Age, Executive Function, and Social Decision Making: A Dorsolateral Prefrontal Theory of Cognitive Aging

Sarah E. MacPherson, Louise H. Phillips, and Sergio Della Sala
University of Aberdeen

Current neuropsychological models propose that some age-related cognitive changes are due to frontal-lobe deterioration. However, these models have not considered the possible subdivision of the frontal lobes into the dorsolateral and ventromedial regions. This study assessed the age effects on 3 tasks of executive function and working memory, tasks dependent on dorsolateral prefrontal dysfunction; and 3 tasks of emotion and social decision making, tasks dependent on ventromedial prefrontal dysfunction. Age-related differences in performance were found on all tasks dependent on dorsolateral prefrontal dysfunction. In contrast, age-related differences were not found on the majority of the tasks dependent on ventromedial prefrontal dysfunction. The results support a specific dorsolateral prefrontal theory of cognitive changes with age, rather than a global decline in frontal-lobe function.

Healthy adult aging is associated with the deterioration of the frontal lobes of the brain, earlier and more severely than other brain areas (Haug et al., 1983). Current neuropsychological models propose that it is this frontal-lobe deterioration that is responsible for many age-related cognitive changes (Daigneault & Braun, 1993; Moscovitch & Winocur, 1995; West, 1996). Furthermore, studies of cognitive aging show that age effects are most evident on cognitive tasks (Mittenberg, Seidenburg, O’Leary, & DiGiulio, 1989) and memory measures (Moscovitch & Winocur, 1995) thought to be sensitive to frontal-lobe dysfunction, although age effects have been reported on measures sensitive to other regions of the brain (Greenwood, 2000).

Yet, although older individuals show poorer performance than younger individuals on typical frontal-lobe measures (Daigneault, Braun, & Whitaker, 1992; Mittenberg et al., 1989), many frontal patients are reported to perform relatively well on these measures and instead demonstrate comprehensive problems in everyday life, such as inappropriate social behavior and poor social decision making (Brazzelli, Colombo, Della Sala, & Spinnler, 1994; Shallice & Burgess, 1991). One way of accounting for this dissociation between the performance of older adults and patients with frontal-lobe damage is to consider the possible subdivisions of the frontal lobes of the brain. The frontal lobes are not a homogeneous cellular region and can be subdivided into distinct areas such as the dorsolateral (DL) and ventromedial (VM) prefrontal regions. Although these regions are considered in the present study, there are a number of ways to structurally and functionally divide the frontal lobes. The DL region is thought to include Brodmann’s Areas 9 and 46 (Pandya & Yeterian, 1996). However, it is less clear which of Brodmann’s areas should be included under the heading of the VM region. It becomes more apparent when the different cortico-cortical connections of the DL and VM regions are considered (Pandya & Yeterian, 1996; Rolls, 1996). Although these regions are closely interconnected with one another, they are also connected to distinct parts of the brain. In particular, the DL region receives projections from the primary sensory and motor regions, as well as the parietal cortex, while the VM region is closely connected to the limbic system (Adolphs, Tranel, Bechara, & Damasio, 1996; Rolls, 1996). Generally, the VM region is categorized as Brodmann’s Areas 10, 11, 12, 13, 14, and 47, as these regions have reciprocal connections with the limbic system (Eslinger, 1999; Price, 1999).

An alternative way of differentiating the DL and VM regions is in terms of the functional domains that each supports. The DL area is thought to be important for cognitive abilities known as “executive functions” and working memory (Petrides & Milner, 1982), whereas the VM area is thought to be involved in the processing of emotions and the regulation of social behavior (Rolls, 1996). In support of this dissociation, Sarazin et al. (1998) used PET to examine patients with frontal-lobe lesions and found that the cerebral glucose metabolism in the DL regions, but not VM regions, correlated with performance on executive abilities whereas the metabolism in the VM regions, but not DL regions, correlated with behavioral and emotional abnormalities.

The current neuropsychological models of aging have failed to consider fully these possible subdivisions of the frontal lobes and have led to the suggestion that there is general age-related frontal-lobe decline (Daigneault & Braun, 1993; Moscovitch & Winocur, 1995). Nevertheless, West (1996, 2000) suggested that there may be differential effects of age on different regions of the frontal lobes. Also, there are a handful of morphological and functional studies that have distinguished between the DL and VM regions in terms of aging. Autopsy studies have shown that there is a significant decrease in brain weight, cortical thickness, and number of large neurons in the DL cortex (Terry, DeTeresa, & Hansen, 1987). In contrast, the neurons in the VM cortex do not tend to...
shrink until much later, and there is a smaller slope in the function of cell loss compared with other brain regions (Haug et al., 1983).

Few neuroimaging studies have distinguished between age effects on the DL and VM areas; however, those studies that do have reported that the glucose metabolism at rest in the DL region is significantly correlated with age while the glucose metabolism at rest in the VM region is not (Duara et al., 1983; Garraux et al., 1999). Nonetheless, one recent structural MRI study has reported age-related differences in the size of both the DL and VM regions (Raz et al., 1997). However, in this study the more ventral and medial Areas 12, 13, and 14 that are essential for influencing social behavior (Rolls, 1996) were not included under the heading of the VM region.

If adult aging is associated with accelerated deterioration of the DL compared with the VM regions, then age effects should begin early and progress rapidly on frontal-lobe tests involving executive function or working memory because these tasks are sensitive to DL prefrontal dysfunction. In contrast, age effects should be less apparent on tasks involving emotional processing or social decision making because these tasks are sensitive to VM prefrontal dysfunction. To our knowledge, there are not any studies in the literature that have compared the effects of healthy adult aging on tests sensitive to DL and VM prefrontal dysfunction. Therefore, the current study attempts to examine the effects of adult aging on tasks sensitive to damage to each of these frontal-lobe regions.

