Early assessment of dementia: the contribution of different memory components
Spaan, P.E.J.
CHAPTER IV: RESULTS AND DISCUSSION OF THE SECOND ADMINISTRATION OF THE MEMORY TEST BATTERY

In this chapter, the results of the second administration of the memory test battery (T2) will be presented. In addition, the effect of the dementia diagnoses, made at T2 according to the CAMDEX administration (i.e., ‘dementia’ defined by a CAMDEX classification of mild, moderate or severe dementia, consistent with the DSM-IV criteria of dementia), will be examined. On the one hand, the development of memory performance in nondemented versus demented elderly subjects over the two-year time period will be examined. On the other hand, the accuracy of the memory test battery in predicting dementia, two years before the clinical (DSM-IV) diagnosis could be made, will be investigated.

The first two sections present general characteristics of the memory test battery, known from the T2 data. The reliability of the battery and the influence of the demographic and the screening variables will be investigated. Differences with the findings on the T1 data, described in Chapter III, will be discussed.

Section 3 will discuss whether the level of memory functioning in normal elderly subjects was originally (at T1) underestimated by not excluding the ‘preclinical dementia’ subjects (i.e., the subjects identified as demented at T2; after Sliwinski, Lipton, Buschke & Stewart, 1996). In this way, norms for memory functioning in ‘normal ageing’ will be provided, free from pathological ageing processes (i.e., preclinical or early-stage dementia).

In section 4, the development of memory performance over time will be examined. The elderly subjects will be subdivided into various clinical subgroups, taking into account the specific classification type of the CAMDEX (‘nondemented’, ‘minimal dementia’, ‘mild or moderate dementia’) and the global cognitive status of the nondemented subjects (according to MMSE: cognitively impaired or cognitively healthy, as was defined in Chapter III). Effects of administration period, clinical subgroup and interaction between these two factors will be analysed for each subtest separately. In addition, the memory measures will be determined that are most sensitive to decline when subjects become demented. Consequently, the degree and nature of decline over time on the memory test battery will be illustrated for elderly subjects diagnosed at T2 as nondemented, minimally demented, or mildly or moderately demented. Thus, the sensitivity of the various memory measures for global cognitive status and severity of dementia will be discussed.

In section 5, the preclinical assessment of dementia will be discussed. By examining memory performance, known from T1, of the subjects that were vs. were not diagnosed as demented two years later, information can be gathered concerning the predictors of dementia. It will be investigated how the elderly subjects diagnosed as demented at T2 can best be differentiated from the nondemented elderly subjects. In other words, which profile of memory measures is indicative for dementia, approximately two years before diagnosis? Thus, the cognitive predictors of developing dementia in the near future will be examined. In addition, it will be investigated whether these memory measures lead to an improvement in the prediction of dementia, above what is possible using the current clinical assessment methods (i.e., MMSE and memory components that are implicated in memory tests used in clinical practice), but also using demographic data of the subjects (age,
education and sex). Furthermore, the profile of the minimally demented subjects, though not officially demented according to DSM-IV criteria, will be examined in further detail.

In section 6, the memory profile of the clinically demented subjects will be described by means of the memory subtest data collected at T2. While section 5 describes the profile of the preclinically demented subjects (two years before the diagnosis was made), section 6 presents cross-sectional analyses of the T2 measures that best discriminate between demented and nondemented subjects, at the time of diagnosis. It will be investigated if and how the memory performance characteristics of the preclinically and the clinically demented subjects differ.

Furthermore, section 7 will focus on the identification of latent subgroups within the sample of clinically nondemented subjects. An attempt will be made to identify subjects that demonstrated (greater than average) decline on the memory test battery from T1 to T2, but nonetheless were not diagnosed as demented at T2. It will be examined whether the subjects that showed greater than average decline from T1 to T2 on the memory test battery, were the same subjects that were ‘falsely’ classified to the demented group, based on the profile of memory measures indicative for dementia (determined in section 5). Thus, the subjects will be defined that may be considered as being at risk for developing dementia in the near future. If possible, it will be examined whether these subjects show a profile of memory performance that is qualitatively and/or quantitatively similar to the profile described for the subjects that were actually diagnosed as demented at T2.

Finally, findings described in the current chapter will be integrated and inconsistencies will be discussed. In addition, shortcomings of the battery in assessing memory performance of elderly subjects (over time) and identifying dementia cases as well as possible improvements will be discussed.

1. Reliability analyses of the memory test battery

In this section, the level of internal consistency (measured by Cronbach’s alpha) of each subtest, administered at T2, will be discussed. In addition, the general level of internal consistency per subtest will be determined by performing the internal consistency analysis over all available data (i.e., the available subtest data collected at T1 and T2 together). Table 1a summarises the internal consistency data of the memory test battery.

It may be noted from Table 1a that the T2 data generally led to higher levels of internal consistency than the T1 data. Specifically, the ‘Ten word list-learning test’ and the ‘Visual Association Test’ showed improved Cronbach’s alpha’s. To a lesser degree, the ‘Category fluency test’, the ‘priming’ measure of the ‘Perceptual identification task’ and the ‘Word stem completion task’ also showed increased levels of internal consistency. However, both priming measures should still be considered unreliable. In addition, it must be noted that the ‘Two-alternative word-recognition test’ demonstrated a clear decrease in internal consistency at T2, relative to T1.

As was done in Chapter III, Table 1b presents the items per subtest that decreased the level of internal consistency of the particular subtest, administered at T2.
Table 1a: Reliability (internal consistency) of each subtest, measured by Cronbach’s alpha; analyses over all subjects tested with valid data available, at T₁, T₂, and T₁ and T₂ together. The items per subtest over which the internal consistency analyses were performed, are described in Table 2a of Chapter III.

<table>
<thead>
<tr>
<th>subtest:</th>
<th>T₁ n: ss. (missings); variables:</th>
<th>T₁ Cronbach’s alpha²⁷</th>
<th>T₂ n: ss. (missings); variables:</th>
<th>T₂ Cronbach’s alpha</th>
<th>T₁+₂ n: ss. (missings); variables:</th>
<th>T₁+₂ Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ten word list-learning test</td>
<td>158 ss. (0); 10 var.;</td>
<td>.6233</td>
<td>103 ss. (0); 10 var.;</td>
<td>.8029</td>
<td>261 ss. (0); 10 var.;</td>
<td>.7183</td>
</tr>
<tr>
<td>Digit span task</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Word-recognition test</td>
<td>157 ss. (1); 20 var.;</td>
<td>.6427</td>
<td>103 ss. (0); 20 var.;</td>
<td>.6778</td>
<td>260 ss. (1); 20 var.;</td>
<td>.6569</td>
</tr>
<tr>
<td>Paired-associate learning test</td>
<td>158 ss. (0); 10 var.;</td>
<td>.8514</td>
<td>102 ss. (1); 10 var.;</td>
<td>.8906</td>
<td>260 ss. (1); 10 var.;</td>
<td>.8724</td>
</tr>
<tr>
<td>Black span task</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Word stem completion task</td>
<td>155 ss. (3); 8 var.;</td>
<td>.2274</td>
<td>102 ss. (1); 8 var.;</td>
<td>.2950</td>
<td>257 ss. (4); 9 var.;</td>
<td>.2583</td>
</tr>
<tr>
<td>Category fluency test</td>
<td>157 ss. (1); 2 var.;</td>
<td>.7026</td>
<td>103 ss. (0); 2 var.;</td>
<td>.7679</td>
<td>260 ss. (1); 2 var.;</td>
<td>.7300</td>
</tr>
<tr>
<td>Mirror-reading task</td>
<td>142 ss. (16); 3 var.;</td>
<td>.8288</td>
<td>83 ss. (20); 3 var.;</td>
<td>.8522</td>
<td>225 ss. (36); 3 var.;</td>
<td>.8369</td>
</tr>
<tr>
<td>Perceptual identification task: ‘semantic memory’</td>
<td>142 ss. (16); 36 var.; 9796</td>
<td>91 ss. (12); 36 var.; 9704</td>
<td>9630</td>
<td>233 ss. (28); 36 var.; 9768</td>
<td>9924</td>
<td></td>
</tr>
<tr>
<td>Perceptual identification task: ‘priming’</td>
<td>146 ss. (12); 12 var.; 2864</td>
<td>92 ss. (11); 12 var.; 3624</td>
<td>.3624</td>
<td>235 ss. (23); 12 var.; .3465</td>
<td>.3097</td>
<td></td>
</tr>
<tr>
<td>Two-alternative word-recognition test</td>
<td>155 ss. (3); 10 var.; 6964</td>
<td>102 ss. (1); 10 var.; 5618</td>
<td>.5572</td>
<td>257 ss. (4); 10 var.; .6717</td>
<td>.6690</td>
<td></td>
</tr>
<tr>
<td>Visual Association Test</td>
<td>157 ss. (1); 6 var.;</td>
<td>.7127</td>
<td>103 ss. (0); 6 var.;</td>
<td>.8104</td>
<td>260 ss. (1); 6 var.;</td>
<td>.7543</td>
</tr>
</tbody>
</table>

Table 1b: Items per subtest that lead to a higher alpha value when the particular item was deleted from the subtest (‘alpha if item deleted’ > ‘alpha’).

<table>
<thead>
<tr>
<th>subtest:</th>
<th>‘alpha’:</th>
<th>number of items in subtest:</th>
<th>items:</th>
<th>‘alpha if item deleted’:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word-recognition test</td>
<td>.6778</td>
<td>20</td>
<td>handdoek</td>
<td>.6864</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>circus</td>
<td>.6958</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>slinger</td>
<td>.6781</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>vuilnis</td>
<td>.6790</td>
</tr>
<tr>
<td>Word stem completion task</td>
<td>.2950</td>
<td>8</td>
<td>ke</td>
<td>.3196</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>za</td>
<td>.3196</td>
</tr>
<tr>
<td>Perceptual identification task: ‘semantic memory’</td>
<td>.9630</td>
<td>36</td>
<td>kieuw</td>
<td>.9639</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>vergiet</td>
<td>.9633</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>lollie</td>
<td>.9635</td>
</tr>
</tbody>
</table>

²⁷ The first value represents ‘alpha’; the second value represents ‘standardised item alpha’, which is the ‘alpha’ calculated over items with equal means and variances (i.e., parallel items).

²⁸ The stems ‘ba’ and ‘lo’ could not be analysed, because variance was equal to 0.
Table 1b (continued)

<table>
<thead>
<tr>
<th>subtest:</th>
<th>‘alpha’:</th>
<th>number of items in subtest:</th>
<th>items:</th>
<th>‘alpha if item deleted’:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceptual identification task: ‘priming’</td>
<td>.3624</td>
<td>12</td>
<td>hak, inzicht, slager, dame</td>
<td>.3834, .3643, .3647, .3990</td>
</tr>
<tr>
<td>Two-alternative word-recognition test</td>
<td>.5618</td>
<td>10</td>
<td>polder – vallei, handdoek – zakdoek</td>
<td>.5755, .5693</td>
</tr>
<tr>
<td>Visual Association Test</td>
<td>.8104</td>
<td>6</td>
<td>dobbelsteen</td>
<td>.8465</td>
</tr>
</tbody>
</table>

Items listed above usually reflect best or worst performance, relative to the other items of the particular subtest. As was noted in Chapter III, both priming measures show relatively many items that decrease Cronbach’s alpha, which is consistent with their low level of internal consistency. Generally, fewer items are listed here than in Table 2b of Chapter III, which may be related to the higher level of reliability of most subtests, administered at T2.

**DISCUSSION**

The generally higher level of internal consistency of the subtests administered at T2 may be explained by the greater range of scores at T2, relative to the range of scores at T1. It may be noted that on the subtests that showed improved reliability, the T2 (mildly or moderately) demented subjects performed worse at T2 than the seven T1 demented subjects did at T1, which resulted in a greater range of scores at T2 than at T1. This difference in performance level by T1 vs. T2 demented subjects cannot be explained by significant differences in age ($r=-.871$, $df=12$, $p=.001$), education ($r=-.665$, $df=12$, $p=.518$) or sex ($\chi^2=.000$, $df=1$, $p=1.000$). Furthermore, the T2 demented subjects scored worse at the MMSE than the T1 demented subjects (mean scores 18.43 (4.86) vs. 22.00 (4.32)), though the difference was not significant ($r=-1.453$, $df=1$, $p=.12$). However, only seven subjects were available for both groups. In addition, the disease duration per subject is unknown, which makes it a heterogeneous sample of patients. The T2 demented group includes five mild and two moderate dementia cases, while the T1 demented group includes six mild and one moderate dementia case. These aspects may explain the differences found between the T1 and the T2 data.

### 2. Influence of the demographic and the screening variables on each subtest

In this section, analyses are presented of the influence of the demographic variables (age, education and sex) and the screening variables – global cognitive status (measured by the MMSE at T2), and depression (measured by the CES-D at T2) – on the various memory subtests (administered at T2). The same analyses were performed with T1 data (see Chapter III, section 3); evident differences in correlational values between T1 and T2 data will be discussed. Table 2 presents the results of the correlational analyses.

---

29 The subjects were also screened at T2 for CVA residual symptoms, but no subjects reported having them.
Table 2: Pearson correlations between the main score of each subtest at T2 and each of the variables age, years of education, MMSE score at T2 and CES-D score at T2 (analysed over the subjects that were tested at T2 and who were not diagnosed as being demented at T2\(^{30}\)); each cell presents the correlation coefficient and the significance level. Eta correlation measures between the nominal variable sex and the main score of each subtest at T2; each cell presents the Eta value and the significance level.

<table>
<thead>
<tr>
<th>subtest:</th>
<th>n</th>
<th>age</th>
<th>education</th>
<th>sex</th>
<th>MMSE-T2</th>
<th>CES-D-T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ten word list-learning test</td>
<td>96</td>
<td>-.329**</td>
<td>-.017</td>
<td>.303**</td>
<td>.581**</td>
<td>-.149</td>
</tr>
<tr>
<td>Digit span task</td>
<td>96</td>
<td>.052</td>
<td>.197</td>
<td>.183</td>
<td>.437**</td>
<td>.043</td>
</tr>
<tr>
<td>Word-recognition test</td>
<td>96</td>
<td>-.175</td>
<td>-.076</td>
<td>.268**</td>
<td>.536**</td>
<td>-.143</td>
</tr>
<tr>
<td>Paired-associate learning test</td>
<td>96</td>
<td>-.079</td>
<td>-.031</td>
<td>.274**</td>
<td>.636**</td>
<td>-.161</td>
</tr>
<tr>
<td>Block span task</td>
<td>96</td>
<td>-.229*</td>
<td>.297**</td>
<td>.219*</td>
<td>.289**</td>
<td>-.115</td>
</tr>
<tr>
<td>Word stem completion task</td>
<td>96</td>
<td>-.308**</td>
<td>.056</td>
<td>.165</td>
<td>-.118</td>
<td>-.161</td>
</tr>
<tr>
<td>Category fluency test</td>
<td>96</td>
<td>-.147</td>
<td>.050</td>
<td>.079</td>
<td>.479**</td>
<td>-.133</td>
</tr>
<tr>
<td>Mirror-reading task</td>
<td>80</td>
<td>-.025</td>
<td>-.019</td>
<td>.075</td>
<td>-.430**</td>
<td>.029</td>
</tr>
<tr>
<td>Perceptual identification task: 'semantic memory'</td>
<td>91</td>
<td>.145</td>
<td>-.008</td>
<td>.094</td>
<td>-.445**</td>
<td>.035</td>
</tr>
<tr>
<td>Perceptual identification task: 'priming'</td>
<td>91</td>
<td>.149</td>
<td>.055</td>
<td>.154</td>
<td>.129</td>
<td>.063</td>
</tr>
<tr>
<td>Two-alternative word-recognition test</td>
<td>96</td>
<td>-.169</td>
<td>.034</td>
<td>.066</td>
<td>.499**</td>
<td>-.220*</td>
</tr>
<tr>
<td>Visual Association Test</td>
<td>96</td>
<td>-.171</td>
<td>.020</td>
<td>.019</td>
<td>.449**</td>
<td>.028</td>
</tr>
</tbody>
</table>

** significant correlation at the .01 level (two-tailed); * significant correlation at the .05 level (two-tailed);

It may be noted that fewer subtests, administered at T2, correlated significantly with age than during the first administration of the memory test battery. T2 data demonstrate that performance level on the ‘Word stem completion task’, the ‘Ten word list-learning test’ (both \( p < .01 \)), and the ‘Block span task’ (\( p < .05 \)) decreased significantly with increasing age. Remarkable is the correlation between the ‘Word stem completion task’ and age, considering its low level of internal consistency and lack of correlation with age at T1. The ‘Paired-associate learning test’, the ‘Category fluency test’, the ‘Two-alternative word-recognition test’, the ‘Visual Association Test’ (at T1 all \( p < .01 \)) and the ‘Word-recognition test’ (at T1 \( p < .05 \)) showed

\(^{30}\) Subjects were considered ‘demented’ in the analyses of this chapter if they were diagnosed with ‘mild’, ‘moderate’ or ‘severe’ dementia according to CAMDEX administered at T2 (i.e., officially demented according to DSM-IV criteria).
decreasing correlation coefficients, leading to non-significant correlations with age at T2.

