controls: 59.3%; patients with AD: 56.8%] and saccade latencies [F(2, 70) = 1.1, p = 0.3; young controls: 240.8 ms, SD: 21; elderly controls: 241.9 ms, SD: 18; patients with AD: 231.1 ms, SD: 19] in this task. There was no significant difference between the scenes [indoor: 58.6% vs. outdoor: 57.5%; F(1, 70) = 0.1, p = 0.72 for accuracy; indoor: 237.9 ms, SD: 22; outdoor: 238.6 ms, SD: 19; F(1, 70) = 0.04, p = 0.84 for saccade latencies]. There was no significant interaction between group and task on accuracy [F(2, 70) = 0.6, p = 0.54] and saccade latencies [F(2, 70) = 0.04, p = 0.95]. Results are shown in figure 3.

**Comparison of Natural/Urban versus Indoor/Outdoor**

On average, accuracy was higher for the natural/urban categorization than for the indoor/outdoor categorization [63.7 vs. 58.1%; F(1, 145) = 6.6, p < 0.05]. This was mainly due to young participants who were better in the natural/urban than in the indoor/outdoor categorization task [by 10.3%; F(1, 49) = 6.2, p < 0.01]. There was no significant difference between the natural/urban and the indoor/outdoor categorization task in the 2 other groups (elderly controls: −2.9% for indoor/outdoor pictures compared to natural/urban pictures; patients with AD: −3.3% of accuracy for indoor/outdoor pictures compared to natural/urban pictures). There was no effect on saccade latencies.

**Discussion**

The experiment reported here was designed to investigate scene categorization in AD. To our knowledge, few studies have explored scene perception in patients with AD [23, 31, 32, 49] and particularly scene gist recognition. In the present study, the participants were asked to make a saccadic eye movement as fast as possible to the side of the target within pairs of scenes. Consistent with the literature on eye movements in AD [23, 50–55], we observed an increased number of saccades in the wrong direction (toward the distracter). This decrease in accuracy affected particularly the categorization of natural scenes for which patients were at chance level. This result contrasts with data on healthy young observers showing, for instance, that categorization of natural versus urban resists the low spatial resolution of very large eccentricities (above 50°), whilst categorization of indoor versus outdoor is at chance at large eccentricities, likely resulting from the need to recognize local features (e.g., pieces of furniture) to decide if a scene is indoor or outdoor [40]. Moreover, Greene and Oliva [37] found that image exposure thresholds were shortest for classifying images according to whether or not they were natural.
Why are patients with AD impaired at rapid categorization of natural scenes? Previous studies using objects as stimuli have shown that a category-specific deficit in AD is characterized by relative impairment in the recognition of natural compared to man-made objects [26, 28, 56, 57]. Our results suggest that this deficit generalizes to scene categorization. Re-organization and neural change in AD can in part explain this result. Indeed, in a categorization task using fMRI [26], participants (young adults and patients with AD) were asked to judge whether pairs of words shared a feature. Data showed activation of the temporoooccipital cortex (TOC) where visuoperceptual features may be represented as well as activations of the prefrontal cortex (PFC), which may contribute to feature selection. Once again patients were impaired in the recognition of natural objects compared to man-made objects. The dorsal white matter in AD projection between the PFC and TOC correlated only with judgments of natural scenes. For man-made objects, patients activated regions outside of the PFC-TOC network that showed limited correlation areas activated by controls, but may help patients for man-made categorization. The authors concluded that patients remained dependent on the same neural network as controls, specifically during judgments of natural scenes, despite disease in these areas. The deficit in this specific category is due in part to this neural network dependence. However, the targets used in the experiment by Grossman et al. [26] were words; brain activation associated with words cannot totally be matched to brain activation associated with pictures of objects or photographs of scenes.

Another hypothesis that could explain the patients’ impairment in our visual task may be that rapid categorization is based on the detection of global physical properties in the scenes [58, 59]. In the present study, urban scenes were structurally homogeneous (skyscrapers and cities), whilst natural scenes were composed of a variety of structurally different scenes (mountains, forests, beaches, rivers, deserts, fields, etc.). This variety of different scenes might have confused the patients and led them to make a saccadic eye movement more frequently towards urban scenes in case of uncertainty. The effect of structural homogeneity on visual categorization has already been observed in previous studies [60, 61]. The high structural homogeneity of the urban scenes compared to the natural scenes could facilitate their categorization. Crouzet et al. [59] explained that the large advantage of the animal category over the vehicle category was difficult to explain using a simple search strategy based on a simple set of features. The authors suggested that one possibility was that processing of certain key biological stimuli such as animals (and faces) could involve faster hard-wired neural mechanisms possibly tuned by expertise or even by ancestral priorities. However, the structural similarity between faces [60] was higher than the structural homogeneity between vehicles. In our study, the structural homogeneity is beneficial to urban pictures (man-made).