To identify tests essentially sensitive to DL or VM prefrontal dysfunction, we conducted an extensive review of the frontal-lobe tests available in the literature. Only those tests with evidence from patient and neuroimaging studies supporting the involvement of either the DL or VM regions were considered. Although no tests exist that specifically tap only the DL or the VM regions, several frontal-lobe tests have been identified as involving one of these regions more than the other, on the basis of patient and neuroimaging data (for a review see Phillips, MacPherson, & Della Sala, 2002).

**DL Prefrontal Measures**

**Wisconsin Card Sorting Task**

The Wisconsin Card Sorting Task (WCST; Grant & Berg, 1948) was devised to assess the ability to identify abstract categories and shift cognitive set. Studies have shown that patients with lesions in the DL region make significantly more errors and achieve fewer sorting categories on the WCST than do patients with lesions elsewhere in the brain (Milner, 1963). In contrast, patients with lesions mainly involving the VM region perform similarly on the WCST to patients with lesions outside the frontal lobes (Ahola, Vilkki, & Servo, 1996; Milner, 1963) and better than patients with frontal damage outside this region (Drewes, 1974). Further evidence from neuroimaging studies shows significant increases in blood flow to the DL but not to the VM region during WCST performance (Rezai et al., 1993).

**Self-Ordered Pointing Task**

The Self-Ordered Pointing Task (SOPT; Petrides & Milner, 1982) was developed to assess the ability to arrange, perform, and monitor a sequence of responses. The few studies that have examined the role of the DL area on SOPT performance have found that patients with DL lesions perform significantly more poorly than patients with temporal-lobe lesions (Petrides & Milner, 1982) and healthy controls (De Zubicaray, Chalk, Rose, Semple, & Smith, 1997). Further support from neuroimaging studies provides evidence of significant increases in rCBF in the DL region in healthy volunteers while performing the SOPT task (Petrides, Alivisatos, Evans, & Meyer, 1993; Petrides, Alivisatos, Meyer, & Evans, 1993).

**Delayed-Response Task**

In the delayed-response task (Hunter, 1913), participants are required to maintain an internal representation during a brief delay. A small number of patient studies have examined the effects of DL lesions on the delayed-response task. These studies have demonstrated that patients with DL lesions perform poorly on the delayed-response task compared with patients with temporal-lobe lesions (Teixeira Ferreira et al., 1998), patients with postcentral lesions, and healthy controls (Vérin et al., 1993). Furthermore, the evidence from neuroimaging studies suggests that there is significant DL but not VM activation during the performance of the task (Goldberg, Berman, Randolph, Gold, & Weinberger, 1996).

In terms of the effects of aging on these DL tasks, age-related differences in performance have been found on both the WCST (Daigneault et al., 1992) and the SOPT (Daigneault et al., 1992). However, there do not appear to be any studies in the literature that have directly examined the effects of adult aging on delayed-response tasks similar to those used in patient populations, although there are age effects reported in nonhuman primate studies (Bartus, Fleming, & Johnson, 1978).

**VM Prefrontal Measures**

**Gambling Task**

Bechara, Damasio, Damasio, and Anderson (1994) devised the gambling task in an attempt to detect impairments in social decision making by simulating real-life decision making. The aim of the task is to try to win as much money as possible by making advantageous card selections rather than disadvantageous card selections from any one of four decks of cards. Several studies have found that patients with VM lesions perform the gambling task more poorly than healthy controls, or patients with lesions in the occipital lobes, temporal lobes, or DL prefrontal regions (Bechara et al., 1994; Bechara, Damasio, Tranel, & Anderson, 1998). There are no neuroimaging studies in the literature that have examined healthy volunteers performing the gambling task; nonetheless, activation in the VM region has been demonstrated in decision-making-type tasks with some overlapping cognitive demands to the gambling task (Rogers et al., 1999).

**Faux Pas Task**

Although there is some agreement that the frontal lobes are involved in theory of mind abilities (Baron-Cohen & Ring, 1994; Stuss, Gallup, & Alexander, 2001), there is only one study that we are aware of that has compared the effects of DL and VM lesions...
on theory of mind tasks. Using the faux pas task, Stone, Baron-Cohen, and Knight (1998) demonstrated that patients with VM lesions were impaired at detecting faux pas or social slips in stories compared with patients with DL lesions and healthy controls. Evidence from the neuroimaging studies indicates that there is VM activation when healthy volunteers perform mental state recognition tasks (Baron-Cohen et al., 1994) or view theory of mind stories and cartoons (Gallagher et al., 2000). However, some neuroimaging studies have also demonstrated more dorsomedial prefrontal activation when healthy volunteers attempt to attribute mental states to others (Fletcher et al., 1995) or attribute intentions to characters in cartoons (Brunet, Sarfati, Hardy-Bayle, & Decety, 2000).

**Emotion Identification Task**

Patients with lesions in the VM region have been found to be significantly impaired at identifying emotions compared with patients with lesions elsewhere in the brain (Hornak, Rolls, & Wade, 1996). Neuroimaging studies have implicated the VM region in emotion identification, although some studies have also found activation in the DL areas (George et al., 1993). There is also evidence that the VM region may be differentially activated, depending on which emotion is observed, in particular when processing angry expressions rather than sad expressions (Blair, Morris, Frith, Perrett, & Dolan, 1999).

There has been relatively little research examining the effects of age on tasks sensitive to VM prefrontal dysfunction. The few studies that have been conducted suggest that abilities such as attributing mental states to others (Happed, Winner, & Brownell, 1998) and perceiving emotions (Blanchard-Fields, Jahnke, & Winner, 2000) are well preserved in old age, although there are also studies that have reported age effects on emotion perception (Malatesta, Izard, Culver, & Nicolic, 1987). As yet, there does not appear to be any work published on the effects of age on the gambling task. However, age differences have not been reported on decision-making tasks, such as investing money (Walsh & Hershey, 1993) or choosing an insurance policy (Hartley, 1989).

**Involvement of Medial Temporal Lobes**

It is important to point out that in addition to deterioration of the frontal lobes, deterioration in regions of the medial temporal lobes has also been associated with healthy adult aging (Golomb et al., 1993). Moscovitch and Winocur (1992) and Mittenberg et al. (1989) claim that both of these regions are responsible for the decline in memory and executive functions associated with age. Therefore, in an effort to demonstrate that any age-related cognitive differences are due to frontal and not medial temporal-lobe deterioration, it may be useful to examine the effects of healthy adult aging on measures thought to have some temporal-lobe involvement and adjust participants’ scores on the frontal tests accordingly for this memory component.