The education effects are similar to the effects found at T1. Both at T1 and at T2, the ‘Block span task’ (p<.01) and the ‘Digit span task’ (p=.055, close to significance) show better performance in more highly educated subjects.

The correlational analyses with sex show some differences with the T1 data. The clear male advantage at the ‘Category fluency test’ (p<.05) at T1 has disappeared, though male subjects still perform significantly better at the ‘Block span task’ (p<.05). Instead, the T2 data show that female subjects perform significantly better than male subjects at the ‘Ten word list-learning test’ and the ‘Word-recognition test’ (both p<.01), and the ‘Paired-associate learning test’ (p<.01; at T1, p<.05).

The correlations between the subtests and the MMSE administered at T2 are similar to the patterns at T1. However, it should be noted that the MMSE score at T1 was based on three MMSE administrations, while at T2 only one score was available. The only difference with T1 is the decreased correlation coefficient of the ‘priming’ measure of the ‘Perceptual identification task’ (T1; p<.05).

The correlational analyses with depression (i.e., CES-D score at T2) show that more ‘depressed’ subjects performed significantly worse at the ‘Two-alternative word-recognition test’ (p<.05). The other subtests do not show significant correlations with depression. Although four subjects exceeded the cut-off value of the CES-D, only one subject was observed to have rather poor concentration and motivation, which might indicate a depression and which may have interfered with her performance level. The other subjects were highly motivated and did their best when administered the memory test battery.

DISCUSSION
First, it must be noted that the current correlational analyses were performed over fewer subjects than the T1 correlational analyses (T1: n=147; T2: n=96). In addition, the subjects that were excluded from the current analyses (because of T2 dementia or drop-out) were significantly older than the included subjects (t=-3.248, df=145, p=.001; mean difference 4.41 years). Furthermore, the T1 MMSE score of the excluded subjects was significantly lower than the T2 MMSE score of the included subjects (t=3.400, df=145, p=.001; mean difference 1.50 points). The other variables did not show differences between the included and excluded subjects. These factors may explain the difference in effects between the T1 and T2 data.

In addition, it may be concluded that the effect of the T2 CES-D score on the ‘Paired-associate learning test’ and the ‘Two-alternative word-recognition test’ is clinically non-significant, because the subjects generally did not demonstrate behavioural symptoms of depression that may have affected their memory performance level.

3. The effects of ‘preclinical dementia’ on the level of memory performance in normal ageing

This section will examine how the level and nature of memory performance of truly nondemented elderly subjects (i.e., the subjects who have not become demented at T2) may be characterised. ‘Preclinical dementia’ cases may affect normative data of cognitive functioning for elderly subjects. Therefore, Sliwinski, Lipton, Buschke and Stewart (1996) investigated the effects of preclinical dementia on the level of cognitive functioning in normal ageing. They defined preclinical dementia as ‘... the
stage of a dementing disease (e.g., Alzheimer’s) in which cognitive decline is so small that individuals still perform within normal limits on measures of cognitive ability’ (p. P217). In this study, all subjects were clinically nondemented at baseline; Sliwinski et al. examined the effect on cognitive functioning in the baseline data after excluding subjects that were diagnosed as demented at follow-up 4 or 8 years later (i.e., excluding the preclinically demented subjects). The authors found that the level of memory functioning in normal elderly subjects was originally underestimated by not excluding preclinical dementia patients: it resulted in underestimation of the mean and overestimation of the variance. In addition, they reported that the effect of age on cognitive measures was an overestimation.

Thus, it seems important to examine the effect of not excluding the subjects that were diagnosed as demented at T2, on the level and nature of the T1 memory test battery data. Table 3a illustrates the level of memory performance at T1 in each sample of elderly subjects, screened vs. not screened for ‘preclinical dementia’. Furthermore, the effect was investigated after excluding all subjects that were diagnosed as minimally demented (at T1 and/or T2), in addition to the exclusion of the subjects that were officially diagnosed as demented at T2 (mild or moderate dementia according to CAMDEX administered at T2). This effect is called the T1 performance excluded from ‘memory disordered’ subjects (see Table 3b). Since only 10 subjects were officially diagnosed demented at T2, these additional analyses provide further information on the effect of the exclusion of subjects that might develop dementia in the near future.

Table 3a: Main score per subtest (mean; SD; minimum; maximum), for each sample of subjects at T1: the sample including vs. excluding subjects that were diagnosed demented at T2 (the ‘preclinically demented’ (PCD) subjects). ‘R age’: Pearson correlation coefficient between the main score of each subtest and age. ‘Difference (in SD)’31: standard deviation estimate from the sample excluding PCD subjects. ‘Overest. age (%)’46: percentage of overestimation of the effect of age: 100 * ((‘R age’ incl. PCD) / (‘R age’ excl. PCD)) - 1).

<table>
<thead>
<tr>
<th>Subtest:</th>
<th>T1 incl. PCD ss.</th>
<th>T1 excl. PCD ss.</th>
<th>diff. (in SD)</th>
<th>overest. age (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n: 136</td>
<td>n: 137</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mini-Mental State Examination [0,30]</td>
<td>26.15 (.261)</td>
<td>26.15 (.261)</td>
<td>-.07</td>
<td>59</td>
</tr>
<tr>
<td>Ten word list-learning test [0,30]</td>
<td>12.27 (4.69)</td>
<td>12.27 (4.69)</td>
<td>-.09</td>
<td>18</td>
</tr>
<tr>
<td>Digit span task [0,10]</td>
<td>17.98 (1.94)</td>
<td>17.98 (1.94)</td>
<td>-.12</td>
<td>43</td>
</tr>
<tr>
<td>Word-recognition test [0,20]</td>
<td>16.07 (5.75)</td>
<td>16.07 (5.75)</td>
<td>-.218</td>
<td>33</td>
</tr>
<tr>
<td>Paired-associate learning test [0,30]</td>
<td>3.79 (.93)</td>
<td>3.79 (.93)</td>
<td>-.173</td>
<td>0</td>
</tr>
<tr>
<td>Block span task [0,10]</td>
<td>.45 (.71)</td>
<td>.45 (.71)</td>
<td>-.092</td>
<td>14</td>
</tr>
</tbody>
</table>

31 after Sliwinski, Lipton, Buschke & Stewart (1996).
...the second administration of the memory test battery

**Table 3a (continued)**

<table>
<thead>
<tr>
<th>Subtest:</th>
<th>T₁ incl. PCD ss.</th>
<th>T₁ excl. PCD ss.</th>
<th>diff. (in SD)</th>
<th>overest. age (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n:</td>
<td>main score:</td>
<td>R age</td>
<td>main score:</td>
</tr>
<tr>
<td>Category fluency test</td>
<td>147</td>
<td>24.32 (6.65) 9 – 46</td>
<td>-265**</td>
<td>137</td>
</tr>
<tr>
<td>Mirror-reading task</td>
<td>135</td>
<td>5.96 (3.06) 1.72 – 20.56</td>
<td>.998</td>
<td>126</td>
</tr>
<tr>
<td>Perceptual identification task: ‘semantic memory’</td>
<td>144</td>
<td>6.98 (2.37) 3.53 – 20.22</td>
<td>.159</td>
<td>134</td>
</tr>
<tr>
<td>Perceptual identification task: ‘priming’</td>
<td>141</td>
<td>.69 (.80) 1.67 – 4.08</td>
<td>.055</td>
<td>131</td>
</tr>
<tr>
<td>Two-alternative word-recognition test [0,10]</td>
<td>146</td>
<td>8.99 (1.46) 2 – 10</td>
<td>-285**</td>
<td>136</td>
</tr>
<tr>
<td>Visual Association Test [0,6]</td>
<td>147</td>
<td>5.15 (1.16) 0 – 6</td>
<td>-256**</td>
<td>137</td>
</tr>
</tbody>
</table>

**Table 3b: Main score per subtest (mean; SD; minimum; maximum), for each sample of subjects at T₁: the sample including vs. excluding subjects that were diagnosed as minimally or more severe demented at T₁ or T₂ (the ‘memory disordered’ subjects). ‘R age’: Pearson correlation coefficient between the main score of each subtest and age. ‘Difference (in SD)’: standard deviation estimate from the sample excluding memory disordered subjects. ‘Overest. age (%)’: percentage of overestimation of the effect of age.**

<table>
<thead>
<tr>
<th>Subtest:</th>
<th>T₁ incl. all ss.</th>
<th>T₁ excl. memory disordered ss.</th>
<th>diff. (in SD)</th>
<th>overest. age (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n:</td>
<td>main score:</td>
<td>R age</td>
<td>main score:</td>
</tr>
<tr>
<td>Mini-Mental State Examination [0,30]</td>
<td>147</td>
<td>25.96 (2.66) 19 – 30</td>
<td>-.162</td>
<td>118</td>
</tr>
<tr>
<td>Ten word list-learning test [0,30]</td>
<td>147</td>
<td>11.86 (4.82) 3 – 26</td>
<td>-.363**</td>
<td>118</td>
</tr>
<tr>
<td>Digit span task [0,10]</td>
<td>146</td>
<td>4.42 (.95) 2 – 7</td>
<td>.064</td>
<td>117</td>
</tr>
<tr>
<td>Word-recognition test [0,20]</td>
<td>146</td>
<td>17.85 (2.04) 10 – 20</td>
<td>-.171*</td>
<td>118</td>
</tr>
<tr>
<td>Paired-associate learning test [0,30]</td>
<td>146</td>
<td>15.39 (6.38) 0 – 30</td>
<td>-.289**</td>
<td>117</td>
</tr>
<tr>
<td>Block span task [0,10]</td>
<td>146</td>
<td>3.80 (.92) 1 – 6</td>
<td>-.173*</td>
<td>118</td>
</tr>
<tr>
<td>Word stem completion task [-10,10]</td>
<td>146</td>
<td>.44 (.69) -1 – 4</td>
<td>-.105</td>
<td>117</td>
</tr>
<tr>
<td>Category fluency test</td>
<td>147</td>
<td>24.33 (6.65) 9 – 46</td>
<td>-.265**</td>
<td>118</td>
</tr>
<tr>
<td>Mirror-reading task</td>
<td>135</td>
<td>5.97 (3.06) 1.72 – 20.56</td>
<td>.098</td>
<td>110</td>
</tr>
<tr>
<td>Perceptual identification task: ‘semantic memory’</td>
<td>144</td>
<td>6.98 (2.37) 3.53 – 20.22</td>
<td>.159</td>
<td>115</td>
</tr>
<tr>
<td>Perceptual identification task: ‘priming’</td>
<td>141</td>
<td>.69 (.80) 1.67 – 4.08</td>
<td>.055</td>
<td>113</td>
</tr>
<tr>
<td>Two-alternative word-recognition test [0,10]</td>
<td>146</td>
<td>8.10 (1.46) 2 – 10</td>
<td>-.285**</td>
<td>117</td>
</tr>
<tr>
<td>Visual Association Test [0,6]</td>
<td>147</td>
<td>5.16 (1.16) 0 – 6</td>
<td>-.256**</td>
<td>118</td>
</tr>
</tbody>
</table>

**significant correlation at the .01 level (two-tailed); * significant correlation at the .05 level (two-tailed);**

120
DISCUSSION  First, it must be noted that only a small number of subjects were excluded in the sample excluding PCD cases (6.8%). In the sample of Sliwinski et al., however, 45% of the cases were excluded (206 out of 458 subjects) in order to create the sample of subjects who did not receive a diagnosis of dementia at follow-up. In addition, the follow-up in the study of Sliwinski et al. was 4 and 8 years later, instead of only 2 years later in the current study. Thus, a larger proportion of subjects might be excluded in the current study when examined several years later. Therefore, additional analyses were performed in order to investigate the effect of excluding minimally demented subjects, who might also develop dementia in the near future.

Sliwinski et al. report that the differences between their two samples were between 0.2 and 0.3 SD. Considering that only a small proportion of subjects was excluded in the current study, the SD estimate for the ‘Visual Association Test’, the ‘Paired-associate learning test’ and the ‘Two-alternative word-recognition test’ is rather high (> 0.1 SD). However, the absolute underestimation of the mean and overestimation of the variance, when including the PCD cases, is extremely small. The SD estimates clearly increase when excluding the entire group of ‘memory disordered’ subjects (i.e., additionally excluding the subjects diagnosed with minimal dementia according to CAMDEX), though the number of excluded subjects is still very small (19.7%) compared to the sample of Sliwinski et al. The SD estimates of, particularly, the Two-alternative word-recognition test, and, to a lesser degree, the ‘Visual Association Test’, the ‘Paired-associate learning test’, the ‘Mini-Mental State Examination’, the ‘Word-recognition test’ and the ‘Ten word list-learning test’ are remarkably high in this sample (i.e., similarly high as in the study by Sliwinski et al.). However, the absolute underestimation of the mean and overestimation of the variance, when including all cases available, is again extremely small.

Focusing attention on the subtests that correlated significantly with age in the sample including PCD cases, the relatively high proportion of overestimation of the effect of age in the ‘Visual Association Test’, the ‘Word-recognition test’ and the ‘Paired-associate learning test’ may be noted. In these subtests, the exclusion of PCD cases leads to the disappearance or decrease of the significance of the age effect. When excluding the entire group of ‘memory disordered’ subjects (i.e., by additionally excluding the subjects diagnosed with minimal dementia according to CAMDEX), the degree of overestimation of the effect of age in the ‘Visual Association Test’ increases, as well as in the ‘Two-alternative word-recognition test’ and the ‘Category fluency test’.

It may be noted that most subtests mentioned above are episodic memory measures. These results will be discussed in more detail in section 5, in which the predictors of dementia are examined. However, it must be noted that investigating the effects of ‘preclinical dementia’ (or ‘minimal dementia’) on cognitive performance is complicated by the issue that one does not know for sure that the subjects that were retrospectively excluded were, in fact, in a preclinical stage of dementia at the baseline measurement time. Particularly when the interval is as large as in the study by Sliwinski et al. (4 to 8 years), conclusions should be drawn very cautiously.

It is, nonetheless, remarkable that in the current study, with an evidently smaller proportion of subjects being retrospectively excluded (6.8% and 19.7% against 45% in Sliwinski et al.), relatively high SD estimates were found, compared to Sliwinski et al. (i.e., between 0.1 and 0.2 SD in the sample excluding the preclinical dementia subjects, and between 0.2 and 0.3 SD in the sample additionally excluding the minimal dementia subjects, corresponding with the results of Sliwinski et al.). Therefore, it might be argued that Sliwinski et al. falsely excluded a certain number of
...the second administration of the memory test battery

subjects: i.e., subjects that were diagnosed as demented many years later but were actually nondemented (i.e., no pathologic process of dementia was present yet) at baseline. Though Sliwinski et al. used different cognitive measures (i.e., verbal and performance IQ measures from the WAIS and the sum of recall on the Selective Reminding Test (Buschke, 1973)), it might be concluded that the effects on cognitive performance of dementia in a preclinical stage are indeed noticeable about two years before the diagnosis was made, but may not extend to four years or longer ago.

Nevertheless, the T1 scores provided by the group of subjects excluded from subjects that were diagnosed demented at T2 or T1, might better be considered as norms for truly normal ageing, free from preclinical dementia influences.