The comparison between performance in the natural/urban task and the indoor/outdoor task showed that accuracy was higher for natural/urban (63.3%) pictures than for indoor/outdoor (58.5%) pictures in young participants only. By manipulating the eccentricity of naturalness and indoor/outdoor pictures (from an eccentricity of 10–70°), Boucart et al. [40] showed that the perception of some global properties (naturalness) was more robust in peripheral vision than the categorization of indoor/outdoor scenes, which requires a more local analysis. In line with this account, Tran et al. [62] found that patients with low vision were better at categorizing natural/urban than indoor/outdoor scenes. As the first saccade was taken into account in the present study, this fast processing benefited more a discrimination between natural versus urban scenes, which were structurally different, than a discrimination between indoor and outdoor scenes, which were structurally similar.

According to Bar’s [63] model, the extraction of sensory visual attributes follows a coarse-to-fine processing scheme, where low spatial frequency, i.e. a coarse spatial scale, represents the global information about the shape and orientation of the stimulus, while high spatial...
frequency, i.e. a fine spatial scale, corresponds to configural information and fine details. The coarse information of a picture is projected early and rapidly from the visual cortex to the PFC and the parahippocampal cortex, possibly through the magnocellular pathway, where it can activate a scene schema. This coarse representation activates, in the PFC, expectations about the most likely interpretations of the input image, which are then back-projected as an ‘initial guess’ to the temporal cortex to be integrated with the bottom-up analysis. The top-down process facilitates recognition by substantially limiting the number of object representations that need to be considered. However, the anterior areas (likely the PFC) do not accomplish the ‘initial guesses of the stimulus’ on the basis of a coarse information but need more physical information. A good candidate to mediate this rapid projection is thought to be the magnocellular pathway, which is known to convey low spatial frequency information [63]. Viggiano et al. [27] showed, with filtered pictures of living and man-made objects, that patients with AD were impaired in performance with regard to both categories compared to controls, but the decline was more rapid for living things and specifically with low-pass filtered pictures. The authors suggested that the magnocellular deficit in AD might result in poor performance in higher visual cognitive function such as categorization. Our results suggest that the reduced accuracy in the recognition of natural scenes in patients with AD, associated with difficulties in extracting the coarse information of a picture [27], is also compatible with a dysfunction of the magnocellular pathway in patients with AD [15, 41, 64]. Though the material and the experimental conditions (optimal contrast, long exposure time and colours) used in our experiment cannot completely isolate the magnocellular pathway, the results from the study by Viggiano et al. [27] with filtered images provide some clues about the repercussion of the magnocellular deficit on the ability to categorize natural pictures.

Except for accuracy, there was no difference in the oculomotor parameter (latency of the first saccade) between the young and elderly controls and the patients with AD. This result contrasts with recent studies which suggested that eye movement characteristics including fixation duration, saccade length, direction and re-fixation patterns can be used to distinguish patients with AD from control subjects [65, 66]. Ocular and visual information processing (from the retina through the visual pathway to the visual cortex) undergo changes in AD [50, 67]. Some aspects of vision loss are typically due to advanced age [68–70], but automatic oculomotor parameters are minimally influenced by normal aging [71] and specific oculomotor impairments appear in AD [52–55, 66, 67] such as increased latency, reduced accuracy and peak velocity, smooth pursuit abnormalities and visual fixation instability [50]. However, other studies consider that the overall increase in saccade reaction time is related to increased motor and sensory processing time in AD [51] and may reflect a global processing speed deficit. Neither pathology nor age affected the latency of the first saccade in our task. Our patients were at a mild stage of the disease, whereas most of the previous studies have included patients ranging from mild to severe stages of AD dementia [51, 66, 67]. Taking into account these results, the severity of the dementia might potentially play a role in the saccade variable. Boucart et al. [23] found no deficit in the oculomotor parameters (amplitude, latency and duration) in patients with a mild stage of AD during a categorization task.

The present results emphasize the importance of visual perception in AD, and information gathered from such assessments may be useful in object recognition, scene perception, the prediction of instrumental abilities and globally in activities of daily living.

Acknowledgements

The authors are grateful to all patients and controls for agreeing to participate. The study was funded by a grant ’ADvise’ from the French National Research Agency to M.B.
References