The temporal-lobe measures included in this study were the Verbal Paired Associates subtest of the Wechsler Memory Scale–III United Kingdom version (WMS-III UK; Wechsler, 1997), which is thought to be sensitive to left temporal-lobe dysfunction (Lezak, 1995), and the Doors subtest of the Doors and People Test (Baddeley, Emslie, & Nimmo-Smith, 1994), which is thought to be sensitive to right temporal-lobe dysfunction (Morris, Abrahams, Baddeley, & Polkey, 1995).

**Method**

**Participants**

Three age groups involving 30 healthy participants in each (15 men, 15 women) were compared in this study: 30 aged between 20 and 38 years ($M = 28.8$, $SD = 6.0$), 30 aged between 40 and 59 years ($M = 50.3$, $SD = 5.7$), and 30 aged between 61 and 80 years ($M = 69.9$, $SD = 5.5$). All participants were recruited through a subject panel at the Psychology Department, University of Aberdeen, and they were reimbursed for their time. None of the participants had any history of the neurological or psychiatric disorders listed in the Wechsler Adult Intelligence Scale–III UK (WAIS-III UK) and WMS-III UK selection criterion (Wechsler, 1997).

The older group had significantly fewer years of education than did the middle-aged and young groups with means of 12.4 years ($SD = 3.2$), 14.0 years ($SD = 3.2$), and 15.0 years ($SD = 2.4$), respectively, $F(2, 87) = 5.91$, $MSE = 8.66$, $p < .005$, $\eta^2_p = .12$. Analyses indicated that education did not explain any of the age differences, therefore education is not considered further. All participants were administered the Vocabulary subtest from the WAIS-III (Wechsler, 1997). The young, middle-aged, and older groups did not significantly differ in terms of their vocabulary scores with means of 44.8 ($SD = 7.4$), 47.1 ($SD = 10.7$), and 47.6 ($SD = 8.2$), respectively, $F(2, 87) = 0.86$, $MSE = 78.73$, $\eta^2_p = .02$.

**Procedure**

Participants performed eight tasks; the order of these tasks was counterbalanced across participants.

**Tests of DL Prefrontal Dysfunction**

**WCST.** The WCST was administered and scored by using the standardized procedure published in the WCST manual (Heaton, 1981). The task involved participants matching cards through trial and error to stimulus cards on the basis of color, shape, or number. Only one of the categories was correct at a time: color in the first instance, then shape, and then number. Participants were not told what the current sorting category was and were simply informed whether they had correctly placed the response card beneath a stimulus card. After every 10 correct consecutive responses, the sorting category was changed without the participants’ knowledge. Therefore, participants had to shift their cognitive set to identify and attend to the new sorting dimension. The procedure was repeated twice or until all 128 response cards were placed. Each participant’s performance was scored in terms of the total number of errors, the number of perseverative responses, the number of perseverative errors, the number of nonperseverative errors, the number of conceptual level responses, the number of categories completed, and the number of failures to maintain set. In addition, a global score was included in the analysis: a measure thought to provide an overall index of WCST performance by taking into account the number of trials administered, the total number of errors, the number of conceptual level responses, and the number of categories completed (Laiacona, Inzaghi, De Tanti, & Capitani, 2000). The higher the global score, the poorer the performance of the WCST.

**SOPT.** Participants were presented with three computerized versions of the SOPT that differed only in the type of stimulus materials used: abstract designs, high-imagery words, and low-imagery words. The stimuli used in the abstract designs version of the SOPT were black and white computer-designed pictures that did not resemble concrete objects. The words used in the verbal versions of the task were taken from norms...
measuring the image evoking properties of words (Gilhooly & Logie, 1980). The high-imagery words were rated above 6.0, and the low-imagery words were rated below 3.2 on a 7-point scale. On each version of the task, participants were repeatedly presented with an array and instructed to choose one of the items in the array by using a light pen, on the condition that they did not choose an item that had already been chosen. Therefore, participants had to monitor their previous choices while they prepared each new response. The task became increasingly more difficult with 3 trials containing 8 items, then 3 trials containing 12 items, and finally 3 trials containing 16 items. As a result, participants had to remember whether responses had already been made within a trial or a block of trials. The set size determined the number of selections that participants were required to make in each trial (e.g., 12 selections in trials with 12 items). The 8 items were arranged in a 4 × 2 matrix, the 12 items were arranged in a 4 × 3 matrix, and the 16 items were arranged in a 4 × 4 matrix. Each set size contained different items, and the position of the items altered across presentations to prevent participants from simply moving in a spatially sequential order to perform each trial. The abstract designs version of the task was always administered prior to the high- and low-imagery versions of the task to discourage the use of verbal strategies on the abstract designs task. However, the presentation order of the high- and low-imagery versions of the SOFT was counterbalanced across participants. Participants were instructed that accuracy was more important than speed. They were not allowed to continually choose items from the same location, which would simply require them to remember whether an item had previously appeared in that location rather than prepare a sequence of responses. Furthermore, in the verbal tasks, participants were not allowed to pick the stimuli using any alphabetical strategies, such as using the first or last letter of each word. Only if the participant attempted either of these strategies did the experimenter indicate to the participant that such strategies should not be used, and once instructed not to do so, none of the participants continued to use either strategy. The computer recorded the order in which the stimuli were chosen, and every time an item was chosen more than once, it was deemed an error.