4. Differences in memory performance (over time) between various clinical subgroups of elderly subjects, as defined at T2

In this section, the development of memory performance over time (from T1 to T2) will be examined for each clinical subgroup of elderly subjects, based on the assessment at T2. The CAMDEX, administered at T2, identified 84 nondemented (‘NonD’) subjects, 6 minimal dementia (‘Min.D’) subjects, and 7 mild or moderate dementia (‘Mild/mod.D’) subjects. In addition, the large group of nondemented subjects, was subdivided into two subgroups: 56 subjects that were classified at T1 to the Normal Control group (the ‘NonD-NC’ group; MMSE 27-30) and 28 subjects that were classified at T1 to the Cognitively Impaired group (the ‘NonD-CI’ group; MMSE 21-25). Appendix E of Chapter II presents further details of these clinical subgroups.

Table 4 presents the average data per subtest, per clinical subgroup, per administration period.

In addition, main effects of between-subjects factor ‘clinical subgroup’ and within-subjects factor ‘administration period’, as well as interaction effects between these factors, were tested for significance by means of the GLM Repeated Measures analysis, for each subtest separately. Figure 1 through 13 illustrate the effects for each subtest separately, and significant effects are described. In this way, the following three aspects are investigated: (1) the degree of difference in performance between the various clinical subgroups, (2) the degree of change from T1 to T2, and (3) how the degree of change from T1 to T2 varies per clinical subgroup. It is expected that decline on a subtest will be greater when the CAMDEX/MMSE diagnosis is more severe. In addition, the subtests will be determined that showed greatest decline of performance in demented subjects. In other words, which memory measures are most sensitive to incident dementia? Dementia is defined here as mild or moderate dementia, according to CAMDEX and consistent with the DSM-IV criteria of dementia. In addition, the role of the minimally demented subjects is investigated in order to determine whether the patterns of decline differ when considering the minimal dementia subjects. However, it must be noted that relatively few subjects developed (minimal) dementia in the two-year time period and, thus, conclusions in this respect should be regarded cautiously.

32 In addition, 6 subjects were diagnosed as nondemented at T3, while they were diagnosed as minimally demented at T1. Because of possible diagnostic errors (either at T1 or at T3), these subjects were excluded, in order to create an unambiguous ‘nondemented’ group and an unambiguous ‘minimally demented’ group.
33 In addition, there are 3 mild dementia subjects and 1 minimal dementia subject that refused or were not able to participate at T2 – these subjects are not included in this section.
Table 4: Average performance per subtest, per clinical subgroup, per administration period (mean, standard deviation, n).

<table>
<thead>
<tr>
<th>subtest:</th>
<th>NonD-NC</th>
<th>NonD-CI</th>
<th>Min.D</th>
<th>Mild/mod.D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_1$:</td>
<td>$T_2$:</td>
<td>$T_1$:</td>
<td>$T_2$:</td>
</tr>
<tr>
<td></td>
<td>(4.43)</td>
<td>(5.12)</td>
<td>(4.27)</td>
<td>(4.88)</td>
</tr>
<tr>
<td></td>
<td>$n=56$</td>
<td>$n=56$</td>
<td>$n=28$</td>
<td>$n=28$</td>
</tr>
<tr>
<td>Digit span task</td>
<td>4.71</td>
<td>4.98</td>
<td>4.18</td>
<td>4.21</td>
</tr>
<tr>
<td></td>
<td>(.94)</td>
<td>(.95)</td>
<td>(.86)</td>
<td>(.79)</td>
</tr>
<tr>
<td></td>
<td>$n=55$</td>
<td>$n=55$</td>
<td>$n=28$</td>
<td>$n=28$</td>
</tr>
<tr>
<td>Word-recognition test</td>
<td>18.57</td>
<td>18.70</td>
<td>18.04</td>
<td>17.96</td>
</tr>
<tr>
<td></td>
<td>(1.61)</td>
<td>(1.72)</td>
<td>(1.57)</td>
<td>(1.57)</td>
</tr>
<tr>
<td></td>
<td>$n=56$</td>
<td>$n=56$</td>
<td>$n=28$</td>
<td>$n=28$</td>
</tr>
<tr>
<td>Paired-associate learning test</td>
<td>19.33</td>
<td>20.64</td>
<td>15.21</td>
<td>16.04</td>
</tr>
<tr>
<td></td>
<td>(5.48)</td>
<td>(5.43)</td>
<td>(4.60)</td>
<td>(6.27)</td>
</tr>
<tr>
<td></td>
<td>$n=55$</td>
<td>$n=55$</td>
<td>$n=28$</td>
<td>$n=28$</td>
</tr>
<tr>
<td>Block span task</td>
<td>4.13</td>
<td>4.27</td>
<td>3.68</td>
<td>3.93</td>
</tr>
<tr>
<td></td>
<td>(.90)</td>
<td>(.90)</td>
<td>(.90)</td>
<td>(.94)</td>
</tr>
<tr>
<td></td>
<td>$n=56$</td>
<td>$n=56$</td>
<td>$n=28$</td>
<td>$n=28$</td>
</tr>
<tr>
<td>Word stem completion task</td>
<td>.43</td>
<td>.46</td>
<td>.61</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>(.60)</td>
<td>(.81)</td>
<td>(.79)</td>
<td>(.104)</td>
</tr>
<tr>
<td></td>
<td>$n=56$</td>
<td>$n=56$</td>
<td>$n=28$</td>
<td>$n=28$</td>
</tr>
<tr>
<td>Category fluency test</td>
<td>27.36</td>
<td>27.75</td>
<td>24.64</td>
<td>24.25</td>
</tr>
<tr>
<td></td>
<td>(6.73)</td>
<td>(7.58)</td>
<td>(4.67)</td>
<td>(4.49)</td>
</tr>
<tr>
<td></td>
<td>$n=56$</td>
<td>$n=56$</td>
<td>$n=28$</td>
<td>$n=28$</td>
</tr>
<tr>
<td>Mirror-reading task</td>
<td>4.58</td>
<td>5.36</td>
<td>7.04</td>
<td>10.85</td>
</tr>
<tr>
<td></td>
<td>(1.82)</td>
<td>(4.29)</td>
<td>(1.34)</td>
<td>(5.18)</td>
</tr>
<tr>
<td></td>
<td>$n=54$</td>
<td>$n=54$</td>
<td>$n=15$</td>
<td>$n=15$</td>
</tr>
<tr>
<td>Perceptual identification task:</td>
<td>5.95</td>
<td>5.75</td>
<td>7.58</td>
<td>7.41</td>
</tr>
<tr>
<td>'semantic memory'</td>
<td>(1.42)</td>
<td>(1.37)</td>
<td>(2.53)</td>
<td>(2.12)</td>
</tr>
<tr>
<td></td>
<td>$n=54$</td>
<td>$n=54$</td>
<td>$n=23$</td>
<td>$n=23$</td>
</tr>
<tr>
<td>Perceptual identification task:</td>
<td>.50</td>
<td>.76</td>
<td>.99</td>
<td>.73</td>
</tr>
<tr>
<td>'priming'</td>
<td>(.55)</td>
<td>(.67)</td>
<td>(.88)</td>
<td>(.84)</td>
</tr>
<tr>
<td></td>
<td>$n=54$</td>
<td>$n=54$</td>
<td>$n=23$</td>
<td>$n=23$</td>
</tr>
<tr>
<td>Two-alternative word-recognition</td>
<td>9.64</td>
<td>9.61</td>
<td>9.48</td>
<td>9.48</td>
</tr>
<tr>
<td>test</td>
<td>(1.64)</td>
<td>(1.73)</td>
<td>(.89)</td>
<td>(1.05)</td>
</tr>
<tr>
<td></td>
<td>$n=56$</td>
<td>$n=56$</td>
<td>$n=27$</td>
<td>$n=27$</td>
</tr>
<tr>
<td>Visual Association Test</td>
<td>5.54</td>
<td>5.61</td>
<td>5.32</td>
<td>5.39</td>
</tr>
<tr>
<td></td>
<td>(.69)</td>
<td>(.68)</td>
<td>(.102)</td>
<td>(.83)</td>
</tr>
<tr>
<td></td>
<td>$n=56$</td>
<td>$n=56$</td>
<td>$n=28$</td>
<td>$n=28$</td>
</tr>
<tr>
<td>Mini-Mental State Examination</td>
<td>28.23</td>
<td>28.05</td>
<td>23.50</td>
<td>25.07</td>
</tr>
<tr>
<td></td>
<td>(1.03)</td>
<td>(1.39)</td>
<td>(1.64)</td>
<td>(2.45)</td>
</tr>
<tr>
<td></td>
<td>$n=56$</td>
<td>$n=56$</td>
<td>$n=28$</td>
<td>$n=28$</td>
</tr>
</tbody>
</table>

4.1. Ten word list-learning test

Significant effects were found for within-subjects factor ‘administration period’ ($F=9.060, df=1, p=.003$) and between-subjects factor ‘clinical subgroup’ ($F=20.026, df=3, p=.000$). Bonferroni tests showed significant differences between all pairs of
means in the expected directions ($p$ levels ranging from .000 to .004), except between the NonD-CI group and the Min.D group, and between the Min.D group and the Mild/mod.D group. As is illustrated in Figure 1, the Min.D and the Mild/mod.D group decline from T1 to T2 (especially the Mild/mod.D group who scores just above 0), while the NonD-NC and the NonD-CI group perform relatively stable over time (significant interaction effect between ‘administration period’ and ‘clinical subgroup’: $F=8.325$, $df=3$, $p=.000$).

**Figure 1:** T1 and T2 performance on the ‘Ten word list-learning test’ by various clinical subgroups of elderly subjects.

4.2. Digit span task
Significant effects were found for ‘administration period’ ($F=5.120$, $df=1$, $p=.026$) and ‘clinical subgroup’ ($F=5.570$, $df=3$, $p=.001$). Bonferroni tests showed that the NonD-NC group performed better than the NonD-CI group ($p=.004$), but no significant differences were found between the other pairs of means. As is illustrated in Figure 2, both the Min.D and the Mild/mod.D group show an evident drop in performance at T2 (especially the Mild/mod.D subjects), while the NonD-NC and the NonD-CI group perform relatively stable over time (significant interaction effect: $F=4.550$, $df=3$, $p=.005$).
Figure 2: $T_1$ and $T_2$ performance on the ‘Digit span task’ by various clinical subgroups of elderly subjects.

4.3. Word-recognition test
A significant effect for ‘administration period’ ($F=8.819$, $df=1$, $p=.004$) and for ‘clinical subgroup’ ($F=16.056$, $df=3$, $p=.000$) was found. This test cannot differentiate significantly between subgroups that are clinically most closely related. Therefore, it does not seem sensitive to severity of dementia or cognitive status of nondemented subjects. It only differentiates subjects that differ greatly in clinical ‘rank’ (i.e., NonD-NC vs. minimally and mildly demented subjects ($p=.009$ and .000), NonD-CI vs. mildly demented subjects ($p=.000$)). As was found in the previous two subtests, the ‘Word-recognition test’ shows a decline in performance over time by the Min.D and the Mild/mod.D subjects, while the NonD-NC and the NonD-CI subjects performed relatively stable over time (significant interaction effect: $F=4.149$, $df=3$, $p=.008$). See Figure 3 for an illustration.

Figure 3: $T_1$ and $T_2$ performance on the ‘Word-recognition test’ by various clinical subgroups of elderly subjects.
4.4. Paired-associate learning test

In general, the ‘Paired-associate learning test’ shows stable performance from T\textsubscript{1} to T\textsubscript{2} \((F=.118, df=1, p=.732)\). The effect of ‘clinical subgroup’ was significant \((F=17.400, df=3, p=.000)\). Bonferroni tests showed significant differences between all pairs of means in the expected directions \((p\text{ levels ranging from .000 to .004})\), except between the NonD-CI group and the Min.D group, and between the Min.D group and the Mild/mod.D group. It may be noted from Figure 4 that the NonD-NC and the NonD-CI subjects slightly improved their performance level from T\textsubscript{1} to T\textsubscript{2}. The Min.D and the Mild/mod.D subjects showed a very modest decrease of performance. However, 4 of the 6 available Mild/mod.D subjects recalled only 2 pairs or less, 1 subject recalled 6 pairs, and 1 subject recalled 25 pairs. Thus, most Mild/mod.D subjects performed close to 0 and, therefore, showed a clear decrease of performance. However, the interaction effect between ‘administration period’ and ‘clinical subgroup’ was not significant \((F=1.570, df=3, p=.202)\).

Figure 4: T\textsubscript{1} and T\textsubscript{2} performance on the ‘Paired-associate learning test’ by various clinical subgroups of elderly subjects.

4.5. Block span task

Close to significant effects were found for ‘administration period’ \((F=3.658, df=1, p=.058)\) and ‘clinical subgroup’ \((F=2.261, df=3, p=.087)\). It may noted from Figure 5 that only the NonD-NC subjects perform at a higher level than the other subjects, but the absolute score difference is only small (half a block; no significant difference were found). It may be concluded that all groups perform relatively stable over time (no significant interaction effect: \(F=.338, df=3, p=.798\)). It may be argued that this subtest is not useful in differentiating between the various clinical subgroups.
Figure 5: $T_1$ and $T_2$ performance on the ‘Block span task’ by various clinical subgroups of elderly subjects.

4.6. Word stem completion task
As was demonstrated at $T_1$, the ‘Word stem completion task’ did not show significant priming effects at $T_2$ either. The subjects hardly ever completed the experimental stems with words from the ‘Ten word list learning test’. Although the NonD-CI subjects showed an increase in priming score (see Figure 6; significant effect of ‘clinical subgroup’: $F=3.739$, $df=3$, $p=.014$; Bonferroni: significantly higher scores than the NonD-NC ($p=.039$) and the Min.D group ($p=.085$)), they still completed only one of the ten experimental stems with a target word (over the whole task). Similar to $T_1$, this task had a very low level of internal consistency. No significant effects were found for ‘administration period’ ($F=.591$, $df=1$, $p=.444$) or the interaction with ‘clinical subgroup’ ($F=1.121$, $df=3$, $p=.345$).

Figure 6: $T_1$ and $T_2$ performance on the ‘Word stem completion task’ by various clinical subgroups of elderly subjects.
4.7. Category fluency test
Significant effects were found for ‘administration period’ \( (F=4.258, df=1, p=.042) \) and ‘clinical subgroup’ \( (F=13.140, df=3, p=.000) \). This measure cannot differentiate significantly between subgroups that are clinically most closely related. It only differentiates subjects that differ greatly in clinical ‘rank’ (i.e., NonD-NC vs. minimally and mildly demented subjects \( (p=.042 \) and \( .000 \)), NonD-CI vs. mildly demented subjects \( (p=.000) \)). As is illustrated in Figure 7, the Mild/mod.D group showed a decrease in performance from T1 to T2; the other groups showed stable performance over time (slightly significant interaction effect: \( F=3.105, df=3, p=.030 \)). Whether other scores than the sum of correct exemplars show a different pattern over the clinical subgroups will be investigated in Chapter V.

Figure 7: T1 and T2 performance on the ‘Category fluency test’ by various clinical subgroups of elderly subjects.

4.8. Mirror-reading task
Significant effects were found for ‘administration period’ \( (F=9.417, df=1, p=.003) \), and ‘clinical subgroup’ \( (F=7.521, df=3, p=.000) \). Remarkably, the greatest difference in reading times was found between the two nondemented groups (i.e., the NonD-NC subjects reading faster than the NonD-CI subjects, \( p=.000 \)), while for the other groups no significant differences were found (see also Figure 8) Note that the ‘Mirror-reading task’ was one of the best discriminating variables regarding the differentiation between the NC and the CI group at T1 (see Chapter III). Nonetheless, the ‘Mirror-reading task’ does not seem useful in discriminating between nondemented and demented subjects. However, it should be mentioned that this task suffered from missing data (e.g., only 3 subjects were available for the Mild/mod.D group). Furthermore, it may be noted from Figure 8 that the NonD-CI and the Mild/mod.D group showed slower reading times from T1 to T2, while the NonD-NC and the Min.D group performed relatively stable over time (significant interaction effect: \( F=3.388, df=3, p=.023 \)).
Figure 8: $T_1$ and $T_2$ performance on the ‘Mirror-reading task’ by various clinical subgroups of elderly subjects.