Delayed-response task. Participants were presented with 12 identical 2-cm blue squares displayed in a pseudorandom order on a 17-in. computer screen. Then a number of these squares, varying between 2 and 5, changed color from blue to white simultaneously. The number of squares that changed color determined how long the squares remained white on the screen before completely disappearing. In trials where 2 squares changed color, the squares remained on the screen for 2 s, and half a second was added for each additional square. Therefore, the squares remained white on the screen for a maximum of 3.5 s. Then, a delay followed where the screen remained blank for 0.5 s, 30 s, or 60 s; and during this delay interval, participants had to remember the spatial location of the squares that changed color. At the end of the delay, all 12 blue squares reappeared, and participants had to indicate which squares had changed color, using a light pen. Participants were able to change their mind as often as they wanted until they pressed an "accept" button. There were 6 trials for each set size (2, 3, 4, and 5 squares), resulting in 24 trials in total. Each set size contained 2 trials with a 0.5-s delay, 2 trials with a 30-s delay, and 2 trials with a 60-s delay. The different set sizes and delays were intermixed. Trials were scored as incorrect if participants added or excluded any squares when identifying the squares that changed color. Prior to performing the task, participants were given 6 shorter practice trials.

Tests of VM Prefrontal Dysfunction

Gambling task. The gambling task was based on that of Bechara et al. (1994) with the exception that the monetary values were converted into Sterling. Participants were presented with four decks of cards labeled A, B, C, and D and a loan of £2,000. The aim of the task was to try to maximize the profit on the £2,000 loan by choosing one card at a time from any of the four decks. Decks A and B were considered high-risk decks because they had both high rewards and high penalties, whereas Decks C and D were considered low-risk decks because they had both low rewards and low penalties. Consequently, it was more profitable in the long run to pick cards from the low-risk decks, C and D. Participants were allowed to switch from deck to deck as often as they desired; however, they were not told how many card selections they would have to make (100 card selections in total). Participants had no way of knowing when a penalty would appear and were unable to calculate the net gain or loss from each deck. To perform well on the task, participants had to develop a hunch that the high-paying decks were bad and the low-paying decks were good. The 100 card selections for each participant were split into 5 blocks of 20 cards. Then the number of cards chosen from the disadvantageous decks was subtracted from the number of cards chosen from the advantageous decks to derive a net score for each of the blocks. A negative score indicated that participants were choosing the cards disadvantageously, whereas a positive score indicated that participants were choosing the cards advantageously.

Faux pas task. In the faux pas task, participants were asked to silently read a series of short stories, 10 of which contained a faux pas or social slip (Stone et al., 1998) and 10 that did not. Each faux pas story contained a situation in which the protagonist in the story said something socially inappropriate that would hurt or insult another character in the story. There were no time limits, and participants were able to read the stories as many times as they required to fully understand them. Each story was followed by a series of questions about the intentions of the protagonist in the story to establish that the participant did in fact realize that a faux pas had been committed and a control question to check that the participant fully understood the story. Moreover, to understand that a faux pas had occurred, participants had to not only understand that the protagonist making the faux

<table>
<thead>
<tr>
<th>WCST score</th>
<th>Young M</th>
<th>SD</th>
<th>Middle-aged M</th>
<th>SD</th>
<th>Older M</th>
<th>SD</th>
<th>F(2, 84)</th>
<th>MSE</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total errors</td>
<td>20.9</td>
<td>12.5</td>
<td>29.8</td>
<td>18.1</td>
<td>35.9</td>
<td>16.4</td>
<td>3.44</td>
<td>247.42</td>
<td>&lt;.05</td>
<td>.08</td>
</tr>
<tr>
<td>Perseverative responses</td>
<td>12.8</td>
<td>9.2</td>
<td>15.0</td>
<td>8.3</td>
<td>22.5</td>
<td>14.8</td>
<td>4.05</td>
<td>123.64</td>
<td>&lt;.05</td>
<td>.09</td>
</tr>
<tr>
<td>Perseverative errors</td>
<td>11.3</td>
<td>7.2</td>
<td>13.8</td>
<td>7.1</td>
<td>20.0</td>
<td>11.7</td>
<td>5.02</td>
<td>79.67</td>
<td>&lt;.01</td>
<td>.11</td>
</tr>
<tr>
<td>Nonperseverative errors</td>
<td>9.5</td>
<td>6.2</td>
<td>15.9</td>
<td>13.4</td>
<td>16.0</td>
<td>9.3</td>
<td>1.95</td>
<td>98.58</td>
<td>.15</td>
<td>.04</td>
</tr>
<tr>
<td>Conceptual level responses</td>
<td>73.4</td>
<td>18.7</td>
<td>61.7</td>
<td>26.0</td>
<td>53.1</td>
<td>21.2</td>
<td>3.31</td>
<td>485.86</td>
<td>&lt;.05</td>
<td>.07</td>
</tr>
<tr>
<td>Global score</td>
<td>37.0</td>
<td>26.8</td>
<td>56.6</td>
<td>43.6</td>
<td>72.0</td>
<td>37.1</td>
<td>2.44</td>
<td>1505.68</td>
<td>&lt;.09</td>
<td>.06</td>
</tr>
<tr>
<td>No. of categories achieved</td>
<td>5.2</td>
<td>1.5</td>
<td>4.5</td>
<td>2.0</td>
<td>4.1</td>
<td>2.0</td>
<td>1.00</td>
<td>3.37</td>
<td>=.37</td>
<td>.02</td>
</tr>
<tr>
<td>Failure to maintain set</td>
<td>0.6</td>
<td>1.0</td>
<td>0.7</td>
<td>1.0</td>
<td>0.9</td>
<td>1.1</td>
<td>0.16</td>
<td>1.10</td>
<td>=.85</td>
<td>.00</td>
</tr>
</tbody>
</table>
pas did not know that they should not have said what they said but that they had some empathetic understanding of how the person in the story would feel. The stories remained in front of the participants while the questions were asked, allowing participants to check their answers. Answers were judged as correct if they indicated a clear understanding that the protagonist had unintentionally said something that would hurt or insult another character in the story. Any responses that inferred that the protagonist’s utterance was intentional or responses that involved incorrect facts about the story were judged as incorrect.

Emotion identification task. Participants were presented with color photographs of faces (Matsumoto & Ekman, 1988) in the middle of a computer screen, underneath which were a series of adjectives describing different emotions: happy, sad, angry, disgusted, frightened, surprised, and contempt. Participants were instructed to choose the adjective that best described the emotion displayed on the face in the photograph. Each photograph remained on the computer screen until the participants made their decision and the participants did not receive any feedback about whether their decision was correct or not. The photographs were presented in a pseudorandom order, and the adjectives always remained in the same position on the computer screen. No expression appeared more than once in succession, and each individual only appeared once in the series. Seven examples of each emotion were presented, resulting in a total of 49 faces. Prior to performing the task, participants were given 7 practice trials consisting of an example of each emotion. If participants were not confident about their response, they were encouraged to choose the adjective that they believed to be the closest to the viewed emotion.

Tests of Medial Temporal-Lobe Dysfunction

Verbal Paired Associates subtest. Participants were presented with 8 pairs of words at a rate of one word pair every 3 s. Four word pairs were semantically related (e.g., metal–iron) and 4 word pairs were not (e.g., cabbage–pen). Participants were then cued with the first word of each pair and asked to remember the word that was associated with it (e.g., metal–?). This procedure was repeated until participants were able to recall all of the word pairs (a minimum of 3 times and a maximum of 6 times). The first 3 times that the task was performed, participants were awarded 1 point for each correctly recalled word pair, resulting in a maximum score of 24 (Verbal Paired Associates I). Approximately 20 min later, participants were again cued with the first word of each pair and asked to remember the word that was associated with it. Once again, they were awarded 1 point for each correctly recalled word pair giving a score out of 8 (Verbal Paired Associates II).

Doors subtest. Participants were shown a series of 12 photographs of individual doors at 3-s intervals. They were then shown 12 pages with four different doors on each and asked to point to the door that they had previously been shown. Each correctly identified door was awarded 1 point, and a maximum of 12 points could be achieved. Then the procedure was repeated but with 12 different doors, giving participants a maximum possible score of 24.

Analysis

To determine that the age-related differences in performance on the tasks were at least partly due to frontal-lobe deterioration and not temporal-lobe deterioration, we derived a memory score for each participant. This memory score consisted of the sum of the standard $z$ scores for the Verbal Paired Associates subtests of the Wechsler Memory Scale–III (Wechsler, 1997) and the standard $z$ score for the Doors subtest of the Doors and People Test (Baddeley et al., 1994).

Multivariate analyses of covariance (MANCOVA) were performed on the WCST measures and the faux pas task measures with the memory score entered into the analyses as a covariate. The remaining measures were analyzed by using separate analyses of variance (ANOVA), and post hoc analyses were conducted by using Tukey’s HSD post hoc comparisons. Separate analyses of covariance (ANCOVA) were also reported to control for the effect of memory.

Results

Tests of DL Prefrontal Dysfunction

WCST. Two of the 90 participants (one middle-aged and one older) refused to complete the WCST, therefore the results were...
analyzed for the remaining 88 participants. A MANCOVA conducted on the WCST measures with memory entered as a covariate revealed a significant main effect of age. Wilks’s $\Lambda = .72$, $F(16, 154) = 1.72$, $p < .05$, $\eta^2_p = .15$. Univariate one-way analyses followed by post hoc $t$ tests revealed that the older group performed significantly more poorly than did the young and middle-aged groups in terms of the perseverative responses and perseverative errors, and the older group performed significantly more poorly than did the young group in terms of the total number of errors and conceptual level responses. Table 1 shows the means, standard deviations, and the separate univariate analyses for the WCST measures for the three age groups.

SOPT. Figure 1 demonstrates the performance of young, middle-aged, and older participants as a function of set size on the abstract designs, high-imagery, and low-imagery versions of the SOPT. A three-way ANOVA was conducted on the number of errors on the SOPT with age, stimulus type, and set size as factors. The analysis showed a significant main effect of age and significant interactions between age and stimulus type and age and set size. Table 2 shows the main effects, interactions, and post hoc analysis. An ANCOVA performed on the total number of errors on the SOPT with memory as a covariate showed that age-related memory differences did not account for the main effect of age on SOPT performance, $F(2, 86) = 9.56$, $MSE = 248.64$, $p < .0001$, $\eta^2_p = .18$.

Delayed-response task. The mean number of incorrect trials and standard errors of the mean for each age group as a function of number of squares and delay are shown in Figure 2. It should be noted that although the means for the 0.5-s delay were very close to zero, the Mauchly Sphericity Test (Mauchly, 1940) demonstrated that the assumption of homogeneity of variance was not violated and there was sufficient variance across the trials to permit the use of ANOVA. A three-way ANOVA compared the performance of the three age groups across the number of squares (2, 3, 4, and 5) and delay length (0.5 s, 30 s, and 60 s). The ANOVA revealed a significant main effect of age and a significant interaction between age and delay. Table 3 shows the main effects, interactions, and the pairwise comparisons. Moreover, an ANCOVA conducted on the total number of incorrect trials on the delayed-response task showed that when memory performance was covaried out of the analysis, the age effect remained, $F(2, 86) = 3.92$, $MSE = 9.58$, $p < .05$, $\eta^2_p = .08$.

Tests of VM Prefrontal Dysfunction

Gambling task. Figure 3 demonstrates the means and standard errors of the mean for the young, middle-aged, and older age groups performing the gambling task. A two-way ANOVA was conducted on the disadvantageous card choices subtracted from the advantageous card choices (net score) for each block of 20 cards for each age group. The analysis revealed a significant main effect of block, $F(4, 348) = 10.56$, $MSE = 55.80$, $p < .0001$, $\eta^2_p = .11$, and post hoc $t$ tests revealed that there was a significant difference between the net score for the first 1 to 20 cards and the net scores for Blocks 41–60, 61–80, and 81–100 and between the net score for Block 21–40 and the net score for Block 81–100. The main effect of age, $F(2, 87) = 0.73$, $MSE = 147.80$, $\eta^2_p = .02$, and the two-way Block $\times$ Age interaction, $F(8, 348) = 0.93$, $MSE = 55.80$, $\eta^2_p = .02$, were not significant.