4.9. Perceptual identification task: ‘semantic memory’
No significant effect was found for ‘administration period’ ($F=.502, df=1, p=.481$), but the effect of ‘clinical subgroup’ was significant ($F=7.256, df=3, p=.000$). As in the ‘Mirror-reading task’, a clearly significant difference was found between both nondemented groups: the NonD-NC subjects identified the words faster than the NonD-CI subjects ($p=.001$). In addition, the increase in identification times by the Mild/mod.D group at $T_2$ (of, on average, 1.10 tics – 17.6 msec.; see also Figure 9) caused a slightly significant difference with the NonD-NC group ($p=.049$). However, also the ‘Perceptual identification task’ suffered from a lot of missing data (e.g., only 4 subjects were available for the Mild/mod.D group). It may be noted that, except for the Mild/mod.D group, this task showed relatively stable performance over time (no significant interaction effect: $F=1.996, df=3, p=.121$). From these data, it may be argued that the ‘semantic memory’ measure of the ‘Perceptual identification task’ is not useful in discriminating between the demented and nondemented groups, despite the increased identification times by the Mild/mod.D group. However, the absolute difference in identification times is rather small and is based on only a few subjects.

Figure 9: $T_1$ and $T_2$ performance on the ‘semantic memory’ measure of the ‘Perceptual identification task’ by various clinical subgroups of elderly subjects.
4.10. Perceptual identification task: ‘priming’
No significant effects were found for ‘administration period’ ($F=3.70$, $df=1$, $p=.544$) and ‘clinical subgroup’ ($F=1.619$, $df=3$, $p=.191$). Figure 10 shows that the NonD-NC and, particularly, the Mild/mod.D group improve, while the NonD-CI and the Min.D group decrease in priming score over time (close to significant interaction effect: $F=2.485$, $df=3$, $p=.066$). This effect is most likely caused by the improved priming score by the four Mild/mod.D subjects. This improvement most likely originates from their generally slower identification times, described in the previous section, which gives them more opportunity to identify the repeated items faster than they did at the first presentation. Thus, it may be argued that the Mild/mod.D subjects did not really benefit to a greater degree from the repetition of items than the other subjects. Nonetheless, it may be concluded that they (implicitly) remembered the repeated items, otherwise their identification times would not differ between the first and the second presentation. However, the effects are not significant, and, thus, conclusions in this respect should be regarded cautiously.

Figure 10: $T_1$ and $T_2$ performance on the ‘priming’ measure of the ‘Perceptual identification task’ by various clinical subgroups of elderly subjects.

4.11. Two-alternative word-recognition test
No significant effect was found for ‘administration period’ ($F=2.267$, $df=1$, $p=.136$), but the effect of ‘clinical subgroup’ was significant ($F=13.633$, $df=3$, $p=.000$). Similar to the ‘Word-recognition test’, this test cannot differentiate between subgroups that are clinically most closely related. Bonferroni tests did not demonstrate significant differences between the NonD-NC and the NonD-CI group ($p=1.000$) or between the Min.D and the Mild/mod.D group ($p=1.000$). Therefore, it does not seem sensitive to severity of dementia or cognitive status of nondemented subjects. Nonetheless, the differences between the other groups, including the difference between the NonD-NC and the Min.D group, were significant ($p$ levels ranging from .000 to .005). Furthermore, a slightly significant interaction effect between ‘administration period’
and ‘clinical subgroup’ was found \((F=3.258, \text{df}=3, p=.025)\), mainly caused by the decrease in performance in the \text{Min.D} subjects, while the other groups perform relatively stable over time (see Figure 11).

**Figure 11:** \(T_1\) and \(T_2\) performance on the ‘Two-alternative word-recognition test’ by various clinical subgroups of elderly subjects.

4.12. Visual Association Test

A significant effect for ‘administration period’ \((F=16.026, \text{df}=1, p=.000)\) and for ‘clinical subgroup’ \((F=27.054, \text{df}=3, p=.000)\) was found. Bonferroni tests showed significant differences between the \text{Mild/mod.D} group and the other three groups \((p=.000)\). No significant differences were found between the other pairs of means, though the difference between the \text{NonD-NC} and the \text{Min.D} group was close to significance \((p=.050)\). The \text{NonD-NC} and the \text{NonD-CI} group performed stable from \(T_1\) to \(T_2\) and obtained near maximum scores (see Figure 12). In contrast, the \text{Min.D} and the \text{Mild/mod.D} group showed clear decline of performance (significant interaction effect: \(F=6.977, \text{df}=3, p=.000\)). However, the \(T_1\) performance level of the \text{Min.D} group resembles the \text{NonD-CI} subjects’ performance level, while the \text{Mild/mod.D} subjects already performed much worse at \(T_1\) (the clinically nondemented stage). It may be concluded that performance on the ‘Visual Association Test’ is affected by incident dementia, but it does not seem sensitive to varying cognitive status of subjects, except when dementia is at a more advanced stage.
...the second administration of the memory test battery

**Figure 12:** $T_1$ and $T_2$ performance on the ‘Visual Association Test’ by various clinical subgroups of elderly subjects.

![Graph showing performance on the 'Visual Association Test' by different groups](image)

4.13. Development of performance on the Mini-Mental State Examination (MMSE)

Since the MMSE items are included in the CAMDEX procedure, it may be expected that the Min.D and, particularly, the Mild/mod.D group will show decline on the MMSE, as opposed to the NonD-NC and the NonD-CI group. In addition, the MMSE determined the group classification at $T_1$ (i.e., NC versus CI group). Nonetheless, for reasons of comparison, the analysis was performed over the MMSE data as well.

A significant effect for ‘administration period’ ($F=24.103$, $df=1$, $p=.000$) and for ‘clinical subgroup’ ($F=79.257$, $df=3$, $p=.000$) was found. Bonferroni analyses showed significant differences between all pairs of means in the expected directions ($p$ levels ranging from .000 to .006), except between the NonD-CI group and the Min.D group ($p=1.000$). Furthermore, it may be noted that, as was expected, the Min.D and, particularly, the Mild/mod.D group decline from $T_1$ to $T_2$, while the nondemented groups perform relatively stable (the NonD-CI subjects even improve; see also Figure 13). A significant interaction effect was found between ‘administration period’ and ‘clinical subgroup’ ($F=22.553$, $df=3$, $p=.000$).
4.14. The subtests most sensitive to decline in incident dementia
It may be noted from section 4.1 through 4.13 that several subtests show decreased performance from $T_1$ to $T_2$ in the Min.D and the Mild/mod.D subjects, while the NonD-NC and the NonD-CI subjects do not decline. These subtests are most episodic memory tests (the ‘Ten word list-learning test’, the ‘Visual Association Test’ and the ‘Word-recognition test’) and the ‘Digit span task’. Another episodic memory measure, the ‘Two-alternative word-recognition test’, only shows decline in the Min.D subjects, while the ‘Category fluency test’ only shows decline in the Mild/mod.D subjects. The ‘Mirror-reading task’, finally, shows decline in the Mild/mod.D and the NonD-CI subjects, whereas the NonD-NC and the Min.D subjects perform stable over time. The remaining memory measures do not show clear differences in patterns of development from $T_1$ to $T_2$ between the four clinical subgroups, as was indicated by non-significant interaction effects between ‘clinical subgroup’ and ‘administration period’.

This section investigates on which subtests the demented subjects show greatest decline from $T_1$ to $T_2$, relative to the nondemented subjects. In other words, how is the decline of performance when subjects develop dementia best characterized? In the first block of analyses, the ‘demented’ group is defined as mild or moderate dementia (according to the CAMDEX administered at $T_2$ and consistent with the DSM-IV criteria of dementia), whereas the ‘nondemented’ group consists of the NonD-NC, the NonD-CI and the Min.D subjects. In the second block of analyses, the Min.D subjects is excluded, because of the ambiguous nature of the ‘diagnosis’ (i.e., they are officially nondemented according to DSM-IV criteria, but they do present additional symptoms of memory deficits, compared to the NonD-NC and the NonD-CI subjects). However, the summary of patterns of development over time, described above, indicates that several measures show clear decline of performance in the Min.D subjects, as opposed to the NonD-NC and the NonD-CI subjects. Thus, it might be argued that the Min.D subjects are in fact in a ‘preclinical stage’ of dementia, while the diagnosis may be confirmed officially within the next few years. Therefore, on a exploratory basis, the group of subjects diagnosed with minimal
dementia at T2 (though not officially demented according to DSM-IV criteria) is classified to the ‘demented’ group in the third block of analyses.

Table 5 shows the results of several ‘stepwise’ discriminant analyses performed over difference scores (T2 score minus T1 score, per subtest). It lists the subtests that best discriminated, in terms of decline, between the demented and the nondemented subjects. As was described above, different analyses (three blocks) were performed for different definitions of ‘nondemented’ and ‘demented’. At each step, the variable that minimized the sum of the unexplained variation for all pairs of groups was entered. More variables were not entered since the significance of $F$ exceeded 0.050. First, the analysis was performed over all subjects with scores available on all subtests. However, 29 subjects were missing in this analysis, because of missing data on one or more discriminating variables. Most of these subjects showed missing data on the ‘Mirror-reading task’ and/or the ‘Perceptual identification task’. Since these two subtests did not show consistent decline in the demented subjects, relative to the nondemented subjects, a second discriminant analysis was performed without these variables. In addition, the third analysis was focused on the subjects that were classified to the Cognitively Impaired (Cl) group at T1 (i.e., the ‘dementia-at-risk’ group). Table 5 displays the results of all these analyses.

Table 5: Results of various ‘stepwise’ discriminant analyses, performed over the difference scores of each subtest (T2-T1), in discriminating between T2 ‘demented’ and ‘nondemented’ subjects. The classification of the T2 ‘minimal dementia’ subjects is manipulated.

<table>
<thead>
<tr>
<th>discriminating between:</th>
<th>performed over:</th>
<th>entered subtests</th>
<th>Residual variance</th>
<th>Exact $F$</th>
<th>df</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>nondemented/ minimal dementia vs. mild/moderate dementia (DSM-IV consistent)</td>
<td>all ss.; all subtests</td>
<td>1. Ten word list-lrn.</td>
<td>.538</td>
<td>9.849</td>
<td>68</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Category fluency</td>
<td>.430</td>
<td>7.494</td>
<td>67</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>all ss.; excl. Mirror-read. &amp; Perc.id.</td>
<td>1. Ten word list-lrn.</td>
<td>.602</td>
<td>14.733</td>
<td>82</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Category fluency</td>
<td>.517</td>
<td>10.291</td>
<td>81</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Two-alt. word-recogn.</td>
<td>.455</td>
<td>8.688</td>
<td>80</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>CI ss.; excl. Mirror-read. &amp; Perc. id.</td>
<td>1. Ten word list-lrn.</td>
<td>.624</td>
<td>9.914</td>
<td>26</td>
<td>.004</td>
</tr>
<tr>
<td>nondemented vs. mild/moderate dementia</td>
<td>all ss.; all subtests</td>
<td>1. Ten word list-lrn.</td>
<td>.511</td>
<td>10.983</td>
<td>64</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Category fluency</td>
<td>.406</td>
<td>8.248</td>
<td>63</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>all ss.; excl. Mirror-read. &amp; Perc.id.</td>
<td>1. Ten word list-lrn.</td>
<td>.550</td>
<td>18.172</td>
<td>77</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Visual Assoc. Test</td>
<td>.459</td>
<td>12.920</td>
<td>76</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Category fluency</td>
<td>.411</td>
<td>10.300</td>
<td>75</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>CI ss.; excl. Mirror-read. &amp; Perc. id.</td>
<td>1. Ten word list-lrn.</td>
<td>.439</td>
<td>20.415</td>
<td>23</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Category fluency</td>
<td>.339</td>
<td>14.951</td>
<td>22</td>
<td>.000</td>
</tr>
<tr>
<td>nondemented vs. minimal/ mild/moderate dementia</td>
<td>all ss.; all subtests</td>
<td>1. Ten word list-lrn.</td>
<td>.650</td>
<td>13.593</td>
<td>68</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Visual Assoc. Test</td>
<td>.543</td>
<td>10.436</td>
<td>67</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Two-alt. Word-recogn.</td>
<td>.485</td>
<td>8.643</td>
<td>66</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>all ss.; excl. Mirror-read. &amp; Perc.id.</td>
<td>1. Ten word list-lrn.</td>
<td>.598</td>
<td>25.745</td>
<td>82</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Visual Assoc. Test</td>
<td>.501</td>
<td>18.806</td>
<td>81</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>CI ss.; excl. Mirror-read. &amp; Perc. id.</td>
<td>3. Ten word list-lrn.</td>
<td>.448</td>
<td>28.200</td>
<td>26</td>
<td>.000</td>
</tr>
</tbody>
</table>

The MMSE was not included in these analyses, because the MMSE items are included in the CAMDEX procedure, which is responsible for the group classification concerning dementia. In addition, the MMSE determined the classification into the NonD-NC or the NonD-CI group.

134
CHAPTER IV

DISCUSSION  It may be noted from Table 5 that the results are hardly changed by classifying the minimal dementia subjects to the ‘nondemented’ group, the ‘demented’ group or by excluding them. In all three options, the ‘Ten word list-learning test’ shows most evident decline in the demented subjects, relative to the nondemented subjects. This is found regardless of where the minimally demented subjects were classified or the degree of cognitive heterogeneity of the group of subjects under investigation (the entire group of subjects or the CI group). Thus, it may be concluded that the ‘Ten word list-learning test’ shows to be sufficiently sensitive to detect decline in the minimally demented subjects (see also Figure 1).

In addition to the consistent decline on the ‘Ten word list-learning test’, some subtle differences should be noted. First, the ‘Category fluency test’ demonstrates to be sensitive to decline in incident dementia cases, but only in officially demented subjects (i.e., the Mild/mod.D group), as was also described above and illustrated by Figure 7. On the contrary, the ‘Two-alternative word-recognition test’ only shows decline in the minimally demented subjects, while the other groups (including the Mild/mod.D subjects) show stable or even improved performance over time (see also Figure 11). The ‘Visual Association Test’, finally, seems particularly sensitive to decline when the minimal dementia subjects are classified to the demented group (see also Figure 12), though the degree of decline is greater in the Mild/mod.D subjects. Nonetheless, these subtleties should be regarded cautiously since only 6 Min.D and 7 Mild/mod.D subjects were available within the entire group of subjects.

In order to examine the causes of decline from $T_1$ to $T_2$ on the ‘Ten word list-learning test’ in the minimally demented subjects and in the mildly and moderately demented subjects, Figure 14 illustrates the patterns of recall performance over the three trials in the various groups of subjects, for each administration period separately. All subjects that were diagnosed nondemented according to the CAMDEX were joined together in one ‘nondemented’ group (since the NonD-NC and the NonD-CI group both performed stable over time; see Figure 1), while the Min.D and the Mild/mod.D group were separately analysed.

Figure 14: $T_1$ and $T_2$ performance on the three trials of the ‘Ten word list-learning test’ by the minimally demented subjects ($n=6$), the mildly or moderately demented subjects ($n=7$) and the nondemented subjects (NonD-NC and NonD-CI subjects together; $n=84$). Group classification according to the CAMDEX administered at $T_2$. 

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure14.png}
\caption{$T_1$ and $T_2$ performance on the three trials of the ‘Ten word list-learning test’ by the minimally demented subjects ($n=6$), the mildly or moderately demented subjects ($n=7$) and the nondemented subjects (NonD-NC and NonD-CI subjects together; $n=84$). Group classification according to the CAMDEX administered at $T_2$.}
\end{figure}
As is illustrated by Figure 14, both the minimal and the mild/moderate dementia subjects have a rather steep learning curve at T1, comparable to the learning curve at T1 and T2 in the nondemented subjects. A GLM Repeated Measures analysis, performed over the T1 data, found no significant interaction effect between ‘trial’ and ‘demented’ (F= 0.841, df= 4, p= 0.501). At T2, the learning curve of the mild/moderate dementia subjects (in their clinical stage now) is evidently less steep – they hardly recall any words at T2. This is in contrast with the minimal dementia subjects, who recall fewer words than they did at T1 but still benefit from the repeated presentation of the words (i.e., they show a similarly steep learning curve as the nondemented subjects). A GLM Repeated Measures analysis, performed over the T2 data, found a significant interaction effect between ‘trial’ and ‘clinical subgroup’ (F= 5.730, df= 4, p= 0.000).