Faux pas task. One of the older participants was excluded from the analysis as he claimed that a faux pas had been committed in all 20 stories. Therefore, the results of the remaining 89 participants were analyzed. In accordance with Stone et al. (1998), Table 4 demonstrates the mean performance of the three age groups on the stories containing a faux pas. Furthermore, to ensure that the three age groups were not only comparable in terms of their sensitivity to faux pas but their response bias, $A'$ (sensitivity) and $B'_{\text{D}}$ (bias) values were also calculated, using nonparametric signal detection analysis of the hit (correct responses to faux pas) and false-alarm (incorrect responses to no faux pas) rates. These values are also presented in Table 4. The performance of the three age groups on the faux pas stories, the control questions, the empathy questions, and the $A'$ and $B'_{\text{D}}$ values from the signal detection analysis were entered into a MANCOVA with memory entered as a covariate. Neither the MANCOVA, Wilks’s $\Lambda = .86$, $F(10, 162) = 1.26$, $\eta^2_p = .07$, nor the univariate analyses were signifi-

Table 2

<table>
<thead>
<tr>
<th>Effect</th>
<th>$df$</th>
<th>$F$</th>
<th>$p$</th>
<th>$MSE$</th>
<th>$\eta^2_p$</th>
<th>Tukey’s HSD$^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>2, 87</td>
<td>17.76</td>
<td>&lt;.0001</td>
<td>27.95</td>
<td>.29</td>
<td>older &gt; middle &amp; young, middle &gt; young</td>
</tr>
<tr>
<td>Stimulus type</td>
<td>2,174</td>
<td>60.63</td>
<td>&lt;.0001</td>
<td>4.10</td>
<td>.41</td>
<td>abstract &gt; low imagery &amp; high imagery, low imagery &gt; high imagery</td>
</tr>
<tr>
<td>Set size</td>
<td>2,174</td>
<td>322.86</td>
<td>&lt;.0001</td>
<td>4.67</td>
<td>.79</td>
<td>16 &gt; 12 &amp; 8, 12 &gt; 8</td>
</tr>
<tr>
<td>Age $\times$ Stimulus Type</td>
<td>4,174</td>
<td>3.08</td>
<td>&lt;.05</td>
<td>4.10</td>
<td>.07</td>
<td>abstract designs: older &gt; middle &amp; young, middle &gt; young; high imagery: older &gt; middle &amp; young, middle &gt; young; low imagery: older &amp; middle &gt; young</td>
</tr>
<tr>
<td>Age $\times$ Set Size</td>
<td>4,174</td>
<td>6.42</td>
<td>&lt;.0001</td>
<td>4.67</td>
<td>.13</td>
<td>8 items: older &amp; middle &gt; young; 12 items: older &amp; middle &gt; young; 16 items: older &gt; middle &amp; young, middle &gt; young</td>
</tr>
<tr>
<td>Stimulus Type $\times$ Set Size</td>
<td>4,348</td>
<td>3.87</td>
<td>&lt;.01</td>
<td>3.11</td>
<td>.04</td>
<td>8 items: abstract &gt; low imagery &amp; high imagery; 12 items: abstract &gt; low imagery &amp; high imagery, low imagery &gt; high imagery; 16 items: abstract &amp; low imagery &gt; high imagery</td>
</tr>
<tr>
<td>Age $\times$ Stimulus Type $\times$ Set Size</td>
<td>8,348</td>
<td>0.85</td>
<td>= .56</td>
<td>3.11</td>
<td>.02</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. The dash indicates that the effect or interaction was not significant, therefore it was unnecessary to conduct a Tukey’s honestly significant difference (HSD) post hoc analysis.

$^*$ Post hoc comparisons in terms of number of errors.
important, indicating that there were no significant age effects on any of the faux pas indices.

Emotion identification task. Figure 4 demonstrates the mean number of errors and the standard errors of the mean for the young, middle-aged, and older groups performing the emotion identification task. A two-way ANOVA with age and emotion as factors demonstrated a significant main effect of age, $F(2, 87) = 3.4, MSE = 4.40, p < .05, \eta^2_p = .07$. Post hoc comparisons demonstrated that older participants were significantly worse than younger participants at recognizing emotions. There was also a significant main effect of emotion, $F(6, 522) = 18.26, MSE = 1.97, p < .001, \eta^2_p = .17$, and post hoc $t$ tests showed that the emotion happiness was significantly easier to recognize and the emotion contempt was significantly harder to recognize than the other emotions. In addition, there was a significant two-way Age $\times$ Emotion interaction, $F(12, 522) = 2.82, MSE = 1.97, p < .001, \eta^2_p = .06$, and post hoc analysis revealed that the older group only performed significantly worse than the younger group when they were asked to identify the emotion sadness. A further ANCOVA was conducted on the total number of errors on the emotion identification task with memory entered as a covariate. The analysis revealed that the significant age effect in the ability to perceive emotions was removed when memory was partialed out of the analysis, $F(2, 86) = 0.76, MSE = 28.99, \eta^2_p = .02$. Therefore, it may be that the age effect is accounted for by medial temporal-lobe deterioration rather than frontal-lobe deterioration.

Summary variables were then selected for each of the DL and VM tests. For the DL tests, the variables were the total number of errors on the WCST, the sum of the errors on the abstract designs and the high-imagery and low-imagery word versions of the SOPT, and the total number of errors on the delayed-response task. For the VM tests, the variables were the total number of card selections from Decks A and B on the gambling task, the sum of the errors on the faux pas stories, the control questions and the empathy questions on the faux pas task, and the total number of errors on the emotion identification task. To understand how healthy adult aging relates to performance on these summary variables, we performed a canonical correlation on the data for the 87 participants who were included in the analyses for all the DL and VM measures. The first canonical correlation was .59 (35% shared variance) and was significant, Wilks’s $\Lambda = .60, F(12, 158) = 3.81, p < .0001$. The second canonical correlation was .26 (7% shared variance) but was not significant, Wilks’s $\Lambda = .93, F(5, 80) = 1.19$. As the data set contained only one significant canonical variate pair, this suggests that there is only one reliable way of combining performance on the DL and VM measures to relate to healthy adult aging. The canonical loadings revealed that the canonical variate (linear combination) for performance on the DL and VM measures was composed of performance on the WCST, SOPT, delayed-response task, and emotion identification task (see Table 5). Therefore, performance on all three DL measures and one of the VM measures was associated with healthy adult aging.