DISCUSSION The predictive value of the ‘Ten word list-learning test’ may be explained by its episodic memory nature. Particularly more active retrieval tasks (free recall), rather than more passive retrieval tasks (i.e., cued recall as in the ‘Paired-associate learning test’ and the ‘Visual Association Test’, or recognition as in the ‘Word-recognition test’ and the ‘Two-alternative word-recognition test’), may be most sensitive to decline when dementia is diagnosed. As was illustrated by Figure 14, the mildly and moderately demented subjects show a flat learning curve at T2: they show a floor effect and they do not benefit from repetition of words, which is in contrast with their performance at T1, the preclinical stage of dementia. The lack of benefit from the repeated presentation of words explains the large decline of recall performance and is in correspondence with literature findings on AD patients, presented in Chapter I. The minimally demented subjects, however, show a similarly steep learning curve as the nondemented subjects, both at T1 and at T2. Though the minimally demented subjects show a general decline of recall performance at T2, they still benefit from the repeated presentation of words, while the officially diagnosed as demented subjects (the Mild/mod.D group) are unable to benefit.

In general, it may concluded that the nondemented subjects did not decline in memory performance in two years of time, regardless of specific memory components. The nondemented subjects (both the NonD-NC and the NonD-CI group) even show generally higher means of scores at T2, compared to the means of scores at T1 (see Table 5), which may be explained by a general practice effect. It is highly unlikely that subjects explicitly (i.e., consciously) remembered the material of the subtests since most subjects reported not to remember the visit two years ago at all (i.e., they have been frequently visited by the LASA study in the past ten years). However, they may have gotten used to the current testing format and this may have made them feel more at ease than during the first administration. The usual LASA studies administer an interview, rather than a test as in the current study. A test may have felt like an exam, which made the subjects more nervous.

5. The profile of memory measures best predicting dementia before the diagnosis can be made

The previous section focused on patterns of memory performance over time in various clinical subgroups of elderly subjects. In this section, it is investigated how the subjects that were diagnosed nondemented at T1 but turned out to be demented at T2 (i.e., a CAMDEX classification of mild or moderate dementia, consistent with the
DSM-IV criteria), can be differentiated, in terms of their T₁ memory performance, from the initially and presently nondemented elderly subjects. By investigating the T₁ data of both groups of subjects (i.e., the stage in which dementia was not present according to the clinical DSM-IV diagnosis), information will be gathered on the specific performance characteristics of preclinically demented subjects. In other words, which profile of memory measures, is indicative of dementia, approximately two years in advance? Does this profile lead to an improvement in the prediction of dementia, above what is possible using the current clinical assessment methods (e.g., the MMSE: how well does this cognitive screening measure alone predict dementia)? Thus, which combination of subtests administered at T₁ predicts dementia at T₂ best, and how accurate is this prediction (i.e., more accurate than, for example, the MMSE score at T₁ can predict)?

In section 5.1, the analyses will be performed according to the official DSM-IV criteria of dementia. Thus, the ‘demented’ group consists of subjects diagnosed as mildly or moderately demented at T₂, while they were classified as nondemented or minimally demented according to CAMDEX at T₁. The ‘nondemented’ group consists of subjects diagnosed as nondemented or minimally dementia at T₂, while they were classified as nondemented according to CAMDEX at T₁. First, analyses will be performed over all subjects that were initially (at T₁) nondemented (n=11736; see section 5.1.1). In addition, it will be investigated how and how well the memory test battery and other variables predict dementia within the group of subjects that were classified to the ‘Cognitively Impaired’ (CI) group at T₁ (i.e., the subgroup of subjects that were suspected to develop dementia because of their impaired global cognitive status, as was determined by the MMSE; see section 5.1.2).

Furthermore, in section 5.2 the role of the minimally demented subjects (i.e., a CAMDEX classification of minimal dementia at T₂, while the CAMDEX indicated these subjects as nondemented at T₁; n=6) will be investigated on an exploratory basis. Though the subjects diagnosed as minimally demented did not meet the DSM-IV criteria of dementia (i.e., in section 5.1 they were classified to the ‘nondemented’ group), it might still be interesting to examine which variables known from T₁ are sensitive to a classification of minimal dementia two years later, at T₂. As was described in section 4.14 of this chapter, several subtests showed decline of performance in the Mild/mod.D and the Min.D subjects, while the NonD-NC and the NonD-CI subjects did not decline. In addition, minimal dementia according to CAMDEX may progress to mild or moderate dementia within a few years. Thus, it might be argued that variables that are sufficiently sensitive to differentiate between nondemented subjects and minimally demented (and more severely demented) subjects are the best predictors of dementia – a diagnosis that will be confirmed in the near future. Therefore, the analyses performed in section 5.1 will be repeated in section 5.2, in which the minimally demented subjects are categorised to the ‘demented’ group. However, conclusions in this respect should be regarded tentatively because of the ambiguous nature of the minimal dementia classification (it

---

35 11 subjects were diagnosed as nondemented at T₂, while they were diagnosed as minimally demented at T₁. Because of possible diagnostic errors (either at T₁ or at T₂), these subjects were excluded, as was also done in section 4 of the present chapter. It must be noted that 5 of these 11 subjects were not tested at T₂ because of refusal or inability to participate. Diagnostic information was provided by the GP, which may only identify the more advanced stage demented subjects, in contrast to the CAMDEX procedure that may be more sensitive to early stage demented subjects (see also Chapter VI). Thus, the ‘nondemented’ status of these subjects is debatable.

36 19 subjects had missing data on one or more subtests and 11 subjects were excluded (see Footnote 35).
is not guaranteed that these subjects will develop clinical dementia in the next few years), and because of the limited number of minimally demented subjects.

5.1. The prediction of dementia, according to the DSM-IV criteria of dementia
In this section, the analyses will be performed according to the official DSM-IV criteria of dementia. Thus, the ‘demented’ group consists of subjects diagnosed as mildly or moderately demented at T₂, while they were classified as nondemented or minimally demented according to CAMDEX at T₁. The ‘nondemented’ group consists of subjects diagnosed as nondemented or minimally dementia at T₂, while they were classified as nondemented according to CAMDEX at T₁.

5.1.1. The prediction of dementia in the entire group of elderly subjects
Table 6 shows the results of several discriminant analyses performed over the T₁ scores per subtest and several subject-related variables (i.e., age, education, sex, MMSE and CES-D). On the one hand, all these variables or meaningful clusters of these variables are entered together, in order to examine the accuracy of these variables to classify the subjects to the nondemented or the demented group. In addition, the characteristics of this classification process are described by means of the proportion of ‘true positive’ classifications (i.e., subjects diagnosed as demented are indeed classified as being demented) and ‘true negative’ classifications (i.e., nondemented subjects are classified as being nondemented). Furthermore, a discrimination measure is calculated from these data (d'), using the ‘True Positive Rate’ and the ‘False Positive Rate’ 37. On the other hand, it is investigated, by means of a ‘stepwise’ analysis, how all these measures may be reduced in order to create a selection of variables that are best able to discriminate between the demented and the nondemented subjects. The classification has been performed with an equal probability (50%) for subjects being classified to the demented or the nondemented group, as well as with the prior probability being computed from the respective group size (see Table 6).

First, it may be noted that the subtests that led to (small) underestimation of T₁ scores when not excluding subjects that were diagnosed demented at T₂ (the ‘Preclinically Demented’ subjects; see section 3) are partly the same measures that came forward from the stepwise discriminant analysis (i.e., the ‘Visual Association Test’ and the ‘Two-alternatives word-recognition test’). However, the ‘priming’ measure of the ‘Perceptual identification task’ was not present in section 3. This may be explained by the relatively large variation of the priming score, leading to a low SD estimate.

In addition, it may be noted from the stepwise discriminant analysis that the T₁ MMSE score was not among the best discriminating variables between T₂ demented vs. nondemented cases. The analysis, using only the MMSE score, showed that only 79.5% of the original grouped subjects were correctly classified (according to classification method A), relative to 90.6% in the stepwise analysis including the ‘Visual Association Test’, the ‘Two-alternatives word-recognition test’, the ‘priming’ measure of the ‘Perceptual identification task’ and age (see also Figure 15a). Also note the large difference in d' between the two analyses (1.36 relative to 2.24; see the grey cells in Table 6 and see Figure 15b).

37 True Positive Rate: (True Positive Predictions) / (Number of subjects diagnosed as ‘demented’); False Positive Rate: (False Positive Predictions) / (Number of subjects diagnosed as ‘nondemented’);
Table 6: Prediction of dementia (T2 CAMDEX: mild or moderate dementia) within the entire group of subjects (NC and CI ss.; n=117).

<table>
<thead>
<tr>
<th>Variables</th>
<th>classification method*</th>
<th>accuracy of classification</th>
<th>proportion of 'true positives'</th>
<th>proportion of 'true negatives'</th>
<th>d'</th>
</tr>
</thead>
<tbody>
<tr>
<td>none (baseline – ‘blind’ classification based on chance)</td>
<td>A</td>
<td>50%</td>
<td>5/10</td>
<td>53.5/107</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>84.6%</td>
<td>1/10</td>
<td>98/107</td>
<td>0.12</td>
</tr>
<tr>
<td>all memory measures together(^{38})</td>
<td>A</td>
<td>93.2%</td>
<td>8/10</td>
<td>101/107</td>
<td>2.39</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>96.6%</td>
<td>6/10</td>
<td>107/107</td>
<td>2.58</td>
</tr>
<tr>
<td>all variables together(^{39})</td>
<td>A</td>
<td>93.2%</td>
<td>8/10</td>
<td>101/107</td>
<td>2.39</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>98.3%</td>
<td>8/10</td>
<td>107/107</td>
<td>3.16</td>
</tr>
<tr>
<td>memory measures, 'stepwise': Vis. Assoc. (Residual variance=.494),</td>
<td>A</td>
<td>89.7%</td>
<td>7/10</td>
<td>98/107</td>
<td>1.92</td>
</tr>
<tr>
<td>Two-alt. word-recogn. (.380) and Perc. ident.: ‘priming’ (.347)</td>
<td>B</td>
<td>94.0%</td>
<td>5/10</td>
<td>105/107</td>
<td>2.05</td>
</tr>
<tr>
<td>all variables, 'stepwise': Vis. Assoc. (.494), Two-alt. word-recogn.</td>
<td>A</td>
<td>90.6%</td>
<td>8/10</td>
<td>98/107</td>
<td>2.24</td>
</tr>
<tr>
<td>(.380), Perc. ident.: ‘priming’ (.347) and age (.323)</td>
<td>B</td>
<td>94.0%</td>
<td>5/10</td>
<td>105/107</td>
<td>2.05</td>
</tr>
<tr>
<td>‘clinical measures’ together: MMSE and Vis. Assoc.</td>
<td>A</td>
<td>86.3%</td>
<td>6/10</td>
<td>95/107</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>91.5%</td>
<td>5/10</td>
<td>102/107</td>
<td>1.64</td>
</tr>
<tr>
<td>T₁ MMSE score</td>
<td>A</td>
<td>79.5%</td>
<td>7/10</td>
<td>86/107 !</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>90.6%</td>
<td>0/10 !!</td>
<td>106/107</td>
<td>0.00</td>
</tr>
<tr>
<td>Visual Association Test</td>
<td>A</td>
<td>82.9%</td>
<td>7/10</td>
<td>90/107</td>
<td>1.52</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>91.5%</td>
<td>6/10</td>
<td>101/107</td>
<td>1.80</td>
</tr>
<tr>
<td>‘episodic memory measures’ together: Ten word list-lrn., Word-</td>
<td>A</td>
<td>83.8%</td>
<td>8/10</td>
<td>90/107</td>
<td>1.83</td>
</tr>
<tr>
<td>recogn., ‘Paired-assoc. lnr.’ and ‘Two-alt. word-recogn.(^{40})</td>
<td>B</td>
<td>92.3%</td>
<td>4/10</td>
<td>104/107</td>
<td>1.62</td>
</tr>
</tbody>
</table>

\(^*\)A=all groups equal probability (50%); B=prior probability computed from group size (‘nondemented’: 107/117→91.5%; ‘demented’: 10/117→8.5%)


\(^{39}\) All memory measures and the variables age, education, sex, T₁ MMSE score and T₁ CES-D score.

\(^{40}\) Though the ‘Visual Association Test’ may also be interpreted as an ‘episodic memory measure’, it was not included here since it was already included in the cluster of ‘clinical measures’. In addition, the factor analyses (Chapter III) questioned the role of this test as an episodic memory measure (see Figures 17 and 18 of Chapter III).
...the second administration of the memory test battery

**Figure 15a:** Prediction of dementia within the entire group of elderly subjects (NC and CI ss.; \( n = 117 \)) according to several clusters of variables entered into a discriminant analysis (presented in Table 6).

**Figure 15b:** Discrimination between demented and nondemented subjects (measured by \( d' \)) within the entire group of elderly subjects (NC and CI ss.; \( n = 117 \)) according to several clusters of variables entered into a discriminant analysis (presented in Table 6).
CHAPTER IV

It may be noted from classification method A in Table 6 that the T₁ MMSE score falsely classifies many subjects to the demented group (many ‘false positives’, resulting in a lower specificity: 80.4%). In addition, the sensitivity of the MMSE (70.0%) is lower than the sensitivity of the four best discriminating variables displayed in Table 6 (80.0%). These two characteristics of the MMSE explain the low ς² value (see Figure 15b).

Overall, when one considers a 50% prior probability of a subject being demented, the demented cases are most accurately detected by worse (T₁) performance on the ‘Visual Association Test’ (an average score of 3.20 (SD 1.75) vs. 5.33 (.97) for the nondemented subjects) and the ‘Two-alternative word-recognition test’ (7.00 (2.00) vs. 9.33 (1.06)), smaller repetition priming effects on the ‘Perceptual identification task’ (.15 (.82) vs. .70 (.77)), and a higher age (86.90 (3.38) vs. 78.80 (7.87)), two years before the diagnosis can be made (90.6% accuracy of classification and a ς² value of 2.24). However, when the frequency distribution of the demented and nondemented subjects is known (i.e., prior probability computed from group size), the accuracy of prediction is not improved by the addition of age as a discriminating variable (94.0% accuracy of classification and a ς² value of 2.05, which is slightly lower than in the analyses using a 50% prior probability; see also Figure 15b).

In sum, poor performance on, particularly, the ‘Visual Association Test’, the ‘Two-alternative word-recognition test’ and the ‘priming’ measure of the ‘Perceptual identification task’, in combination with a high age are key variables in predicting the demented cases within a heterogeneous sample of elderly subjects. This selection of variables leads to a prediction accuracy which is only slightly less than the accuracy by all variables together (see Table 6 and Figure 15a).

5.1.2. The prediction of dementia in the ‘Cognitively Impaired’ (CI) group
In this section, it will be investigated how and how well the memory test battery and other variables predict dementia within the group of subjects that were classified to the CI group at T₁ (i.e., the ‘dementia-at-risk’ group). The same analyses were performed over the CI subjects as the analyses described in the previous section. See Table 7 for the results.

It may be noted from Table 7 that different measures were selected by the ‘stepwise’ analysis in discriminating between demented and nondemented subjects within the CI group than within the entire group of subjects (see Table 6). Within the CI group, the ‘Paired-associate learning test’ was the most important variable, while within the entire group of subjects the ‘Visual Association Test’, the ‘Two-alternative word-recognition test’ and age were the best discriminating variables. The ‘priming’ measure of the ‘Perceptual identification task’ came forward as a discriminating variable within both the entire group of subjects and the CI group. Thus, worse performance at T₁ on the ‘Paired-associate learning test’ (4.22 (5.04) vs. 13.06 (5.29)) and smaller repetition priming effects in the ‘Perceptual identification task’ (.12 (.86) vs. 1.13 (.91)), predict dementia best, two years before the diagnosis was made. See

---

41 A (stepwise) logistic regression analysis showed the following best predicting variables: 1. Vis. Assoc., 2. Two-alt. word-recogn., 3. Perc. ident. ‘priming’, 4. Ten word list-Ⅵn., 5. Two-alt. word-recogn. ‘removed’, 6. age. This analysis (classification cut-off .50) showed high correlations in predicting subjects as demented/nondemented with the best discriminating memory measures according to the stepwise discriminant analysis (Vis. Assoc., Two-alt. word-recogn., Perc. ident. ‘priming’ and age: χ²=48.706, df=1, p=.000). Thus, another choice of analysis does not essentially change the results.
Figure 16a for a comparison of the prediction accuracy according to the various clusters of variables and see Figure 16b for the discrimination measure $d'$.  

**Figure 16a:** Prediction of dementia within the CI group of elderly subjects ($n=42$) according to several clusters of variables entered into a discriminant analysis (presented in Table 7).  

![Graph showing prediction accuracy](image)

**Figure 16b:** Discrimination between demented and nondemented subjects (measured by $d'$) within the CI group of elderly subjects ($n=42$) according to several clusters of variables entered into a discriminant analysis (presented in Table 7).  

![Graph showing discrimination](image)
Table 7: Prediction of dementia (i.e., T2 CAMDEX: mild or moderate dementia) within the CI group of elderly subjects (n=42).

<table>
<thead>
<tr>
<th>Variables</th>
<th>classification method*</th>
<th>accuracy of classification</th>
<th>proportion of 'true positives'</th>
<th>proportion of 'true negatives'</th>
<th>(d^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>none (baseline – ‘blind’ classification based on chance)</td>
<td>A</td>
<td>50%</td>
<td>4.5/9</td>
<td>16.5/33</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>66.7%</td>
<td>2/9</td>
<td>26/33</td>
<td>0.04</td>
</tr>
<tr>
<td>all memory measures together</td>
<td>A</td>
<td>90.5%</td>
<td>9/9</td>
<td>29/33</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>90.5%</td>
<td>6/9</td>
<td>32/33</td>
<td>2.32</td>
</tr>
<tr>
<td>all variables together</td>
<td>A</td>
<td>92.9%</td>
<td>9/9</td>
<td>30/33</td>
<td>3.66</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>95.2%</td>
<td>8/9</td>
<td>32/33</td>
<td>3.11</td>
</tr>
<tr>
<td>memory measures, 'stepwise': Paired-assoc. Irm. (Residual variance=.585) and Perc. ident.: ‘priming’ (.504)</td>
<td>A</td>
<td>78.6%</td>
<td>7/9</td>
<td>26/33</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>90.5%</td>
<td>6/9</td>
<td>32/33</td>
<td>2.32</td>
</tr>
<tr>
<td>all variables, 'stepwise': Paired-assoc. Irm. (.585) and Perc. ident.: ‘priming’ (.504)</td>
<td>A</td>
<td>78.6%</td>
<td>7/9</td>
<td>26/33</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>90.5%</td>
<td>6/9</td>
<td>32/33</td>
<td>2.32</td>
</tr>
<tr>
<td>‘clinical measures’ together: MMSE and Vis. Assoc.</td>
<td>A</td>
<td>71.4%</td>
<td>6/9 !</td>
<td>24/33</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>81.0%</td>
<td>4/9 !</td>
<td>30/33</td>
<td>1.19</td>
</tr>
<tr>
<td>T1 MMSE score</td>
<td>A</td>
<td>59.5% !</td>
<td>6/9 !</td>
<td>19/33 !</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>78.6%</td>
<td>0/9 !!</td>
<td>33/33</td>
<td>0.00 !</td>
</tr>
<tr>
<td>Visual Association Test</td>
<td>A</td>
<td>66.7%</td>
<td>6/9 !</td>
<td>22/33</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>81.0%</td>
<td>4/9 !</td>
<td>30/33</td>
<td>1.19</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>88.1%</td>
<td>5/9</td>
<td>32/33</td>
<td>2.03</td>
</tr>
<tr>
<td>Paired-associate learning test</td>
<td>A</td>
<td>85.7%</td>
<td>7/9</td>
<td>29/33</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>88.1%</td>
<td>6/9</td>
<td>31/33</td>
<td>1.99</td>
</tr>
</tbody>
</table>

*classification method:
A=all groups equal probability (50%)
B=prior probability computed from group size (non-demented group: 33/42=78.6%; demented group: 9/42=21.4%)
It may be noted from Figures 16a and 16b that the clinical measures have an evidently worse accuracy of classification, as well as a lower level of $d'$, than the two best discriminating variables (the ‘Paired-associate learning test’ and the ‘priming’ measure of the ‘Perceptual identification task’). Particularly the MMSE, using method A, leads to an only slightly above baseline accuracy (see Figure 16a). In addition, the $d'$ value for the MMSE when using method B is very low (0), caused by classifying all subjects to the ‘nondemented’ group. It may be noted that the sensitivity of each of the clinical measures to identify demented cases is inadequate (see the proportions of ‘true positives’ in Table 7). Specifically, the MMSE shows proportions of ‘true positives’ just above or even below baseline. In contrast, the ‘Paired-associate learning test’ shows good classification characteristics. Furthermore, the difference in prediction accuracy and discrimination value $d'$ between the two best discriminating variables and all variables together is only minor when using method B (4.7% and 0.79, respectively). However, the differences are evidently greater when using classification method A (14.3% and 2.08, respectively; compare the grey cells in Table 7).

In order to further examine the demented subjects’ characteristics of performance on the ‘Paired-associate learning test’ and the ‘priming’ measure of the ‘Perceptual identification task’, Figure 17 and 18 illustrate the patterns of performance across different trials and conditions in these two tasks (by the CI subjects – in the entire group of subjects, differences between the demented and nondemented subjects are even greater).

**Figure 17:** T₁ performance on the ‘Paired-associated learning test’, across the three trials and two conditions (semantic vs. non-semantic pairs), by the demented vs. the nondemented (CI) subjects.
CHAPTER IV

GLM Repeated Measures analyses found significant interaction effects between ‘trial’ and ‘demented’ for the semantic pairs ($F=4.941$, $df=2$, $p=.011$) and the non-semantic pairs ($F=3.326$, $df=2$, $p=.044$)\(^{42}\). As is illustrated in Figure 17, in both conditions the nondemented subjects improve their recall performance over trials, while the demented subjects show an almost flat learning curve (significant ‘trial’ by ‘demented’ interaction effect: $F=9.176$, $df=2$, $p=.000$). In addition, a significant interaction effect between ‘condition’ (semantic or non-semantic pairs) and ‘demented’ ($F=8.492$, $df=1$, $p=.005$) shows that the nondemented subjects benefit to a greater degree from the semantic relations within the material to be learned than the demented subjects.

Figure 18: T\(_1\) average identification times (in tics of 16 msec.) by the demented vs. the nondemented (CI) subjects in the ‘Perceptual identification task’: on the first (‘MF’) vs. the second presentation of the MF words (‘rep-MF’).

![Identification Times Graph](image)

GLM Repeated Measures analysis found a significant effect of within-subjects factor ‘condition’ (MF vs. rep-MF words; $F=12.239$, $df=1$, $p=.001$), but no significant effect of between-subjects factor ‘demented’ ($F=.016$, $df=1$, $p=.899$). In addition, a significant interaction effect between ‘condition’ and ‘demented’ was found ($F=8.008$, $df=1$, $p=.007$). Thus, as Figure 18 illustrates, the nondemented subjects benefited from the repetition of words, while the demented subjects did not\(^{43}\).

\(^{42}\) In both conditions, the sphericity assumption was not met (Mauchly’s Test of Sphericity: semantic pairs: $p=.019$, e (Greenhouse-Geisser)=.872; non-semantic pairs: $p=.019$, e (Greenhouse-Geisser)=.872). Adjusted tests showed significant interaction effects between ‘trial’ and ‘demented’ for both the semantic pairs ($F=6.970$, $df=1.744$, $p=.002$) and the non-semantic pairs ($F=4.486$, $df=1.744$, $p=.018$).

\(^{43}\) It may be noted that none of the priming effects that could be derived from the ‘Mirror-reading task’ (i.e., difference in reading times between conditions ‘2’ and ‘22’, ‘3’ and ‘32’, ‘32’ and ‘33’, or ‘3’ and ‘33’) showed significant differences between the demented and the nondemented subjects.
5.2. The role of the ‘minimal dementia’ subjects

Since the $T_2$ CAMDEX diagnosis of ‘minimal dementia’ may be considered as an intermediate stage of the disease, alternative stepwise discriminant analyses were performed in order to examine whether excluding these subjects or classifying them to the demented group (as was also done in section 4.14) led to different results (e.g., different discriminating variables). As was noted at the beginning of section 5, the ‘minimal dementia’ classification of the CAMDEX does not officially meet the DSM-IV criteria of dementia. Therefore, these analyses should be seen as an exploratory investigation of early predictors of dementia. See Table 8 for an overview of the results.

From Table 8, it may be noted that nearly the same variables were selected by the discriminant analyses, irrespective of where the minimal dementia subjects were classified. The only exception was the selection of the ‘Paired-associate learning test’, instead of age, in the analysis performed over the entire group of subjects in which the minimal dementia subjects were classified to the ‘demented’ group. This may be explained by the relatively poor $T_1$ performance of the minimal dementia subjects on this subtest, relative to the nondemented subjects (the NonD-NC and the NonD-CI group; see Figure 4). An important characteristic of the ‘Paired-associate learning test’ is that poor $T_1$ performance also occurs in the minimal dementia subjects, while at most other subtests, for example in the ‘Visual Association Test’, poor performance is above all found in the subjects diagnosed with mild or moderate dementia at $T_2$. This is also illustrated in Table 9 and corresponding Figure 19, which reflect average $z$-scores of all subtests per clinical subgroup (defined in section 4), in order to present an overview of the respective performance levels per subtest.

DISCUSSION

It may be concluded that the most important variables in predicting dementia within the CI group are the ‘Paired-associate learning test’ and the ‘priming’ measure of the ‘Perceptual identification task’. Thus, a verbal cued recall task (e.g., demanding benefit of semantic relations within material) together with repetition priming effects in identifying words are variables sufficiently sensitive to detect dementia within a dementia-at-risk group, two years before the clinical diagnosis could be made. In addition, age came forward as an important predictive variable, which is no surprise since the occurrence of dementia increases exponentially with increasing age (e.g., Ott et al., 1995).

The ‘Visual Association Test’ (also a cued recall task) and the ‘Two-alternative word-recognition test’ only seem useful regarding the prediction of dementia within a cognitively more heterogeneous group. Within the entire group, most nondemented subjects obtain (almost) the maximum score at these subtests, while the demented subjects obtain a relatively low score. Within the CI group, relatively low scores are more common, also in nondemented subjects. Note that $T_1$ performance on the ‘Visual Association Test’ by the Min.D group resembled the NonD-CI subjects’ performance (see section 4.12). Therefore, these rather ‘easy’ subtests are less useful in predicting demented cases in a more cognitively impaired group of subjects (i.e., in dementia-at-risk subjects). Only the Mild/mod.D subjects showed deficient performance at $T_1$. Thus, the ‘Visual Association Test’ and the ‘Two-alternative word-recognition test’ are not sufficiently sensitive to differentiate between demented and nondemented subjects within a more homogeneous group of subjects that are suspected to develop dementia.
Table 8: Results of the ‘stepwise’ discriminant analyses, performed over the T1 scores of each subtest and the subject-related variables, in discriminating between T2 ‘demented’ and ‘nondemented’ subjects. The classification of the T2 ‘minimal dementia’ subjects is manipulated\(^{44}\).

<table>
<thead>
<tr>
<th>discriminating between:</th>
<th>performed over:</th>
<th>entered subtests</th>
<th>Residual variance</th>
<th>Exact F</th>
<th>df</th>
<th>p</th>
<th>accuracy of classification</th>
<th>‘true positives’</th>
<th>‘true negatives’</th>
<th>d'</th>
</tr>
</thead>
<tbody>
<tr>
<td>nondemented/ minimal dementia</td>
<td>all ss. (n=117)</td>
<td>1. Visual Assoc. Test</td>
<td>.494</td>
<td>37.424</td>
<td>115</td>
<td>.000</td>
<td>90.6% (A)</td>
<td>8/10</td>
<td>98/107</td>
<td>2.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Two-alt. word-recogn.</td>
<td>.380</td>
<td>29.629</td>
<td>114</td>
<td>.000</td>
<td>94.0% (B)</td>
<td>5/10</td>
<td>105/107</td>
<td>2.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Perc. id.: priming</td>
<td>.347</td>
<td>22.589</td>
<td>113</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. age</td>
<td>.323</td>
<td>18.706</td>
<td>112</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vs. mild/moderate dementia</td>
<td>CI ss. (n=42)</td>
<td>1. Paired-assoc. Irm.</td>
<td>.585</td>
<td>20.098</td>
<td>40</td>
<td>.000</td>
<td>78.6% (A)</td>
<td>7/9</td>
<td>26/33</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Perc. id.: priming</td>
<td>.504</td>
<td>13.550</td>
<td>39</td>
<td>.000</td>
<td>90.5% (B)</td>
<td>6/9</td>
<td>32/33</td>
<td>2.32</td>
</tr>
<tr>
<td>nondemented</td>
<td>all ss. (n=112)</td>
<td>1. Visual Assoc. Test</td>
<td>.484</td>
<td>38.799</td>
<td>110</td>
<td>.000</td>
<td>91.1% (A)</td>
<td>8/10</td>
<td>94/102</td>
<td>2.24</td>
</tr>
<tr>
<td>vs. mild/moderate dementia</td>
<td>CI ss. (n=39)</td>
<td>1. Paired-assoc. Irm.</td>
<td>.550</td>
<td>22.618</td>
<td>37</td>
<td>.000</td>
<td>84.6% (A)</td>
<td>8/9</td>
<td>25/30</td>
<td>2.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Perc. id.: priming</td>
<td>.451</td>
<td>16.371</td>
<td>36</td>
<td>.000</td>
<td>89.7% (B)</td>
<td>7/9</td>
<td>28/30</td>
<td>2.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. age</td>
<td>.386</td>
<td>13.874</td>
<td>35</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{44}\) The 11 subjects that were diagnosed as minimally demented according to CAMDEX at T1, but were diagnosed as nondemented according to CAMDEX at T2, as well as the subject that was diagnosed as minimally demented according to CAMDEX at T1 and of whom no information was available at T2, were excluded (as was done in previous sections).
Table 9: Average z-scores per subtest administered at T₁, for each clinical subgroup (standard error within parentheses)\(^44\). Z-scores are based on the mean scores and standard deviations of the group of subjects that were classified as ‘nondemented’ according to CAMDEX or GP information at T₂ (the NonD-NC and NonD-CI group; see second column).