Discussion

The current study investigated the effects of age on tasks thought to be dependent on dysfunction of the DL and VM regions. Age effects were found on all three measures thought to be dependent on DL prefrontal dysfunction but were generally not found on the tasks thought to be dependent on VM prefrontal dysfunction. These findings suggest that age-related differences in frontal-lobe functions may show different trajectories, depending on the area within the frontal lobes that a task taps. There does not appear to be uniform frontal-lobe decline, as a general frontal-lobe hypothesis of aging would predict. Instead, the evidence suggests that DL prefrontal functions may be more sensitive to healthy adult aging than VM prefrontal functions.

The age-related differences in performance found on the DL tasks are in line with the previous findings in the literature demonstrating that executive function and working memory are subject
to the effects of healthy adult aging (Daigneault & Braun, 1993; Mittenberg et al., 1989; West, 1996). As these age effects remained even when performance on the temporal-lobe memory measures was controlled for, the age effects are unlikely to be due to deterioration of the medial temporal lobes as opposed to the frontal lobes. In contrast, emotional functioning and the regulation of social behavior appear to remain relatively intact with age. But until now, few studies have examined the effects of healthy adult aging on emotional processing and social decision making, using tasks analogous to those used in patient studies.

The claim that there are specific DL prefrontal changes with age relies on the assumption that performance on the measures used in this study essentially reflect either the functioning of the DL or VM prefrontal regions. By and large, the evidence from patient and neuroimaging data does point toward each test being largely sensitive to one of these regions more than the other. However, given the complexity of the tasks involved and the interconnections among these regions of the frontal lobes, there are probably no tests in the literature that exclusively tax only DL or VM functions. Some studies have suggested that DL lesions do not selectively affect WCST performance (Anderson, Damasio, Jones, & Tranel, 1991; Goldstein, Bernard, Fenwick, Burgess, & McNeil, 1993). There is also evidence that some patients with lesions involving the VM regions may be impaired on the delayed-response task (Bechara et al., 1998; Freedman & Oscar-Berman, 1986). Some neuroimaging studies have shown activation in both the DL and VM regions while healthy participants perform the WCST task (Berman et al., 1995) and a decision-making task similar to the gambling task (Elliot, Rees, & Dolan, 1999).

In terms of overall performance on the six tasks reported here, the methodological issue of task difficulty can also be raised with regard to the interpretation of these findings. It might be argued that the age differences on the DL tasks could result from a high overall task complexity rather than a specific age-related impairment. However, as Baddeley and Della Sala (1996) have pointed out, interpreting findings in terms of difficulty can be somewhat circular as difficulty is often attributed to age effects post hoc. Task difficulty per se cannot explain the age effects found on the DL tasks, as it does not explain the underlying processes involved in the tasks that make the tasks more difficult. In addition, there are many reports in the literature of patients with frontal-lobe damage who perform well within normal limits on tests of executive function and working memory but who show inappropriate social behavior and decision making (Brazzelli et al., 1994; Shallice & Burgess, 1991). In contrast, there are frontal-lobe patients who cope relatively well in everyday social situations but perform poorly on tests of executive function and working memory (Bechara et al., 1998).

Furthermore, although the DL and VM prefrontal regions are strongly interconnected with one another, they are also differentially connected with other brain regions. In a recent review, age effects were reported on tasks involving nonfrontal brain regions such as the parietal lobes, but which have strong reciprocal connections with the DL area (Greenwood, 2000). Therefore, as the
DL region does not function in isolation, age-related differences in performance on measures sensitive to DL prefrontal dysfunction may not be solely due to DL deterioration, but deterioration in nonfrontal regions projecting into the DL region may also have an impact on performance.

An overall analysis of the pattern of age effects for DL and VM tasks using canonical correlational analysis revealed that performance on the emotion identification task is associated with healthy adult aging. Closer examination of the emotion identification task performance, using ANOVA, revealed that this age impairment was restricted to labeling the emotion sadness, but none of the other emotions. This finding receives support from previous studies in the literature, demonstrating that the perception of negative emotions specifically changes with age, for example, sadness (Moreno, Borod, Welkowitz, & Alpert, 1993) or anger (Malatesta et al., 1987). Yet, the removal of the age effect on the emotion identification task when the temporal-lobe memory measures were controlled for could lend support for the involvement of the temporal-lobe region in emotion perception, a finding previously reported in the literature (Streit et al., 1999; Weddell, 1994). As the medial temporal-lobe region is sensitive to the effects of healthy adult aging, medial temporal-lobe deterioration as opposed to VM deterioration may be responsible for the age differences in performance.

The lack of an age effect on several of the WCST measures, such as the total number of nonperseverative responses, number of categories achieved, and failure to maintain set does not necessarily provide evidence against a DL prefrontal theory of aging. Stuss et al. (2000) claimed that performance on different WCST indices may be associated with different regions of the frontal lobes, therefore DL and VM lesions cause different patterns of deficits on the task. Damage to the DL region is more likely to produce an increase in the number of perseverative errors on the WCST, whereas damage to the VM region is more likely to produce failures to maintain set: This would fit with the current pattern of age stability in maintaining set but decline in perseverative errors.

Although age-related differences in performance on the gambling task were not found, the performance of the participants in this study was not consistent with the performance of healthy participants in previous studies of the gambling task (Bechara et al., 1994, 1998). Bechara and colleagues reported that healthy

Table 4: Means and Standard Deviations for the Measures on the Faux Pas Task for the Young, Middle-Aged, and Older Groups, and the Univariate Analyses

<table>
<thead>
<tr>
<th>Error</th>
<th>Young</th>
<th>Middle-aged</th>
<th>Older</th>
<th>F(2, 85)</th>
<th>MSE</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faux pas stories</td>
<td>8.6 5.4</td>
<td>9.7 5.7</td>
<td>7.9 4.0</td>
<td>1.06</td>
<td>8.70</td>
<td>.35</td>
<td>.02</td>
</tr>
<tr>
<td>Detecting faux pas</td>
<td>0.3 0.6</td>
<td>0.2 0.4</td>
<td>0.4 0.9</td>
<td>0.75</td>
<td>9.24</td>
<td>.48</td>
<td>.02</td>
</tr>
<tr>
<td>Control questions</td>
<td>1.8 1.6</td>
<td>2.1 1.6</td>
<td>1.8 1.5</td>
<td>2.21</td>
<td>5.78</td>
<td>.12</td>
<td>.05</td>
</tr>
<tr>
<td>Empathy questions</td>
<td>0.96 0.04</td>
<td>0.95 0.06</td>
<td>0.95 0.06</td>
<td>1.35</td>
<td>2.39</td>
<td>.26</td>
<td>.03</td>
</tr>
<tr>
<td>A' (sensitivity)</td>
<td>0.25 0.62</td>
<td>0.28 0.66</td>
<td>−0.19 0.68</td>
<td>0.88</td>
<td>3.02</td>
<td>.42</td>
<td>.02</td>
</tr>
<tr>
<td>B_D (bias)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Canonical Loadings Between the Dorsolateral (DL) and Ventromedial (VM) Measures and the Corresponding Canonical Variate (Linear Combination) of the DL and VM Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Canonical loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCSTa</td>
<td>.54</td>
</tr>
<tr>
<td>SOPTb</td>
<td>.86</td>
</tr>
<tr>
<td>Delayed response taska</td>
<td>.45</td>
</tr>
<tr>
<td>Gambling taskb</td>
<td>.09</td>
</tr>
<tr>
<td>Faux pas taska</td>
<td>−.01</td>
</tr>
<tr>
<td>Emotion identification taska</td>
<td>.39</td>
</tr>
</tbody>
</table>

Note: Canonical loadings above .30 demonstrate a significant correlation between the respective dependent variable and the linear combination of the DL and VM measures. WCST = Wisconsin Card Sorting Task; SOPT = Self-Ordered Pointing Task.

*Total errors.  b Total selections from Decks A and B.
participants tend to show a preference for Decks C and D when performing the gambling task, whereas in this experiment participants showed a preference for Decks B and D. However, the current findings of a preference for Decks B and D have also been reported in healthy participants by Wilder, Weinberger, and Goldberg (1998). This pattern of performance indicates that participants showed a preference for Decks B and D. However, the current findings of a preference for Decks B and D have also been reported in healthy participants by Wilder, Weinberger, and Goldberg (1998). This pattern of performance indicates that participants were influenced by instant rewards rather than by the longer term profits and losses associated with the task. Participants seemed to believe that the small frequent penalties had a more negative impact on their long-term profit than the large infrequent penalties. Indeed, studies of risk-taking behavior and gambling have demonstrated that an individual’s choices are not affected by the amount of money already won or lost on previous trials but the ratio of wins to losses (Greenberg & Weiner, 1966). Furthermore, when individuals are faced with ambiguous information and have to make decisions, they rely on the encoding of frequency information rather than the amount of reward (Hascher & Zacks, 1984).

It could be argued that the age-related cognitive differences in performance on the DL tasks are due to a general reduction in information-processing speed rather than localized changes in the DL region per se (Saltzhouse, 1996). Accordingly, the DL tasks would involve speed factors that are affected by aging whereas the VM tasks would involve speed factors that are insensitive to aging. Indeed, studies have demonstrated that speed of processing is a determinant of age-related variance on frontal-lobe measures (Parkin & Java, 1999). However, studies have also shown that significant age effects are still observable on frontal-lobe tasks such as fluency (Keys & White, 2000), the Trail Making Test (Keys & White, 2000), and the WCST (Saltzhouse, Fristoe, & Rhee, 1996), even when speed of processing is controlled for. Therefore, although speed of processing may play a part in age-related cognitive deficits in performance, it cannot exclusively account for the age effects (Keys & White, 2000).

Another possible explanation for the age-related differences in performance on the DL tasks is that older adults depend on the cognitive functions mediated by the DL regions to perform the DL tasks whereas younger adults do not. Therefore, older adults perform tasks such as the delayed-response task, SOPT, or WCST, demands are placed on the DL region, the region becomes overloaded, and task performance deteriorates. Indeed, the few neuroimaging studies that have examined the effects of aging on brain activity have shown that older adults recruit different brain regions compared with younger adults while performing the same cognitive tasks (for a review see Grady, 2000). However, although increases in DL activation with age have been found during a face working memory task (Grady et al., 1998), other neuroimaging studies have demonstrated reductions in DL activation with age during the WCST (Esposito, Kirkby, Van Horn, Ellmore, & Berman, 1999; Nagahama et al., 1997).

The selective age effects in terms of certain frontal-lobe measures may be accounted for in terms of an age-related decline in the functional integrity of the dopaminergic system. Dopamine is a neurotransmitter that is found in high levels in the prefrontal cortex and studies suggest that the dopaminergic projections to the frontal lobes may be important for cognitive functions (Brozoski, Brown, Rosvold, & Goldman, 1979; Williams & Goldman-Rakic, 1995). More specifically, the dopaminergic system has been associated with performance on the SOPT (West, Ergis, Winocur, & Saint-Cyr, 1998). Age-related deterioration in the dopaminergic system may explain the age effects on the DL tasks, but no conclusion can be reached on the basis of the current data.

In summary, the results from the current study support a specific DL prefrontal theory of cognitive changes with age, rather than a global decline in frontal-lobe function. Healthy adult aging mainly affects abilities mediated by the DL prefrontal area, such as executive functions and working memory, leaving abilities mediated by the VM prefrontal area, such as emotional processing and social behavior, relatively intact. This finding has implications for the current neuropsychological models of aging, which tend to suggest that global frontal-lobe deterioration is responsible for age-related cognitive decline.

References