<table>
<thead>
<tr>
<th>Subtest:</th>
<th>Norms NonD: average (sd)</th>
<th>z-score NonD-NC</th>
<th>z-score NonD-CI</th>
<th>z-score Min.D</th>
<th>z-score Mild/mod.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ten words list-learning test</td>
<td>12.79 (4.66)</td>
<td>.20 (.11)</td>
<td>-39 (.14)</td>
<td>-78 (.17)</td>
<td>-1.39 (.18)</td>
</tr>
<tr>
<td></td>
<td>n=119</td>
<td></td>
<td></td>
<td>n=40</td>
<td>n=6</td>
</tr>
<tr>
<td>Word-recognition test</td>
<td>18.20 (1.70)</td>
<td>.13 (.11)</td>
<td>-25 (.16)</td>
<td>-71 (.43)</td>
<td>-1.71 (.53)</td>
</tr>
<tr>
<td></td>
<td>n=118</td>
<td></td>
<td></td>
<td>n=39</td>
<td>n=6</td>
</tr>
<tr>
<td>Two-alternative word-recognition test</td>
<td>9.31 (1.08)</td>
<td>.08 (.10)</td>
<td>-14 (.18)</td>
<td>-44 (.37)</td>
<td>-2.14 (.59)</td>
</tr>
<tr>
<td></td>
<td>n=119</td>
<td></td>
<td></td>
<td>n=40</td>
<td>n=6</td>
</tr>
<tr>
<td>Visual Association Test</td>
<td>5.36 (.92)</td>
<td>.15 (.09)</td>
<td>-.28 (.21)</td>
<td>-.21 (.59)</td>
<td>-2.35 (.60)</td>
</tr>
<tr>
<td></td>
<td>n=119</td>
<td></td>
<td></td>
<td>n=40</td>
<td>n=6</td>
</tr>
<tr>
<td>Paired-associate learning test</td>
<td>16.71 (5.67)</td>
<td>.25 (.11)</td>
<td>-.49 (.14)</td>
<td>-.98 (.33)</td>
<td>-1.89 (.41)</td>
</tr>
<tr>
<td></td>
<td>n=118</td>
<td></td>
<td></td>
<td>n=40</td>
<td>n=6</td>
</tr>
<tr>
<td>Category fluency test</td>
<td>25.31 (6.52)</td>
<td>.17 (.12)</td>
<td>-.34 (.13)</td>
<td>-.66 (.35)</td>
<td>-.89 (.33)</td>
</tr>
<tr>
<td></td>
<td>n=119</td>
<td></td>
<td></td>
<td>n=40</td>
<td>n=6</td>
</tr>
<tr>
<td>Perceptual identification task: ‘semantic memory’(^45)</td>
<td>13.44 (2.33)</td>
<td>.30 (.07)</td>
<td>-.62 (.21)</td>
<td>-.11 (.25)</td>
<td>-.35 (.19)</td>
</tr>
<tr>
<td></td>
<td>n=116</td>
<td></td>
<td></td>
<td>n=38</td>
<td>n=6</td>
</tr>
<tr>
<td>Perceptual identification task: ‘priming’</td>
<td>.72 (.81)</td>
<td>-.25 (.09)</td>
<td>.54 (.20)</td>
<td>-.07 (.38)</td>
<td>-.71 (.32)</td>
</tr>
<tr>
<td></td>
<td>n=114</td>
<td></td>
<td></td>
<td>n=37</td>
<td>n=6</td>
</tr>
<tr>
<td>Mirror-reading task(^9)</td>
<td>14.70 (2.91)</td>
<td>.36 (.08)</td>
<td>-.84 (.19)</td>
<td>-.08 (.37)</td>
<td>-.35 (.56)</td>
</tr>
<tr>
<td></td>
<td>n=110</td>
<td></td>
<td></td>
<td>n=33</td>
<td>n=5</td>
</tr>
<tr>
<td>Block span task</td>
<td>3.87 (.94)</td>
<td>.14 (.12)</td>
<td>-.26 (.14)</td>
<td>-.39 (.24)</td>
<td>-.39 (.29)</td>
</tr>
<tr>
<td></td>
<td>n=119</td>
<td></td>
<td></td>
<td>n=40</td>
<td>n=6</td>
</tr>
<tr>
<td>Digit span task</td>
<td>4.41 (.97)</td>
<td>.20 (.11)</td>
<td>-.40 (.14)</td>
<td>-.08 (.43)</td>
<td>.51 (.19)</td>
</tr>
<tr>
<td></td>
<td>n=118</td>
<td></td>
<td></td>
<td>n=40</td>
<td>n=6</td>
</tr>
<tr>
<td>Mini-Mental State Examination</td>
<td>26.49 (2.53)</td>
<td>.61 (.05)</td>
<td>-1.21 (.10)</td>
<td>-5.59 (.27)</td>
<td>-1.26 (.24)</td>
</tr>
<tr>
<td></td>
<td>n=119</td>
<td></td>
<td></td>
<td>n=40</td>
<td>n=6</td>
</tr>
</tbody>
</table>

\(^45\) Tasks measuring reaction times: transformed for reasons of comparison (i.e., high scores indicating fast reaction times in this table).
Figure 19: Average z-scores per clinical subgroup of the best discriminating subtests, with respect to dementia (see Table 8), administered at T1. The z-scores of the MMSE, known from T1, are also displayed, as a reference point of global cognitive status.
Furthermore, the MMSE was found even less valuable than other variables since it was not among the discriminating variables according to any stepwise discriminant analysis. On its own, the MMSE classified too many subjects incorrectly to the demented group, even within the large heterogeneous group of elderly subjects. Nonetheless, the MMSE does seem useful in functioning as a first and global screening instrument in order to identify cognitively at-risk subjects (high sensitivity when considering method A), for which it indeed was designed. However, the ‘Visual Association Test’ was found to be similarly useful as a first screening instrument (equally sensitive) and even led to a more specific selection (Table 6).

In sum, other measures than the variables commonly used in clinical practice to assess dementia were found to be most accurate in predicting dementia, two years before the diagnosis could be made according the DSM-IV definition of dementia — particularly within a group of dementia-at-risk subjects. It may be concluded from the data presented here that the following three features are the best indication of early dementia: little benefit from semantic relations in a cued recall task (and hardly any improvement when words are repeatedly presented), absent implicit remembering of words presented previously, and a high age.

6. The profile of memory measures best discriminating between demented and nondemented elderly subjects at the time of diagnosis

Section 5 focused on the prediction of dementia before the diagnosis can be made and, thus, provided information of the memory profile in preclinical dementia. Though this information is most important in the present research, it is also interesting to examine which measures of the memory test battery discriminate best between demented and nondemented subjects once the diagnosis has been made (i.e., a ‘cross-sectional’ analysis). While the results of the analyses performed in the previous section over the T₁ memory subtest data provide information on preclinical predictors of dementia, the analyses that will be described in the current section identify the memory measures that are most useful in differential assessment once dementia can be diagnosed. For this reason, the analyses that were performed in section 5 will be repeated in this section, using the T₂ memory subtest data. In the ‘stepwise’ discriminant analyses, presented in Table 10, the classification of the ‘minimal dementia’ subjects is manipulated again: classifying them to the ‘nondemented’ group (consistent with the DSM-IV criteria), excluding them, and classifying them to the ‘demented’ group. In addition, Table 11 and corresponding Figure 20 present an overview of the performance level of the various subtests for each clinical subgroup separately, by means of z-scores.
Table 10: Results of the 'stepwise' discriminant analyses, performed over the T2 scores of each subtest (except for the ‘Mirror-reading task’ and both measures of the ‘Perceptual identification task’ because of missing data) and the subject-related variables (except for the MMSE because of interference with group classification), in discriminating between T2 ‘demented’ and ‘nondemented’ subjects. The classification of the T2 ‘minimal dementia’ subjects is manipulated.

<table>
<thead>
<tr>
<th>discriminating between:</th>
<th>performed over:</th>
<th>entered subtests</th>
<th>Residual variance</th>
<th>Exact $F$</th>
<th>$df$</th>
<th>$p$</th>
<th>accuracy of classification</th>
<th>‘true positives’</th>
<th>‘true negatives’</th>
<th>$d'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>nondemented/ minimal dementia vs. mild/moderate dementia</td>
<td>all ss. (n=84)</td>
<td>1. Visual Assoc. Test</td>
<td>.329</td>
<td>45.462</td>
<td>82</td>
<td>.000</td>
<td>95.2% (A)</td>
<td>5/6</td>
<td>75/78</td>
<td>2.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Ten word list-Im.</td>
<td>.263</td>
<td>30.811</td>
<td>81</td>
<td>.000</td>
<td>95.2% (B)</td>
<td>4/6</td>
<td>76/78</td>
<td>2.32</td>
</tr>
<tr>
<td></td>
<td>CI ss. (n=28)</td>
<td>1. Category fluency</td>
<td>.343</td>
<td>31.457</td>
<td>26</td>
<td>.000</td>
<td>96.4% (A)</td>
<td>5/5</td>
<td>22/23</td>
<td>4.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Word-recogn. test</td>
<td>.251</td>
<td>23.590</td>
<td>25</td>
<td>.000</td>
<td>92.9% (B)</td>
<td>4/5</td>
<td>22/23</td>
<td>2.59</td>
</tr>
<tr>
<td>nondemented vs. mild/moderate dementia</td>
<td>all ss. (n=79)</td>
<td>1. Visual Assoc. Test</td>
<td>.237</td>
<td>71.292</td>
<td>77</td>
<td>.000</td>
<td>97.5% (A)</td>
<td>5/6</td>
<td>72/73</td>
<td>3.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Ten word list-Im.</td>
<td>.173</td>
<td>52.233</td>
<td>76</td>
<td>.000</td>
<td>97.5% (B)</td>
<td>4/6</td>
<td>73/73</td>
<td>2.76</td>
</tr>
<tr>
<td></td>
<td>CI ss. (n=25)</td>
<td>1. Word-recogn. test</td>
<td>.295</td>
<td>38.240</td>
<td>23</td>
<td>.000</td>
<td>100.0% (A)</td>
<td>5/5</td>
<td>20/20</td>
<td>4.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Category fluency</td>
<td>.162</td>
<td>39.530</td>
<td>22</td>
<td>.000</td>
<td>100.0% (B)</td>
<td>5/5</td>
<td>20/20</td>
<td>4.64</td>
</tr>
<tr>
<td>nondemented vs. minimal/mild/ moderate dementia</td>
<td>all ss. (n=84)</td>
<td>1. Ten word list-Im.</td>
<td>.412</td>
<td>54.514</td>
<td>82</td>
<td>.000</td>
<td>95.2% (A)</td>
<td>9/11</td>
<td>71/73</td>
<td>2.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Visual Assoc. Test</td>
<td>.294</td>
<td>45.314</td>
<td>81</td>
<td>.000</td>
<td>95.2% (B)</td>
<td>7/11</td>
<td>73/73</td>
<td>2.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Two-alt. word-recogn.</td>
<td>.273</td>
<td>33.173</td>
<td>80</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. CES-D*</td>
<td>.254</td>
<td>26.991</td>
<td>79</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CI ss. (n=28)</td>
<td>1. Ten word list-Im.</td>
<td>.383</td>
<td>36.836</td>
<td>26</td>
<td>.000</td>
<td>96.4% (A)</td>
<td>7/8</td>
<td>20/20</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Visual Assoc. Test</td>
<td>.258</td>
<td>31.673</td>
<td>25</td>
<td>.000</td>
<td>96.4% (B)</td>
<td>7/8</td>
<td>20/20</td>
<td>3.50</td>
</tr>
</tbody>
</table>

*more depressive symptoms in nondemented ss. (average score 6.79) than in demented ss. (6.45).
Table 11: Average z-scores per subtest administered at T₂, for each clinical subgroup (standard error within parentheses)\(^{44}\). Z-scores are based on the mean scores and standard deviations of the group of subjects that were classified as ‘nondemented’ according to CAMDEX or GP information at T₂ (the NonD-NC and NonD-CI group; see second column).

<table>
<thead>
<tr>
<th>Subtest:</th>
<th>Norms NonD: average (sd)</th>
<th>z-score NonD-NC</th>
<th>z-score NonD-CI</th>
<th>z-score Min.D</th>
<th>z-score Mild/mod.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ten words list-learning test</td>
<td>14.44 (5.32)</td>
<td>.24 (.13)</td>
<td>-1.46 (.39)</td>
<td>-2.50 (.10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=84</td>
<td>n=56</td>
<td>n=28</td>
<td>n=6</td>
<td>n=7</td>
</tr>
<tr>
<td>Word-recognition test</td>
<td>18.45 (1.70)</td>
<td>.15 (.13)</td>
<td>-1.44 (.50)</td>
<td>-2.87 (.68)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=84</td>
<td>n=56</td>
<td>n=28</td>
<td>n=6</td>
<td>n=7</td>
</tr>
<tr>
<td>Two-alternative word-recognition test</td>
<td>9.55 (.86)</td>
<td>.07 (.11)</td>
<td>-1.14 (.23)</td>
<td>-1.80 (.85)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=84</td>
<td>n=56</td>
<td>n=28</td>
<td>n=6</td>
<td>n=6</td>
</tr>
<tr>
<td>Visual Association Test</td>
<td>5.54 (.74)</td>
<td>.09 (.12)</td>
<td>-1.00 (.35)</td>
<td>-4.98 (1.20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=84</td>
<td>n=56</td>
<td>n=28</td>
<td>n=6</td>
<td>n=7</td>
</tr>
<tr>
<td>Paired-associate learning test</td>
<td>19.21 (6.17)</td>
<td>.26 (.12)</td>
<td>-1.57 (.47)</td>
<td>-2.22 (.65)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=84</td>
<td>n=56</td>
<td>n=28</td>
<td>n=6</td>
<td>n=6</td>
</tr>
<tr>
<td>Category fluency test</td>
<td>26.58 (6.88)</td>
<td>.17 (.15)</td>
<td>-1.88 (.15)</td>
<td>-2.24 (.32)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=84</td>
<td>n=56</td>
<td>n=28</td>
<td>n=6</td>
<td>n=7</td>
</tr>
<tr>
<td>Perceptual identification task: ‘semantic memory’(^{45})</td>
<td>5.72 (1.77)</td>
<td>.30 (.10)</td>
<td>-1.30 (.36)</td>
<td>-1.35 (.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=80</td>
<td>n=55</td>
<td>n=25</td>
<td>n=6</td>
<td>n=4</td>
</tr>
<tr>
<td>Perceptual identification task: ‘priming’</td>
<td>.77 (.71)</td>
<td>.01 (.13)</td>
<td>-.42 (.39)</td>
<td>.79 (.72)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=80</td>
<td>n=55</td>
<td>n=25</td>
<td>n=6</td>
<td>n=4</td>
</tr>
<tr>
<td>Mirror-reading task</td>
<td>21.73 (5.58)</td>
<td>.25 (.10)</td>
<td>-.14 (.16)</td>
<td>.05 (.44)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=72</td>
<td>n=54</td>
<td>n=18</td>
<td>n=4</td>
<td>n=3</td>
</tr>
<tr>
<td>Block span task</td>
<td>4.15 (.92)</td>
<td>.13 (.13)</td>
<td>-.16 (.28)</td>
<td>-.34 (.18)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=84</td>
<td>n=56</td>
<td>n=28</td>
<td>n=6</td>
<td>n=6</td>
</tr>
<tr>
<td>Digit span task</td>
<td>4.74 (.97)</td>
<td>.27 (.13)</td>
<td>-.111 (.34)</td>
<td>-.91 (.61)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=84</td>
<td>n=56</td>
<td>n=28</td>
<td>n=6</td>
<td>n=7</td>
</tr>
<tr>
<td>Mini-Mental State Examination</td>
<td>27.06 (2.29)</td>
<td>.43 (.08)</td>
<td>-.192 (.46)</td>
<td>-3.77 (.80)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=84</td>
<td>n=56</td>
<td>n=28</td>
<td>n=6</td>
<td>n=7</td>
</tr>
</tbody>
</table>
Figure 20: Average z-scores per clinical subgroup of the best discriminating subtests with respect to dementia (see Table 10), administered at T2.
DISCUSSION  It may be noted from Table 10 that the ‘Ten word list-learning test’ and the ‘Visual Association Test’ discriminated best within the entire group of subjects, regardless of where the minimally demented subjects were classified. These two measures were also best at discriminating between the nondemented subjects and the minimally or more severely demented subjects within the CI group. However, when the ‘demented’ group within the CI group consisted of mildly or moderately demented subjects (the minimally demented subjects were excluded or were classified to the ‘nondemented’ group), the ‘Category fluency test’ and the ‘Word-recognition test’ were the best discriminating variables. Furthermore, it must be noted that the discriminative value of these measures at T2 largely corresponds with the measures that showed greatest decline of performance from T1 to T2 in incident demented cases, relative to nondemented subjects (see section 4.14, Table 5). In addition, the $d'$ values of the measures discriminating between the ‘demented’ and the ‘nondemented’ group within the CI group (the ‘dementia-at-risk’ group) are remarkably high (i.e., particularly the ‘Category fluency test’ and the ‘Word-recognition test’ discriminate very well: $d'$ values > 4).

7. Identification and characterization of nondemented but memory-declined subjects

In this final section, an attempt will be made to identify the subjects that showed (greater than average) decline on the memory test battery from T1 to T2, but were not officially diagnosed, according to DSM-IV criteria, as demented at T2 (i.e., all subjects that were tested at T1 and T2 and were classified as nondemented or minimally demented according to CAMDEX at T2). This group of subjects will be called the ‘Memory Declined’ (MD) group. Table 12 presents, per subtest, the average difference score $d(T2 - T1)$. For each subtest, the subjects can be detected that showed a greater than average reduction in performance, according to this equation: $d \leq (M_d - s_d)$ (if $d<0$, is equal or smaller than the mean $d$ score minus the $d$ standard deviation)$^{46}$. Table 12 presents, per subtest, the number of (nondemented) subjects that showed a greater than average reduction in memory performance (the MD cases) vs. the number of subjects that demonstrated relatively stable (or even improved) performance from T1 to T2 (the ‘Memory Nondeclined’ (MND) group, subjects whose $d$ score did not satisfy the equation described above).

---

$^{46}$ The two exceptions are the ‘Mirror-reading task’ and the ‘semantic memory’ measure of the ‘Perceptual identification task’, since here a higher score indicates worse performance (i.e., slower reaction times). For these measures, the equation is: $d \geq (M_d + s_d)$ (if $d>0$).
Table 12: Mean difference score ($d$) per subtest, for the MD group (greater than average decline from $T_1$ to $T_2$) and the MND group (stable or improved performance from $T_1$ to $T_2$) separately.

<table>
<thead>
<tr>
<th>Subtest:</th>
<th>mean $d$ ($T_2$-$T_1$) (sd)</th>
<th>cut-off: $M_d - s_d$</th>
<th>MD group: $n$; mean $d$ (SD)</th>
<th>MND group: $n$; mean $d$ (SD)</th>
<th>% of ss. declined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mini-Mental State Examination</td>
<td>.22 (2.18)</td>
<td>-1.96</td>
<td>$n=15$; -3.07 (1.39)</td>
<td>$n=81$; .83 (1.70)</td>
<td>15.6%</td>
</tr>
<tr>
<td>Ten word list-learning test</td>
<td>.32 (3.61)</td>
<td>-3.29</td>
<td>$n=13$; -5.23 (1.36)</td>
<td>$n=83$; 1.19 (3.03)</td>
<td>13.5%</td>
</tr>
<tr>
<td>Digit span task</td>
<td>.17 (.99)</td>
<td>-.82</td>
<td>$n=21$; -1.14 (.36)</td>
<td>$n=74$; .54 (.76)</td>
<td>22.1%</td>
</tr>
<tr>
<td>Word-recognition test</td>
<td>.00 (1.78)</td>
<td>-1.78</td>
<td>$n=17$; -2.59 (.80)</td>
<td>$n=79$; .56 (1.40)</td>
<td>17.7%</td>
</tr>
<tr>
<td>Paired-associate learning test</td>
<td>.85 (4.17)</td>
<td>-3.32</td>
<td>$n=13$; -5.77 (1.54)</td>
<td>$n=82$; 1.90 (3.42)</td>
<td>13.7%</td>
</tr>
<tr>
<td>Block span task</td>
<td>.21 (.87)</td>
<td>-.66</td>
<td>$n=15$; -1.20 (.41)</td>
<td>$n=81$; .47 (1.65)</td>
<td>15.6%</td>
</tr>
<tr>
<td>Word stem completion task</td>
<td>.17 (.96)</td>
<td>-.79</td>
<td>$n=20$; -1.15 (.37)</td>
<td>$n=76$; .51 (.74)</td>
<td>20.8%</td>
</tr>
<tr>
<td>Category fluency test</td>
<td>-.11 (5.47)</td>
<td>-5.58</td>
<td>$n=16$; -8.44 (2.53)</td>
<td>$n=80$; 1.55 (4.24)</td>
<td>16.7%</td>
</tr>
<tr>
<td>Mirror-reading task</td>
<td>1.57 (3.65)</td>
<td>+5.22</td>
<td>$n=8$; 10.76 (3.65)</td>
<td>$n=69$; .50 (1.56)</td>
<td>10.4%</td>
</tr>
<tr>
<td>Perceptual identification task: 'semantic memory'</td>
<td>-.17 (1.03)</td>
<td>+.86</td>
<td>$n=7$; 1.98 (.76)</td>
<td>$n=81$; -.36 (.82)</td>
<td>8.0%</td>
</tr>
<tr>
<td>Perceptual identification task: 'priming'</td>
<td>.05 (.90)</td>
<td>-.85</td>
<td>$n=12$; -1.39 (.94)</td>
<td>$n=76$; .28 (.65)</td>
<td>13.6%</td>
</tr>
<tr>
<td>Two-alternative word-recognition test</td>
<td>-.07 (.89)</td>
<td>-.96</td>
<td>$n=20$; -1.35 (.67)</td>
<td>$n=75$; .27 (.58)</td>
<td>21.1%</td>
</tr>
<tr>
<td>Visual Association Test</td>
<td>-.07 (1.03)</td>
<td>-1.10</td>
<td>$n=4$; -3.25 (1.50)</td>
<td>$n=92$; .07 (.75)</td>
<td>4.2%</td>
</tr>
</tbody>
</table>

From Table 12, it is computed that per subtest, on average, 14.8% of the available nondemented subjects declined according to the criteria of decline described above. Eleven subjects declined on zero subtests, 26 subjects declined on one subtest, 15 subjects declined on two subtests, 8 subjects declined on three subtests, 5 subjects declined on four subtests, 3 subjects declined on five subtests, 2 subjects declined on six subtests, and 1 subject declined on eight subtests (25 subjects were not considered here because they had missing values on one or more subtests). Thus, 19 subjects showed decline on at least three subtests, 11 subjects on at least four subtests, and 6 subjects on at least five subtests.

Nonetheless, it is difficult to determine the minimum number of subtests that should indicate a ‘declined performance’ over the entire test battery: three, four or five subtests. Therefore, the baseline number of subjects was determined that theoretically could decline, based on chance alone (i.e., a subject had, per subtest, a prior probability of being classified to the ‘Memory Declined’ group of 14.8%). This investigation showed that 21 subjects would decline on at least three subtests, 8 subjects on at least four subtests, and 2 subjects on at least five subtests. These baseline figures hardly differ from the number of subjects that actually declined on three, four and five subtests, according the criteria of decline proposed above. Therefore, it does not seem meaningful to investigate the nature of decline and the predictive value with respect to decline of the ‘Memory Declined’ subjects any further.

47 ‘Mirror-reading task’ and ‘Perceptual identification task: semantic memory’: $M_d + s_d$ (if $d>0$).
The only examination that was made was whether the subjects that declined on the memory test battery (indicated by decline at a minimum of three or four subtests) were the same subjects that were 'falsely' classified to the 'demented' group (mild or moderate dementia) according to the four best predicting variables indicated by the stepwise discriminant analysis performed in section 5.1.1 (i.e., the 'Visual Association Test', the 'Two-alternative word-recognition test', the 'priming' measure of the 'Perceptual identification task' and age; classification method A). Tables 13a and 13b illustrate the relation between the degree of decline on the battery from T1 to T2 ('declined' or 'nondeclined') and the predicted group according to the stepwise discriminant analysis ('demented': mild or moderate dementia; or 'non-demented': no or minimal dementia). A Pearson Chi-Square Test demonstrated a significant effect between these two variables, both in the three-subtests-cut-off ($\chi^2=5.632$, df=1, $p=.018$) and in the four-subtests-cut-off ($\chi^2=11.225$, df=1, $p=.001$). The five-subtests-cut-off was not tested because of the limited number of subjects classified to the MD group according to decline on at least five subtests.

Table 13a: Distribution of subjects according to decline on at least three subtests and predicted group (mild/moderate dementia vs. no/minimal dementia) according to the T1 scores on the 'Visual Association Test', the 'Two-alternative word-recognition test', the 'priming' measure of the 'Perceptual identification task' and age (classification with all groups having equal probability)

<table>
<thead>
<tr>
<th>MD/MND classification:</th>
<th>Predicted group:</th>
<th>total:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nondemented</td>
<td>demented</td>
</tr>
<tr>
<td>nondeclined</td>
<td>52 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>declined</td>
<td>17 (89.5%)</td>
<td>2 (10.5%)</td>
</tr>
<tr>
<td>total:</td>
<td>69</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 13b: Distribution of subjects according to decline on at least four subtests and predicted group (mild/moderate dementia vs. no/minimal dementia) according to the T1 scores on the 'Visual Association Test', the 'Two-alternative word-recognition test', the 'priming' measure of the 'Perceptual identification task' and age (classification with all groups having equal probability)

<table>
<thead>
<tr>
<th>MD/MND classification:</th>
<th>Predicted group:</th>
<th>total:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nondemented</td>
<td>demented</td>
</tr>
<tr>
<td>nondeclined</td>
<td>60 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>declined</td>
<td>9 (81.8%)</td>
<td>2 (18.2%)</td>
</tr>
<tr>
<td>total:</td>
<td>69</td>
<td>2</td>
</tr>
</tbody>
</table>

DISCUSSION It may be noted that the effects described above are rather weak, which may be partly due to the limited number of subjects that were predicted to be demented or were classified to the MD group. Whether the two subjects that were predicted demented by the best discriminating variables and showed greater than

---

48 Note that the 'nondemented' subjects that are investigated in this section also include the subjects that received a minimal dementia classification.
average decline on at least three or four subtests (see Tables 13a and 13b) will indeed develop dementia in the near future remains to be seen. However, these subjects may certainly be considered as being at risk for becoming demented. It may be noted that one of these two subjects received a minimal dementia classification at T2, while the other subject received a minimal dementia classification at T1 (and was classified as nondemented according to the CAMDEX at T2). The 17 subjects that were predicted to be nondemented by the discriminant analysis, but were classified to the MD group according to decline on at least three subtests, included 3 subjects that once received a minimal dementia classification (either at T1 or at T2), 4 NonD-CI subjects and 10 NonD-NC subjects.

8. Conclusion

The most important finding presented in the present chapter, is the strong predictive value regarding the ‘preclinical’ assessment of dementia of, first, the ‘Paired-associate learning test’ and, secondly, the ‘priming’ measure of the ‘Perceptual identification task’. Within a rather homogeneous group of subjects with cognitive deficits (which makes them an ‘at-risk group’ for developing dementia), these measures show high accuracy of prediction, approximately two years before the diagnosis of dementia was made. The predictive value of these measures was found, irrespective of where subjects that received a ‘minimal dementia’ diagnosis were classified. However, these two measures were found to be of specific good value when minimal dementia cases were classified to the ‘demented’ group. In other words, the ‘Paired-associate learning test’ and the ‘priming’ measure of the ‘Perceptual identification task’ may be particularly sensitive to detect dementia cases who are, two years later, still in a very early stage of their disease and do not yet show an evident decline in daily functioning as a result of their disease. This is in contrast with, particularly, the ‘Visual Association Test’ and the ‘Two-alternatives word-recognition test’. These two measures are mainly sensitive to predict demented cases that are, two years later, already in a mild or more advanced stage of their disease. In addition, these two measures only prove to be of good predictive value within a cognitively heterogeneous group of elderly subjects – within a dementia-at-risk group they lose their predictive ability. Therefore, it may be argued that other measures than the variables commonly used in clinical practice to assess dementia (e.g., MMSE, Visual Association Test or purely episodic memory tests) were found to be most accurate in predicting dementia, two years before the diagnosis could be made. It may be concluded that the following memory components showed deficits in ‘preclinically’ demented elderly subjects (at T1, two years before diagnosis), relative to elderly subjects that did not develop dementia two years later: little benefit from semantic relations in a cued recall task (and hardly any improvement when words are repeatedly presented), absent implicit remembering of words presented previously, and a high age.

Section 4 showed that other memory components than the ones described above showed the greatest degree of decline of performance in the two-year time period in the incident dementia cases, compared with the nondemented subjects: the ‘Ten word list-learning test’ and the ‘Category fluency test’. In other words, when a patient or client that has developed dementia in the mean time is tested in a follow-up examination two years later, decline on these measures should be most evident. Thus,
...the second administration of the memory test battery
decline on a verbal free recall task and a task of retrieval speed from semantic
memory are most obvious once subjects become demented. Nonetheless, these
memory components do not show real deficits before the diagnosis can be made. On
the other hand, decline on the ‘Paired-associate learning test’ is less obvious at
follow-up since performance on this task was already close to 0 at baseline. Finally, it
may be noted that subjects that did not develop dementia generally did not show
decline of memory performance in the two-year time period at all.

From section 6, it may be concluded that different measures discriminated best
between demented and nondemented subjects when using the T_1 data, compared to
using the T_2 data. The only exception was the ‘Visual Association Test’, which was a
useful measure in both situations (i.e., two years before the diagnosis was made, as
well as simultaneously with the time of diagnosis). In addition, in both situations, the
results hardly changed by a different classification of the minimally demented
subjects. This may be explained by the limited number of subjects that were classified
as minimally demented (n=6). The only difference was the importance of the ‘Paired-
associate learning test’ administered at T_1, and the ‘Two-alternative word-recognition
test’ test’ administered at T_2, when discriminating between nondemented subjects and
minimally or more severely demented subjects. More substantial differences occurred
when the analyses were focused on the CI group. These analyses showed the
importance of the ‘Paired-associate learning test’ and the ‘priming’ measure of the
‘Perceptual identification task’ when using the T_1 data, and the ‘Category fluency test’
and the ‘Word-recognition test’ when using the T_2 data.

The evident difference between the ‘preclinical’ dementia subjects and the
nondemented subjects in benefit from semantic relations within the material to be
learned is an issue frequently investigated in studies testing AD patients (see Chapter I
for a review). However, most of these studies use subjects that are at least mildly
demented, rather than in a ‘preclinical’ stage of dementia. Furthermore, these subjects
are many years younger and have a higher level of education than in the current study.
These factors enlarge the difference between pathological and normal ageing
processes in these studies and, thus, lead more easily to significant differences
between patients and controls. Therefore, the strong predictive value of the ‘Paired-
associate learning test’ in the current study – with much smaller differences
(concerning effects of age, education and severity of disease) between the two groups
– may be regarded as an important finding for the early assessment of dementia. It
may be concluded that poor semantic encoding of to be learned information, as was
found for AD patients relative to normal elderly subjects (e.g., Chertkow & Bub,
1990; Russo & Spinnler, 1994; Monti et al., 1996), is also detectable when subjects
are still officially nondemented. Sailor, Bramwell and Griesing (1998) suggested that
AD patients have a specific deficit in the ability to evaluate semantic relations. They
are no longer able to discriminate between two related concepts, because the attribute
knowledge that distinguishes these two concepts is lost. This explanation may also
characterise performance by the preclinical dementia subjects in the current study.
Some subjects noted that the target word ‘had something to do with’ the cue, but
somehow they could not name the correct word. Some subjects named semantically
related intrusions (see also Chapter III, section 4.4.3), but they repeatedly did not
succeed in naming the correct word. In contrast with the preclinically demented
subjects, the nondemented subjects benefited normally from the semantic relation
between the pairs and they showed a normal learning curve over trials. The ‘Paired-
associate learning test’ was easier for these normal subjects than, particularly, the
"Ten word list-learning test", which corresponds with literature findings – ageing effects are found to be greater in free recall tasks than in cued recall or recognition formats (see Chapter I). It may be concluded that, as opposed to cued recall tasks demanding semantic processing, normal elderly subjects show an impaired performance on free recall of semantically unrelated words, which results in a relatively small performance difference with preclinically demented subjects. Not until dementia has progressed, the performance difference with normal elderly subjects on free recall tasks may increase and become more significant (as was also described in section 6).

However, the predictive value of the ‘priming’ measure of the ‘Perceptual identification task’ should be regarded cautiously, considering the absent differences in priming effects between the demented and nondemented subjects in the ‘Mirror-reading task’. Though these two tasks require different cognitive operations, the difference in priming effect derived from the ‘Perceptual identification task’ may also have been affected by the slower initial reaction times by the nondemented subjects, compared to the demented subjects (see Figure 18). Consequently, the nondemented subjects may have had more opportunity to improve their reaction times when items were repeated than the demented subjects, although no significant overall group difference was found. Thus, the predictive value of the priming score in this task seems promising, but further research is needed to investigate the validity of this effect. Remarkably, several studies report normal repetition priming effects for both normal elderly subjects and AD patients, when tested by means of perceptual identification of words (see Chapter I for a review; e.g., Abbenhuis, Raaijmakers, Raaijmakers & Van Woerden, 1990; Meiran & Jelicic, 1995). In contrast, priming experiments based on more conceptually (i.e., semantically) based encoding tasks do reflect deficits in AD patients’ performance, relative to normal elderly controls (e.g., Butters, Heindel & Salmon, 1990; Gabrieli et al., 1994).

In sum, promising tasks regarding the early assessment of dementia are explicit memory tests requiring semantic processing, and possibly as well implicit memory tests from which repetition priming effects can be derived. Tasks typically used in clinical practice to assess dementia, such as free recall of semantically unrelated items and recognition of words with semantically related distractors, seem of less accurate predictive value, and may only be able to differentiate between pathological and normal ageing when dementia has progressed to a more advanced stage. The usefulness of priming effects needs further investigation, but seems promising. Furthermore, the predictive value of semantic (category) fluency tasks will be examined in more detail in the next chapter.
Chapter V has led to the following paper, submitted for publication: