Early assessment of dementia: the contribution of different memory components
Spaan, P.E.J.

Citation for published version (APA):
Spaan, P. E. J. (2002). Early assessment of dementia: the contribution of different memory components

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER I: REVIEW OF CLINICAL AND EXPERIMENTAL MEMORY TESTING AND THEIR APPLICATION TO THE STUDY OF AGEING AND DEMENTIA

Human memory functioning is one of the most frequently studied research topics in experimental psychology. Many different tasks have been devised to test models about the architecture of memory. These tasks have also been employed to explore memory performance in many different research populations, varying from young to old age and from cognitively healthy to cognitively disordered. Brain-damaged patients have become increasingly important in experimental research.

Also in clinical research, a growing amount of attention is paid to the investigation of brain-damaged patients: to date there is an abundance of neuropsychological tests constructed to measure a broad range of cognitive abilities, including memory performance. These tests are aimed at improving assessment of cognitive dysfunctions in many different patients. Memory functioning plays an important role in neuropsychological assessment, particularly in ageing and dementia cases. The same is true for experimental research: memory functioning is frequently studied, especially in dementia patients, using specifically designed tasks to test various hypotheses about the features and severity of memory impairments.

Nonetheless, it is remarkable that the development of experimental neuropsychological research on the one hand and of clinical neuropsychological testing on the other have gone rather separate ways. This will be shown in part I of this chapter, in which I will present an outline of clinically administered memory tests, varying from the oldest tests to more recent examples and representing different modalities and different ways to retrieve information (section 1). In addition, in section 2, I will briefly discuss the most important experimental memory paradigms that have influenced the conceptualisation of memory. Although memory was originally considered as a single unitary system, many experimental studies of the past decade showed that memory might be better viewed as consisting of various subsystems that are most likely subserved by different (sets of) brain structures (Schacter & Tulving, 1994; Squire, 1992). Furthermore, the most frequently used experimental tasks, devised to investigate these different memory subsystems, will be discussed. Subsequently, I will compare the type of measurements integrated in the clinical tests and the experimental tasks and discuss its gains and shortcomings (section 3).

In part II of this chapter, attention will be focused on the differentiation between dementia (Alzheimer's Disease (AD) in particular) and normal ageing. It is well known that elderly people as well as dementia patients exhibit a decline in their functioning of memory. In an early stage of dementia, and especially in a preclinical phase, the distinction with memory problems common at an advanced age is difficult. Clinical memory testing is almost exclusively based on the measurement of episodic memory – the conscious recollection of previous personal experiences or episodes – tested by, for example, word list learning or recognising pictures. However, episodic memory performance of cognitively healthy elderly subjects and early phase demented patients tends to be quite similar. Other memory systems might differentiate better between normal ageing and (early) dementia: e.g., semantic memory (the store
of facts and general knowledge, including the mental lexicon) and implicit memory (the nonconscious influence of past experiences on subsequent performance). Semantic memory in particular, but also certain forms of implicit memory, are alleged to be relatively spared in normal ageing, contrary to early phase dementia (i.e., AD in particular; e.g., Nebes, 1992; Salmon & Heindel, 1992; Fleischman & Gabrieli, 1998). However, semantic and implicit memory are hardly ever explicitly implemented in clinical memory tests, as will also be evident from part I of this chapter. It should be noted that there are several clinical tests that do measure aspects of semantic memory (e.g., several subtests of verbal intelligence of the WAIS-R (Wechsler, 1981), the Boston Naming Test (Kaplan, Goodglass & Weintraub, 1983), or various tests of category fluency). However, these tests are hardly ever interpreted as measures of semantic memory. Experimental neuropsychological research, on the other hand, pays more and more attention to these memory systems. To illustrate this, part II of this chapter will present a review of relevant experimental findings concerning the performance of normal elderly subjects and AD patients on tasks representing these different memory systems. In addition, the syndrome of dementia will be discussed in clinical terms, according to criteria proposed by the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association, 1994). The ability of the DSM-IV criteria and the available clinical memory tests to describe the cognitive profile of AD, known from the experimental studies, will be discussed. Furthermore, an attempt will be made to describe the impairments occurring in a very early (or even preclinical) phase of the disease, and the corresponding measures to detect them.

At the end of this chapter, the aim of the present research and the benefits that it might have will be discussed (part III). This research investigates the efficiency of a new computerised memory test battery in detecting preclinical dementia. This battery (in which the various memory subsystems are implemented) is administered to a population-based sample of initially nondemented elderly subjects. By means of a longitudinally research design it is possible to determine the predictive value of the various memory measures (episodic, semantic, implicit). The official clinical (DSM-IV) diagnosis of dementia is largely based on episodic memory functioning. The current approach investigates memory functioning in a more comprehensive and theoretically founded way, which might aid to a better and earlier assessment of dementia.

I Clinical and experimental memory testing

1. Clinically administered memory tests in neuropsychological assessment

In this section, various types of clinical memory tests will be considered. Detailed information for each test is given in the Appendix. The selection of tests is to a large extent based on Lezak’s (1995) comprehensive reference work on neuropsychological assessment. Information about the earliest memory tests, used in clinical practice, is derived from the review by Erickson and Scott (1977).
The classification of tests appears to be strongly influenced, in the first place, by perceptual modality, mainly verbal and visual (see Table 1 and 2). However, many alleged visual memory tests contain stimuli susceptible to verbal encoding. For example, in Rey’s Visual Design Learning Test (Rey, 1968), each item contains two elements that are easily verbalisable (e.g., a dot in a circle, a triangle above a horizontal line). The Visual Spatial Learning Test (Malec, Ivnik & Hinkeldey, 1991) on the other hand, makes use of a 6 by 4 grid and seven stimulus items placed on the grid, which provides different nonsense designs that are difficult to verbalise. In short, tests using geometric figures are susceptible to verbal encoding because they are nameable. Consequently, one should be cautious of interpreting such a memory test as a ‘visual’ memory test. Nevertheless, perceptual modality is typically the first classification criterion of memory tests.

Furthermore, the tests are classified according to the degree of inherent organization of the material being memorised and to the mode of reproduction. Successful free reproduction (i.e., without giving any cues for retrieval) of semantically and phonologically unrelated, items requires the initiation of elaborative encoding strategies and active retrieval processes. A relevant and frequently administered example is the Rey Auditory Verbal Learning Test (Rey, 1964). Recalling a previously told story – as in the California Discourse Memory Test (Kramer, Delis & Kaplan, 1988) – requires less self-initiated encoding, because the material to be reproduced possesses a meaningful structure itself (which usually aids the retrieval process). Similarly, learning semantically associated pairs of words will take less time and effort than learning unrelated pairs of words (as in the Verbal Paired Associates subtest of the Wechsler Memory Scale-Revised (Wechsler, 1987)). Tasks asking for the recognition of previously presented items out of an array of old and new items demand the least amount of retrieval effort (provided that the items were sufficiently encoded) and this form of testing memory is typically referred to as ‘passive’ retrieval. At the same time, accurate recognition of target items may be complicated by the introduction of semantically and/or phonologically related non-target items (e.g., implemented in the Hopkins Verbal Learning Test (Brandt, 1991)).

In addition, many tests include a delayed recall trial in order to determine the rate of forgetting of newly learned information. The degree of retention of material is usually tested after 20 to 30 minutes, but some tests include a 24-hour period of testing retention (e.g., the New Word Learning and Retention Test (Shapiro & Nelson, 1955)).

Most tests measure the ability to learn new information. Some other tests require the retrieval of old, previously learned information, as in tests concerning autobiographical memory (Kopelman, Wilson & Baddeley, 1989) or remote memory (see Table 4). Figure 1 illustrates a model for classification of clinical memory tests (e.g., see Lezak (1995) or the Appendix for examples of specific tests). In Table 5, a selection is presented of memory batteries that consist of different ways, described above and displayed in Figure 1, to measure memory performance.

The first memory test was developed by Wells and Martin in 1923. The test involved both verbal and visual modalities and old information as well as new learning, but none of the tasks called upon the subject to retain newly learned information over a period of time. In addition, it correlated highly with intelligence. Erickson and Scott reviewed several other memory tests, published between 1923 and 1968. From this review, it may be noticed that paired-associate paradigms were frequently applied, as well as recognition procedures. In addition, testing learning under different modalities and adding a delayed recall trial had become usual practice.
On the other hand, specific visual memory tests (in which stimuli are not verballisable) had not been developed yet. The objective of constructing these new memory tests was to obtain a measure that could distinguish performance of brain damaged versus non-brain damaged patients. In addition, attention was paid to possibly confounding effects of age and intelligence. More detailed information about these earliest tests can be found in the Appendix.

Figure 1: Classification of clinical memory tests (e.g., after Lezak, 1995).

The collection of memory tests published after 1968 considering, it may be concluded that an increasing amount of new tests has been developed over the years. Each test tries to compensate for the shortcomings of previously published tests, for example the effort to construct visual stimuli that cannot be encoded verbally in order to obtain a pure measure of visual memory. However, few tests fundamentally change their path by trying to explore really different aspects of memory. In fact, most tests are variations of already existing tests, though more recent tests investigate memory aspects in a more specific manner. Perhaps an exception to this observation is the introduction of prospective memory items in, for example the Rivermead Behavioural Memory Test (Wilson, Cockburn & Baddeley, 1985) - prospective memory referring to 'the capacity to remember to do things'.

Nevertheless, Berg and Deelman (1997) noted that memory tests are entirely limited to the domain of declarative memory (which can be subdivided into episodic and semantic memory). They distinguish tests that measure 'learning new information' and tests that require the 'retrieval of old information' (see also Figure 1). Berg and Deelman conclude that in neuropsychological assessment, the ability to learn new information is solely measured by episodic memory tests. In sum, the clinical memory tests presented here do not reflect a clear development of changing
ideas about how to measure memory performance. This will be illustrated in the next section with a brief description of the most important theories and concepts about memory.

2. Memory models and experimentally administered memory tasks in neuropsychological research

Originally, memory was conceptualised as a single unitary system. Attention was focused on the stages of learning: acquisition or encoding (the initial processing of the information that is to be learned or memorised), storage (holding the acquired information in 'memory') and retrieval (accessing the information so it can be used).

In 1968, Atkinson and Shiffrin introduced their modal model of memory, which included sensory memories, a short-term memory and a long-term memory. Sensory memories contain information similar to the sensory input itself, acquired when the information first reached the eyes or ears. Sensory memories are very short-lived, but last long enough to allow interpretation and recognition of incoming information. This analysis of incoming information can be recoded into short-term memory. Short-term memory is small in size, containing only 7 ± 2 items (Miller, 1956). Information in short-term memory is readily and easily available. This information is lost quickly (within 30 seconds), but can be maintained by means of repetition. Information can be stored in long-term memory by the activity of control processes, such as repetition and encoding. Contrary to short-term memory, long-term memory is vast in size, but information can be entered from short-term memory into long-term memory only through the expenditure of time and effort.

In contrast with the modal model’s emphasis on short-term and long-term memory, an alternative approach to memory functioning, the levels of processing hypothesis by Craik and Lockhart (1972), stresses the importance of depth of encoding. They found that elaborative rehearsal, encoding material richly in terms of existing knowledge, leads to better recall than maintenance rehearsal, the simple process of repetitive rehearsal. In terms of the modal model, maintenance rehearsal appears to maintain information in short-term memory, but not to facilitate transfer into long-term storage. Elaborative rehearsal does place material in more permanent storage. Elaborative rehearsal can be done at any of several levels, with deep processing (thinking about meaning) leading to better memory retention than shallow processing (thinking about appearances). However, depth of processing depends both on the subject’s goals and intentions and on the expertise that the subject brings to the learning situation.

In response to the modal model, Baddeley and Hitch (1974) proposed the working memory model, in which they abandoned the modal model’s assumption of a unitary short-term store and accepted instead a multi-component working memory, consisting of three subsystems. The central executive forms an interface between long-term memory and two slave systems. These slave systems combine the capacity for the temporary storage of information with an active set of control processes that allow information to be intentionally registered and maintained within the subsystem. One of these, the visuospatial sketchpad, is specialised for maintaining visuospatial information, while verbal information is held using the phonological or articulatory loop. The central executive is assumed to be responsible for the selection and operation of strategies and for maintaining and switching attention as the need arises. Working memory (or short term memory) is frequently investigated by determining
its span or the maximum amount of information it can hold available for application. This span can be measured with auditory-verbal material (e.g., recalling a certain sequence of presented digits) or with visuospatial material (e.g., reproducing a sequence of presented positions on a board).

As a result of frequent observations that amnesic patients showed preserved unconscious learning and remembering, despite their failure to consciously recollect the experience of learning the information (as in amnesic patient H. M.), the assumed unitary long-term store (Atkinson & Shiffrin, 1968) is rejected. Long-term memory is now often divided into explicit (conscious) and implicit (unconscious) memory (Schacter, 1992; Tulving & Schacter, 1990; see also Figure 2). This division is also termed the declarative-nondeclarative distinction, where declarative (explicit) memory refers to conscious recollections of facts and events, while nondeclarative (implicit) memory refers to a heterogeneous collection of nonconscious memory abilities including skills and habits, priming and simple conditioning (Squire, 1992).

**Figure 2**: Components of memory (e.g., after Tulving, 1972; Baddeley & Hitch, 1974; Squire, 1992; Schacter & Tulving, 1994).

Within the concept of explicit memory, or declarative memory, two subsystems may be distinguished, originally defined by Tulving (1972) as episodic and semantic memory. **Episodic memory** refers to the system involved in recollecting particular personal experiences or episodes (events) and is most closely associated with the traditional (clinical) approaches to learning and memory. Episodic memory can be measured by means of free recall (as in list-learning tests), cued recall and recognition tests – whether verbal or nonverbal. These tests were discussed extensively in the previous section.

**Semantic memory** refers to the store of facts and general knowledge, including mental lexicon (the meaning of words and concepts, vocabulary) or “...the associative network of permanent knowledge about the world which has been built up over one’s lifetime” (Ober, Dronkers, Koss, Delis & Friedland, 1986, p. 76). As opposed to episodic memory, semantic memory is not learning-context dependent: it is not necessary to remember the specific occasion on which one had acquired the particular knowledge item. Examples of semantic memory tasks are verbal (or category) fluency, word identification tasks and (object) naming. In category fluency, the subject is presented a word that refers to a broad category (e.g., animals, occupations, music instruments) with the task to name as many exemplars belonging to this...
category word as possible within a certain time limit (usually 60 seconds). It is assumed that subjects who are able to generate more exemplars are faster in retrieving relevant items from their semantic or lexical memory, probably facilitated by an efficient organisation of their associative (semantic) network. In word identification tasks, words of different lexical frequencies are presented for a very short period of time. The subject is instructed to read them aloud as quickly as possible. It is assumed that subjects with better access to their semantic or lexical memory will identify the words faster. Less frequent words will be more difficult to identify than more frequent words, which is again supposed to be an effect of semantic memory functioning. In object naming tasks, subjects must name visually presented objects (by means of line drawings or pictures). Less frequently encountered objects will be more difficult to name than more familiar objects.

Implicit (or nondeclarative) memory can be divided into several learning processes, such as priming, procedural memory and more primitive learning processes (e.g., classical conditioning). These learning processes all share in common that experience alters behaviour nonconsciously without providing access to any memory content (Squire, 1992). Priming refers to the improvement in performance on information that one has recently processed without the need to consciously remember this previous processing (repetition priming). Priming effects are measured by giving the subject a task (the test phase), which is not supposed to be a memory task, but actually refers to stimuli processed in another task some time before (the study phase). Priming effects are greater when the required type of processing of stimuli during the study phase corresponds better with the required processing during the test phase (Roediger & McDermott, 1993). A first processing of items on physical characteristics leads to a faster response when items are later briefly presented in a perceptual identification task (perceptual priming effect). Similarly, processing of presented words on semantic level leads to a higher probability of naming these words in a free association task (conceptual priming effect). Repetition priming is usually item-specific (i.e., item priming): faster or more accurate performance is on the exact same stimuli that were processed before. Item priming may be contrasted with associative priming, which refers to facilitation from repetition of connections between stimuli. Repetition priming is occasionally divided in identification vs. generation priming to explain patterns of results observed in Alzheimer Disease patients (e.g., Gabrieli et al., 1994; Vaidya et al., 1997).

Other examples of tasks that intend to measure (item) priming effects are word stem completion, word fragment completion and picture completion. In these tasks incomplete stimuli are presented, half of them referring to previously studied items and half of them referring to new items that have not been presented before. The subject is instructed to complete the stimuli with the first item that comes to mind. Priming is indicated by the increased likelihood of completing the stimuli with items that had been studied before, relative to an unstudied baseline condition, while it is not necessary to consciously remember the study phase. In another frequently applied priming task, the lexical decision task, subjects are presented strings of letters; the task is to decide whether the string of letters represents an existing word or not. Lexical decisions are significantly faster if subjects have recently seen the test word, even if the subjects have no explicit memory for the previous encounter.

Procedural memory refers to the acquisition of skills, whether motor, perceptuo-motor, (verbal-)perceptual or cognitive skills. ‘Skill learning’ implicates the acquisition of procedures and operations that occurs gradually with repeated exposure to items. Skill learning is demonstrated by general improvement on a task as a
function of practice rather than improvement on a specific item within that task (e.g.,
when new words that have not been presented before in the experiment are read faster
across successive trials). Perceptuomotor skill learning is frequently tested by the
serial reaction time task, in which subjects respond (e.g., by pressing a button) to each
stimulus (e.g., onset of one of several lights) in a long series. Learning is shown by
faster responding in a condition in which the stimuli appear in a repeating sequence
versus a condition in which the stimuli appear in a random order. Another
perceptuomotor skill learning task is mirror tracing, in which the subject must draw
specific lines while watching his movements in a mirror. Motor skill learning is
frequently tested by the pursuit rotor task. In this task, the subject holds the stylus of a
rotary pursuit apparatus in his preferred hand, while it rests on a small metal target.
The subject must try to maintain contact with the target as it rotates at a certain speed.
An example of verbal-perceptual skill learning is reading transformed script, which
measures the (improvement in) speed and accuracy of reading mirror-text or upside
down written words. A cognitive(-perceptual) skill learning test is maze learning, in
which the subject must solve increasingly difficult mazes and improvement in speed
and capacity is registered.

3. Conclusion

Comparing the clinical memory tests with the experimental tasks, it is evident that the
clinical measures are dominated by free recall, cued recall (or paired-associate) and
recognition testing. The emphasis of interpretation lies on the different stages of
memory, mainly encoding versus retrieval. Within the field of retrieval, short-term vs.
long-term retention and perceptual modality are investigated. In clinical memory
testing, however, short-term memory usually refers to immediate free recall trials and
long-term memory to delayed recall trials (in list-learning tests for example). Actually, the only short-term memory measures are recency effects found in list-
learning tests. Therefore, delayed recall trials are not the only long-term memory
measures, but also the immediate learning trials may be included as reflecting long-
term memory. With respect to perceptual modalities, several so-called visual memory
tests are often affected by verbal encoding possibilities, as was discussed before.
Thus, interpretations of clinical memory tests in terms of deficits in short-term or
long-term memory or conclusions concerning perceptual modalities should be
considered cautiously. Nevertheless, the Doors and People Test (Baddeley, Emslie &
Nimmo-Smith, 1994), outlined in Table 5 of the Appendix, may be noted as a
clinically administered memory test which systematically contrasts visual and verbal
memory measures and includes free recall, cued recall and recognition procedures. In
addition, the Digit Span tests (Wechsler, 1945, 1955, 1981, 1987) are evident
examples of verbal short-term memory tests and the frequently administered Corsi
Block Tapping Test (Milner, 1971) unambiguously examines visuospatial short-term
memory. These two tests may also be interpreted in terms of the phonological loop
and the visuospatial sketchpad of the working memory model by Baddeley and Hitch
(1974).

Nevertheless, the emphasis in experimental memory research on the explicit-
implicit division and its corresponding subsystems is not represented in clinical
memory tests. Apparently, clinical memory tests still assume long-term memory to
represent a single, unitary system – comparable with Tulving’s episodic memory –
which comprises the stages of encoding, storage and retrieval and can be measured
with verbal and nonverbal stimuli. In fact, clinical memory tests – except for the Doors and People Test – never use the term ‘episodic memory’.

Implicit memory is officially not implemented in clinical memory tests. Clinical tests measuring priming effects do not exist – either in memory tests or in other neuropsychological tests. However, one could argue that the procedural memory system or skill learning is reflected in, for example, a concept learning test like the Wisconsin Card Sorting Test (Berg, 1948; Grant & Berg, 1948), in which an underlying rule must be discovered (i.e., learned) on the basis of feedback when sorting cards that differ in colour, number and type of figures. Another clinical test which includes procedural memory aspects is the Porteus Maze Test (Porteus, 1959, 1965), a cognitive(-perceptual) skill learning test. Maze learning is indeed used in experimental tasks aimed at measuring skill learning effects. These two tests are, in fact, cognitive skill learning tests; purely perceptual or motor skill learning has never been implemented in clinical neuropsychological tests. However, these tests are classified as executive functioning tests, with planning and cognitive flexibility as important characteristics. Their interpretation does not emphasise the learning aspect. Nonetheless, skill learning or procedural memory is indeed measured in these tests, but practical usage of them as memory tests is complicated by a difference in terminology.

As a representation of semantic memory, the Boston Naming Test (Kaplan, Goodglass & Weintraub, 1983) could be put forward, but this test is usually considered to be an aphasia test (confrontation naming ability) rather than a memory test. Another frequently used ‘semantic memory’ test is verbal fluency (category naming), but this test is also not interpreted as reflecting semantic memory. Instead, it is usually explained in terms of executive functioning or language production. At the same time, naming and verbal fluency are used as semantic memory measures in experimental research. Also in clinical paired-associate learning tests (or cued recall tests), semantic memory plays an important role, since without awareness of the semantic relations within the material to be learned, an advantage in recalling it could not occur. This recall advantage refers to the levels of processing hypothesis by Craik and Lockhart (1972), discussed previously.

In sum, there seems to be a difference in terminology (and therefore in test interpretation) between clinical and experimental tasks. Certain aspects of several clinical tests indeed measure different types of memory, but they are used to measure other cognitive functions (such as language or the executive functions). Of course, certain tests could be used to measure semantic or procedural memory aspects after all, but interpretation of performance will be complicated by difference in scoring methods and a lack of normative data. The overlap of different cognitive functions (memory, language, executive functions) again illustrates the necessity to interpret neuropsychological test results with great caution. Nonetheless, the lack of actual emphasis in clinical testing on semantic memory, and especially on implicit memory (priming and procedural memory), might well be a serious shortcoming in differential assessment. In other words, the lack of implementation of recent theoretical concepts in clinical memory tests may lead to ignorance of important differences between clinically resembling disorders. Clinical memory tests and differential assessment might benefit from a new approach, in which various explicit and implicit subtests are included. This will be illustrated in the second part of this chapter, in light of the ageing/dementia differentiation, in which memory functioning plays an essential role and shows subtle differences.
II Ageing and dementia

The term *dementia* defines the neuropsychological and behavioural consequences of an acquired pathology of the brain in previously normal individuals (Spinnler & Della Sala, 1999). Dementia differs from the acute confusional state, known as *delirium*, which identifies an episode of neuropsychological disorder of acute onset and short duration, mainly characterised by disorientation in various domains (time, space, family, self). Spinnler and Della Sala summarise the definition of dementia as a complex set of modifications in behaviour, to which heterogeneous neuropsychological deficits combine, giving rise to the picture of cognitive incompetence. The main feature of dementia is its slow, steady deterioration, the progression of which should be detectable within a 6-12 month follow-up.

Ageing processes are highly relevant in the study of dementia, because of the increased occurrence of dementia in late life. A comprehensive population-based study by Ott et al. (1995) found a prevalence of dementia cases of 6.3% in the population of 55 years and older. Between age 55 and 65, dementia is very rare (approximately four cases to a thousand people). But from age 65, the prevalence increases from 0.9% to 34% in the population of 85 years and older. Ageing processes are especially important to consider, since the early development of dementia is regularly mistaken for a normal ageing effect. Differential assessment is complicated, because of the occurrence of memory problems in normal ageing as well as in dementia.

In section 1, the concept of dementia will be clarified in more detail from a clinical point of view, according to the official diagnostic features proposed by the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV; American Psychiatric Association, 1994). Subsequently, in section 2, findings from experimental memory research, investigating dementia patients and normal elderly subjects, will be reviewed. Findings are organised according to the explicit and implicit memory subsystems (i.e., episodic memory, semantic memory, tasks measuring priming effects, procedural memory) and findings concerning short-term memory will be reviewed as well. Attention will be focussed on AD, the type of dementia with the highest prevalence. AD is characterised by a gradual onset and a slow progressive cognitive decline. Thus, the course of AD makes it particularly difficult to detect it at an early stage. At the same time, early assessment is crucial for the efficacy of future treatment opportunities. Therefore, in section 3, an attempt will be made to describe, according to literature findings, the impairments occurring in an early (or even preclinical) phase of the disease, and the corresponding measures to detect them.

As was discussed above, dementia patients as well as normal elderly subjects exhibit impaired episodic memory performance (the type of memory extensively tested in clinical assessment). Thus, especially in an early or preclinical stage of dementia, it is difficult to detect pathological episodic memory performance, which is not merely the effect of a normal ageing process. Therefore, the literature review will summarise performance characteristics of (early) AD patients and normal elderly subjects on the other memory subsystems as well. Finally, the experimental findings and the clinical DSM-IV features of dementia will be compared, and the ability of the DSM-IV features and the available clinical memory tests to describe the cognitive profile of dementia, known from the experimental studies, will be discussed.
CHAPTER I

1. DSM-IV

The essential feature of a dementia is the development of multiple cognitive deficits that include memory impairment and at least one of the following cognitive disturbances: aphasia, apraxia, agnosia or a disturbance in executive functioning. DSM-IV reports that memory impairment in dementia implicates "an impaired ability to learn new material or the forgetting of previously learned material" (p. 134). Note that clinical memory tests, for example used to assess dementia, are constructed according to this definition of memory impairments (see also Figure 1). Deterioration of language function (aphasia) can be manifested by a difficulty in producing the names of individuals and objects. Furthermore, it may be noted that speech may become vague or empty. Comprehension of spoken and written language and repetition of language may also be compromised. Apraxia in dementia implies the impaired ability to pantomime the use of objects or to execute known motor acts. Apraxia may contribute to deficits in cooking, dressing and drawing (such as copying intersecting pentagons). Agnosia may be manifested by the failure to recognise or identify objects despite intact sensory function. Disturbances in executive functioning can be demonstrated by an impairment in abstract thinking (e.g., difficulty coping with novel tasks) or by a reduced ability to shift mental sets, to generate novel verbal or nonverbal information and to execute serial motor activities.

DSM-IV emphasises that the cognitive deficits must be sufficiently severe to cause impairment in occupational or social functioning and must represent a decline from a previously higher level of functioning. In addition, a diagnosis of dementia should not be made if the cognitive deficits occur exclusively during the course of a delirium.

DSM-IV warns that dementia should be distinguished from the normal decline in cognitive functioning that occurs with ageing (as in Age-Related Cognitive Decline). The diagnosis of dementia is warranted only if there is "... demonstrable evidence of greater memory and other cognitive impairment than would be expected due to normal ageing processes" (p. 139). Yet, it is questionable how this last feature can be demonstrated reliably. Demented elderly persons with a high premorbid level are not easily identified: their level of cognitive performance will still exceed cut-off values of neuropsychological tests. And even when one disposes of a premorbid measure, it is still not certain that an eventual decline represents the effect of a dementia. More specific information than DSM-IV provides, is needed to draw valid conclusions.

The DSM-IV features of dementia described above clearly reveal a different cognitive terminology than is common in experimental psychology. It might be informative to 'translate' the cognitive DSM-IV features into experimental memory terms. Particularly the features concerning memory, aphasia and executive functioning - rather than apraxia and agnosia - lend themselves to translation into experimental memory terms. One could argue that the described memory impairments are equivalent to episodic memory disorders (i.e., learning new material or recalling previously learned material). Aphasia may be compared to semantic memory problems, characterised by difficulty producing names of individuals and objects and vague and empty speech (i.e., loss of vocabulary). Disturbances in executive functioning may reflect, among others, procedural memory dysfunction (concerning abstract thinking abilities or coping with novel tasks). Naturally, these are rather global comparisons, arguing at the same time that DSM-IV provides poorly specified guidelines with respect to cognitive dysfunctions. In sum, some similarities between
experimental memory terms and these clinical features of dementia may exist (except for priming effects and short-term or working memory performance), but the clinical features are much less elaborately specified than is common in experimental research, as will also follow from the literature review in the next section.

DSM-IV also distinguishes different types of dementia, that all share in common the requirements described above. DSM-IV does not provide specific distinguishing cognitive features, but merely describes global characteristics of the course of the several types. The type with the highest prevalence is *Alzheimer’s Disease* (AD), which is also the type focused on in this chapter. AD is characterised by a gradual onset and a slowly progressive cognitive decline. A common pattern is an insidious onset, with early deficits in ‘recent memory’ followed by the development of aphasia, apraxia and agnosia after several years. Other examples of dementia types are Vascular Dementia (formerly Multi-Infarct Dementia), Parkinson’s Disease and Huntington’s Disease, but these types are much less prevalent than AD. AD ranks between 45% (Cummings & Benson, 1992) and 60-85% of all dementias (Evans et al., 1989). A recent study (Spinnler & Della Sala, 1999) even reported 86.2% of their demented cases (*n*=456) to be suffering from AD. Therefore, in the remaining part of this chapter, attention will be focused on memory functioning of AD patients.

2. Literature review of memory functioning in normal ageing and AD

This section presents a review of relevant experimental research findings on memory performance in normal ageing and AD. Journal articles were obtained from computerised database searches of PsychInfo between 1983 and 1998. The key words were: Alzheimer’s Disease, aging, episodic memory, cued recall, free recall, semantic memory, fluency, short-term memory, priming, word identification, skill learning, and mirror reading. Some articles were also located by citation. Studies were included only if a group of AD patients or normal elderly subjects was directly compared with an appropriate control group (i.e., nondemented normal elderly subjects and younger subjects, respectively). The findings will be organised according to the division of memory components, as was illustrated in Figure 2.

2.1. Episodic memory

Relatively large age effects between older and younger subjects were found when comparing their performance on *free recall* of word lists (e.g., Spinnler, Della Sala, Bandera & Baddeley, 1988; Java & Gardiner, 1991; Jelicic, Craik & Moscovitch, 1996). In a longitudinal study, Giambra, Arenberg, Zonderman, Kawas and Costa Jr. (1995) found that immediate visual memory performance, tested by the Benton Visual Retention Test (see also Appendix, Table 2), declined significantly from the age of 65.

However, AD patients performed even worse than older normal control subjects on free recall of word lists (e.g., Eslinger & Damasio, 1986; Spinnler et al., 1988; Deweer et al., 1992). Martin, Brouwers, Cox and Fedio (1985) found that word list learning in older normal control subjects and AD patients showed comparable patterns of performance (concerning primacy and recency effects), but AD patients scored on a lower reproduction level. AD patients seemed incapable of learning by deficient encoding rather than by impaired retrieval since their free recall performance was as poor as their recognition performance (Greene, Baddeley & Hodges, 1996).
Greene et al., using the Doors and People Test (see Appendix, Table 5), did not find a difference between verbal and visual recall. Furthermore, Greene et al. reported no differences in forgetting rate between older normal control subjects and AD patients, when comparing immediate and delayed recall trials. In addition, AD patients exhibited many intrusions in word recall (Helkala, Laulumaa, Soininen & Riekkinen, 1989).

Grober and Kawas (1997) reported that age-associated memory deficits are due to impaired attention, reduced processing capacity or the use of inefficient strategies. However, using the free and cued selective reminding procedure (see also Appendix, Table 1) proved that age-associated memory deficits are reduced when cognitive processing is controlled. Generally, elderly subjects appeared to perform worse than younger subjects on cued recall tasks (e.g., Java & Gardiner, 1991), but no age difference could be demonstrated in benefit from inherent structure of to be learned material: elderly as well as younger subjects exhibited better recall after deep encoding, or with semantically clustered material, than after superficial encoding or with unclustered material (e.g., Hart, Colenda, Dougherty & Wade, 1992; Bäckman & Wahlin, 1995; Monti et al., 1996).

Nevertheless, semantic cueing could not elevate recall performance of AD patients, because of deficient semantic encoding (e.g., Chertkow & Bub, 1990; Bird & Luszcz, 1991; Bondi & Kasznik, 1991; Hart et al., 1992; Russo & Spinnler, 1994; Monti et al., 1996). To be more precise, 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.

Abbenhuis, Raaijmakers, Raaijmakers and Van Woerden (1990) found that older normal control subjects performed worse than younger subjects on recognition of words that were presented in a perceptual identification task. The meta-analysis by LaVoie & Light (1994) showed that elderly subjects exhibited lower recognition levels of performance than younger subjects, but this effect turned out to be much smaller than the age difference found with recall conditions.

AD patients, on the other hand, did not perform better on verbal recognition tasks than on recall conditions (e.g., Eslinger & Damasio, 1986; Heindel et al., 1988; Abbenhuis et al., 1990; Grosse, Wilson & Fox, 1990; Deweer et al., 1992; Deweer, Pillon, Michon & Dubois, 1993; Russo & Spinnler, 1994; Fleischman et al., 1996; Koivisto, Portin & Rinne, 1996). Eslinger and Damasio reported that AD patients were also unable to recognise pictures of unfamiliar faces. In addition, Green et al. (1996) found that AD patients were equally impaired on visual and verbal recognition trials of the Doors and People test. Furthermore, Keane, Gabrieli, Growdon and Corkin (1994) found that AD patients performed poorly on the recognition of pseudowords. From these findings, it may be concluded that AD patients’ poor performance on episodic memory tasks is a matter of poor encoding of information rather than deficient access to it. Noteworthy for the recognition performance of AD patients is that they made relatively many false positive errors, compared to false negative errors (Deweer et al., 1994). They appeared not to be able to inhibit irrelevant associations (Helkala et al., 1989). Brandt, Corwin and Krafft (1992) explained the AD patients’ high number of false positive errors by their inability to discriminate between different semantic relations in the presented material: they are
sensitive to category membership of words, but they cannot discriminate between different semantic attributes of words within a given category.

In sum, normal elderly subjects as well as AD patients exhibit episodic memory problems. Normal elderly subjects typically have problems with free recall conditions, but they exhibit normal benefit from (semantic) cueing and inherently structured material. Their recognition performance is a little less efficient than that of younger subjects, but this difference is negligible compared to their impairment on free recall tasks. AD patients suffer from a general episodic memory deficit: they cannot benefit from cueing or inherent structure and their recognition ability is as impaired as their free recall performance. AD patients show impaired learning rather than accelerated forgetting or disrupted retrieval. In addition, perceptual modality of stimuli does not seem to be an important factor in episodic memory tests.

It may be concluded that normal elderly subjects have difficulty initiating retrieval processes and, thus, will improve when offered cues or structure within the material to be memorised. In other words, when retrieval requires less active (i.e., more passive) strategies, normal elderly subjects will exhibit more accurate reproduction of information. Thus, normal elderly subjects must be able to encode material efficiently, otherwise they would not be able to benefit from semantic cues or recognition formats.

On the other hand, AD patients do not show any retrieval improvement when aided by means of semantic cueing or recognition conditions. Therefore, their problems must start in a previous phase of memorising: they seem unable to encode information efficiently, which may be caused by poor (access to) semantic knowledge. Nonetheless, studies using cued recall conditions focus on semantic cueing; it has never been explicitly investigated whether phonological cues lead to some retrieval benefit for AD patients. Martin et al. (1985) found that AD patients exhibited more discrimination problems in a recognition task with semantically related distractors than with phonologically related distractors, although this finding does not directly refer to a retrieval benefit when phonologically cued. Nevertheless, a recognition measure or semantic cueing when investigating episodic memory seems necessary for the differentiation between AD and normal ageing. Particularly, the large number of false positive errors may be considered a typical performance characteristic for AD. Delayed recall trials or examining material-specific memory disturbances may not be relevant when differentiating AD and normal ageing. However, it is rather unlikely that the deficit to benefit from semantic cueing or recognition (including false positive errors) occurs in a very early stage of AD (e.g., Hodges & Patterson, 1995). Studies reviewed in this section usually examine AD patients whose diagnosis is based on symptoms occurring for several years, rather than very early (or even preclinical) stage AD patients whose diagnosis was confirmed subsequently.

2.2. Semantic memory
As a measure of semantic memory, many studies use the verbal (or category) fluency task, in which the subject must name as many exemplars of a particular category — e.g., animals, vehicles, fruits and vegetables or supermarket items — as he can think of within a certain time limit (usually one minute). Chertkow and Bub (1990) explicitly warn that verbal fluency is not a direct measure of semantic memory store. They found that verbal fluency was impaired in their AD subjects due to two major constraints: deterioration of semantic memory store and difficulties in semantic
search. Semantic search problems might be affected by attentional problems, motivation and the ability to devise strategies for breaking down categories into smaller sections for efficient search. However, one should always be careful with interpreting test results – not just specifically in the case of verbal fluency paradigms. Nevertheless, AD patients consistently performed deficiently on verbal fluency, compared to normal elderly controls (e.g., Hodges et al., 1990; Monsch et al., 1992; Salmon, Heindel & Butters, in Bäckman, 1992; Mickanin Grossman, Onishi, Auriacombe & Clark, 1994; Hodges & Patterson, 1995; Beatty, Testa, English & Winn, 1997; Sailor et al., 1998). AD may lead to a specific disruption in semantic knowledge (i.e., deterioration of semantic store – Chertkow & Bub, 1990), characterised by a difficulty in differentiating between items within the same semantic category concurrent with relative preservation of broader categorical information (Martin & Fedio, 1983). This is called the bottom-up breakdown of semantic knowledge and is often used to explain the nature of performance by AD patients (e.g., Ober, Dronkers, Koss, Delis & Friedland, 1986; Salmon, Shimamura, Butters & Smith, 1988; Tröster, Salmon, McCullough & Butters, 1989; Monsch et al., 1994; Rosser & Hodges, 1994; Binetti et al., 1995). Weingartner, Kawas, Rawlings and Shapiro (1993) noted that changes in semantic memory are detectable before the diagnosis of AD can be made: first, the patients are not able to name low-frequency exemplars and later they lose more common elements. It is concluded that in AD, loss of knowledge is the cause of impaired verbal fluency performance, rather than deficient initiation of retrieval or semantic search problems (e.g., Randolph, Braun, Goldberg & Chase, 1993; Rosser & Hodges, 1994; Monsch et al., 1994).

Normal elderly subjects typically performed better on the category fluency task than on the letter fluency task, in which as many words starting with a particular letter as one can think of must be named within a certain time limit. AD patients showed the reverse pattern: they performed better on letter fluency than on category fluency, although both types of performance were definitely impaired (e.g., Mickanin et al., 1994; Monsch et al., 1994; Rosser & Hodges, 1994). This pattern of performance usually differentiates AD patients and normal elderly subjects and illustrates the clear semantic memory problems of AD patients, relative to normal elderly controls. In addition, AD patients may be detected by their qualitative performance on the category fluency task: in addition to naming few correct exemplars in general, they typically named the most common elements (i.e., preservation of broad category information) and produced few different subcategories and few items per subcategory and relatively many category labels (e.g., Martin & Fedio, 1983; Ober et al., 1986; Tröster et al., 1989; Beatty et al., 1997).

As was mentioned before, verbal fluency is the most frequently and extensively used task to examine semantic memory performance. Alternative tasks include: confrontation naming (e.g., the Boston Naming Test - Kaplan, Goodglass & Weintraub, 1983); vocabulary (e.g., in WAIS-R - Wechsler, 1981); naming in response to verbal description and semantic feature questions (Hodges & Patterson, 1995); sentence verification (answering category and property statements; Sailor, Bramwell & Griesing, 1998). All of these tasks showed deficits in AD patients, compared with normal elderly subjects.

Hodges and Patterson concluded that semantic memory is also impaired in a very early stage – ‘minimal’ AD patients with a Mini-Mental State Examination score (Folstein, Folstein & McHugh, 1975) above 23. Thus, it may be argued that semantic memory performance might be a better early marker for AD than (purely) episodic memory performance (i.e., episodic memory performance that is relatively unaffected
by semantic processing capacities, as in free recall of inherently unstructured lists of words). This suggestion is supported by the finding that normal elderly controls also show impaired performance in free recall conditions, (as was discussed in section 2.1). However, elderly controls perform normally on tasks sensitive to semantic processing capacities, which is in contrast with AD patients’ performance. In section 3 (on very early and preclinical AD), these issues will be discussed in more detail.

2.3. Short-term (or working) memory
With respect to the auditory/verbal span, it has been established that normal elderly subjects of 60 or 65 years and older exhibited a smaller span than younger subjects (e.g., Orsini et al., 1986; Grégoire & Van der Linden, 1997). Jurden, Laipple and Jones (1993) found that recalling the precise sequence of presented digits – ‘serial position processing’ – turned out to be most age sensitive. Recalling digits when the presented sequence is irrelevant appeared to be less age sensitive. In early or minimal AD, no evident differences in performance with normal elderly controls were found (e.g., Carlesimo, Fadda, Lorusso & Caltagirone, 1994; Morris, 1994; Hodges & Patterson, 1995). Only in moderate dementia patients, performance was significantly poorer than in normal elderly subjects (Orsini, Trojano, Chiacchio & Grossi, 1988).

Normal elderly subjects of 60 years and older exhibited a smaller visuospatial span than younger subjects (Orsini et al., 1986). AD patients (also when in an early stage) performed significantly worse than normal elderly controls, which was in contrast with their relatively good performance on the auditory/verbal span (e.g., Orsini et al., 1988; Spinnler et al., 1988; Carlesimo et al., 1994; Trojano, Chiacchio, De Luca & Grossi, 1994). This deficit was also demonstrated when the sequence of to be recalled patterns was not important (Grossi, Becker, Smith & Trojano, 1993).

In sum, from the age of 60, auditory/verbal as well as visuospatial span decline. Furthermore, on auditory/verbal span normal elderly subjects and AD patients do not differ, whereas AD patients exhibit a significantly smaller visuospatial span than normal elderly controls. Apparently, AD patients suffer from a specific disorder of visual working memory.

2.4. Implicit memory: priming effects
Most studies did not demonstrate a difference between younger and older subjects in benefit from repetition of specific items on verbal-perceptual tasks (e.g., inverted-reading, partial-word identification or mirror reading): elderly subjects showed a normal (perceptual) repetition priming effect (e.g., Hashtroudi, Chrosniak & Schwartz, 1991; Grober, Ausubel, Sliwinski & Gordon, 1992; Jelicic et al., 1996). This effect occurred despite the finding that the same material was poorly explicitly recognised by these subjects. However, Abbenhuis et al. (1990) found a greater facilitation effect for younger subjects than for elderly subjects on repeated items in a perceptual identification task. But this advantage only emerged for low-frequency words, which may be affected by the lower educational level of the older subjects (11.2 years vs. 16.6 years for the younger subjects). The meta-analysis by LaVoie and Light (1994) reported a small age effect on repetition priming, but this effect was smaller than in recall or recognition tasks.

Relative to normal elderly subjects, AD patients exhibited a normal repetition priming effect as well, when tested by means of perceptual identification of words (e.g., Abbenhuis et al., 1990; Keane, Gabrieli, Fennema, Growdon & Corkin, 1991; Gabrieli et al., 1994; Russo & Spinlner, 1994; Meiran & Jelicic, 1995; Koivisto et al., 1996). Gabrieli, et al. (1994) also found perceptual identification of incomplete
pictures to be intact in AD patients. In addition, Keane et al., (1994) used pseudowords, instead of real words, which also resulted in a preserved priming effect. Furthermore, the application of a mirror reading task also lead to a normal repetition priming effect in AD patients (e.g., Grober et al., 1992; Deweer et al., 1993, 1994). Vrbancic, Rouleau, Butters and Moscovitch (1992) administered a picture-rotation naming task and an inverted-script word-pairs reading task: both devises demonstrated preserved repetition priming effects. However, the meta-analysis by Meiran and Jelicic (1995) showed that AD patients performed deficiently on nonverbal perceptual priming tasks, such as the identification of degraded pictures.

In sum, perceptual priming is relatively intact in normal ageing as well as in AD. The effect may be smaller than found in younger subjects, but relative to episodic memory disorders this reduced effect size may be negligible. Particularly on word-based perceptual priming tasks, intact priming effects are found. The reduced priming effects in AD patients on nonverbal perceptual tasks (Meiran & Jelicic) seem inconsistent with the findings by Gabrieli et al. (1994), who reported intact priming effects on perceptual identification of incomplete pictures. Gabrieli et al. investigated whether priming on their Gollin Incomplete-Pictures task was perceptual or conceptual in nature. The experiment led to the conclusion that their task was indeed perceptual in nature. However, most nonverbal perceptual priming studies, reviewed by Meiran and Jelicic, were supposed to be conceptual in nature, because a verbal answer was required in response to the degraded picture. It is believed that generating a verbal response to a nonverbal stimulus requires access to meaning and this may cause problems for AD patients. Thus, it may be concluded that perceptual priming effects are preserved in normal ageing as well as in AD, despite poor explicit memory for the stimuli used.

Another frequently used measurement for priming is the word stem completion task: the subject is exposed to two- or three-letter word stems, some of which can be completed with previously presented words (while the subject is not alerted to this possibility). Though the priming effect found for healthy subjects of 70 years and older was smaller than for younger subjects, the effect was still significant (e.g., Davis et al., 1990; Java & Gardiner, 1991; Jelicic et al., 1996). In addition, Monti et al. (1996) found normal elderly subjects (as opposed to AD patients) exhibiting normal conceptual priming for category exemplars and they benefited normally from deep over superficial encoding.

Priming effects on word stem completion tasks for AD patients were less consistent. Most studies concluded that AD patients perform inefficiently (e.g., Shimamura et al., 1987; Salmon, Shimamura, Butters & Smith, 1988; Butters, Heindel & Salmon, 1990; Bondi & Kaszniak, 1991; Christensen & Birrell, 1991; Keane et al., 1991; Gabrieli et al., 1994; Keane et al., 1994; Carlesimo, Fadda, Marfia & Caltagirone, 1995; Meiran & Jelicic, 1995). However, some studies report the contrary. For example, Grosse et al. (1990) as well as Fleischman et al. (1996) demonstrated intact word stem completion priming in AD patients after a semantic/conceptual encoding task – though numerous studies mentioned above also used semantic encoding tasks – despite poor recognition of the material.

The contradictory findings for AD patients in word stem completion tasks might be explained by the impaired semantic capacities of AD patients. Most studies, including those who reported intact word stem completion priming, demanded semantic encoding of stimuli during the study phase. However, it is questionable whether the stimuli were indeed efficiently (semantically) encoded. When elaborate
encoding does not take place, poor word stem completion priming is to be expected. Grosse et al. applied a sentence-frame completion task to their AD patients during the study phase. In this task, subjects had to complete sentence frames (e.g., “He hit the nail with a ___”) with single, best-fit words (e.g., “hammer”), which served as targets for the test phase. Grosse et al. explained the normal priming effects of the AD patients by the nature of the encoding task: “... it may have provided the sort of rich semantic contextual stimulus needed to activate the defective semantic network in AD” (p. 305).

On the other hand, Fleischman et al. (1996) found intact priming effects for AD patients after perceptual as well as after conceptual processing (reading words aloud vs. generating words from definitions), despite deficient performance on the generating condition. Fleischman et al. could not explain their findings by differences between studies in study-test interval, number of study-phase exposures, encoding modalities engaged or item type. They suggested, instead, the possibility that the pattern of intact and impaired word stem completion priming across independent studies of AD may reflect individual differences in the locus and extent of brain dysfunctions. Another explanation lies in the characteristics of the normal control subjects. Davis et al. (1990) found that age-related word stem completion priming deficits began to appear around age 70. Fleischman et al. concluded that repetition priming in control subjects at this age may be as variable as it is in early-stage AD, which could complicate the search for group differences in word stem completion priming.

Overall, in addition to intact perceptual (identification) priming effects, normal elderly subjects exhibit intact priming for category exemplars and word stem completions. AD patients show preserved perceptual (identification) priming effects as well, but deficient priming in word stem completion and category exemplars tasks. The dissociation in the priming performance of AD patients is consistent with the perceptual/conceptual (e.g., Keane et al., 1991) and the identification/generation hypotheses (Gabrieli et al., 1994), which were also mentioned in part I of this chapter. Thus, conceptual or generation priming tasks may better differentiate between AD and normal ageing than perceptual or identification priming tasks.

2.5. Implicit memory: procedural memory or skill learning
AD patients exhibit a, relative to normal elderly controls, normal improvement over trials on perceptual- motor learning tasks (e.g., pursuit rotor, mirror tracing, serial reaction time or more cognitively mediated tasks such as maze learning and weight biasing) and show normal transfer to comparable trials – despite absent explicit learning (e.g., Eslinger & Damasio, 1986; Knopman & Nissen, 1987; Heindel et al., 1988, 1989; Butters et al., 1990; Grosse, Wilson & Fox, 1991; Knopman, 1991; Gabrieli, Corkin, Mickel & Growdon, 1993).

Verbal-perceptual skills lead to quite similar results: in addition to motor skills, AD patients are also able to learn mirror reading in a normal way (comparable to normal elderly controls), despite impaired explicit recognition of presented items (e.g., Salmon et al., 1992; Deweer et al., 1993, 1994). However, Grober et al. (1992) found no learning ability for AD patients on the mirror reading skill. They explained this result in terms of the patients’ underlying deficit in abstract reasoning that precludes the development of appropriate pattern analysing strategies needed to transform rotated text. In contrast to the AD patients, older normal control subjects in this study were found to be able to learn the mirror reading skill. However, Grober et
al. did not examine skill learning abilities in a younger control group, so it is unclear whether the normal older subjects performed as well as younger subjects.

Hashtroudi et al. (1991), on the other hand, did compare younger and older normal control subjects and found that the older subjects showed less improvement in the skill of reading inverted words than did the younger subjects, although repetition priming effects did not differ. Hashtroudi et al. concluded that skill learning is impaired under data-limited conditions, while priming is unaffected under these conditions. They proposed that the age deficit in skill learning is related to a deficit in perceptual organisation and reorganisation.

In sum, most studies report that AD patients are capable of skill learning – despite their failure to consciously remember the learning experience – irrelevant whether motor, perceptual, cognitive or verbal components are involved. The contradictory finding on mirror reading by Grober et al. may be explained in light of reported spatial rotation deficits in ageing and dementia (Flicker, Ferris, Crook, Reisberg & Bartus, 1988; Armstrong & Cloud, 1998), spatial rotation ability being tested by the Money Road Map Test (Money, Alexander & Walker, 1965). Noteworthy is that the AD patients in the experiment by Grober et al. were significantly older than the older normal control subjects (83.4 and 76.9 respectively). Most subjects in the other studies were around 70 years old. Thus, the reported difference in mirror reading ability might well represent an ageing effect, rather than a disease associated effect.

The study by Hashtroudi et al. was the only study that compared a form of skill learning between younger and older subjects. From this study, it may be concluded that normal elderly subjects are slower in learning verbal-perceptual skills than younger persons, which may refer to age-associated impairments of visuospatial abilities. As was mentioned previously, Orsini et al. (1986) found a smaller visuospatial span in elderly subjects. Nonetheless, skill learning or procedural memory will not be able to differentiate between AD and normal ageing, as opposed to other implicit memory aspects like (conceptual) priming effects.

3. Focus on very early and preclinical AD

Though some remarks on the subject of early assessment of AD were already made in the previous section, this section will pay more specific attention to it. Unfortunately, most experimental studies reviewed above use AD patients who are in an advanced stage of their disease. Studies investigating preclinical AD patients usually recruit a large cohort of non-demented older subjects, who are administered a battery of neuropsychological tests at several times of measurement. This cohort is assessed longitudinally until a sufficient number of subjects have been diagnosed clinically with probable or possible AD (Collie & Maruff, 2000). Deficits on the following tests may be indicative for developing AD, several years before the diagnosis was made: (delayed) story recall (Elias et al., 2000; Collie & Maruff, 2000), Similarities (WAIS-R; Elias et al.), (verbal) paired-associate learning (Elias et al.; Collie & Maruff), and free recall (and recognition) of words (Bäckman, Small & Fratiglioni, 2001; Grober, Lipton, Hall & Crystal, 2000; Collie & Maruff). However, the neuropsychological test battery in this class of studies is usually limited to measures of (verbal) episodic memory (i.e., clinical memory tests), while other memory systems are not investigated. Nonetheless, one may argue that semantic memory plays an important role in several tests listed above: Similarities (i.e., capacity of mental lexicon), paired-
associate learning (i.e., recall benefit in pairs of words that are semantically related), and story recall (i.e., memory for meaningful material, dependent on degree of text comprehension). However, in these studies, these tests are usually not interpreted as measures of semantic memory. Thus, it may be argued that important information could be missed by concluding that (purely) episodic memory processes are crucial for the prediction of dementia. In addition, other than so-called episodic memory tests are rarely administered in these studies – at least, other memory components are not explicitly investigated in order to determine the most sensitive measures of preclinical dementia. Therefore, conclusions in this respect should be drawn cautiously.

Since the studies investigating preclinical AD patients do not examine memory according to the ‘explicit-implicit’ view (Schacter, 1992; Tulving & Schacter, 1990), a selection was made from the experimental studies reviewed above that used diagnosed AD patients but in their earliest stages (‘minimal’ AD) or with high scores on cognitive screening tests (i.e., MMSE>23). The study of Weingartner, Kawas, Rawlings and Shapiro (1993), reviewed in the section on semantic memory, was the only study that investigated preclinical AD patients (in a category fluency task, two years before the diagnosis was made). They concluded that one of the early cognitive symptoms of AD is changes in availability of uncommon exemplars of semantic networks.

Hodges and Patterson (1995) and Greene, Baddeley and Hodges (1996) both define a group of ‘minimal’ AD patients: patients diagnosed with AD who scored in the 24-30 range of the MMSE. Greene et al. present a systematic investigation of various aspects of episodic memory. They found that the ‘minimal’ AD patients performed significantly worse than normal elderly controls on immediate and delayed trials of story recall, free recall and recognition (verbal and nonverbal tests). Greene et al. conclude that the ‘minimal’ AD patients suffer from general episodic memory disorders, characterised by defective learning processes rather than faster forgetting or impaired retrieval. No effect was found for modality of material to be memorised.

Hodges and Patterson investigated how early in the course of the disease and how consistently semantic memory problems occur in AD. The ‘minimal’ AD patients demonstrated impaired performance on various tests of semantic memory (e.g., category fluency, naming, naming to verbal description, semantic feature questions) and on episodic memory (i.e., delayed story recall). Recognition memory was less impaired in ‘minimal’ AD, but may be a better index of severity of the disease. Tests of visuospatial ability and verbal short-term memory (i.e., digit span) did not show significant differences with normal elderly controls. Hodges and Patterson conclude that semantic memory is affected very early in the course of AD, though there was considerable variability in the extent of semantic impairment among patients with the same overall level of dementia. This finding regarding the early semantic memory impairment in AD is supported by Rosser and Hodges (1994), who examined category fluency performance in early AD patients (mean score Dementia Rating Scale (Mattis, 1976): 121.4). In addition, Sailor, Bramwell and Griesing (1998) found their early AD patients (mean MMSE score: 23.7) to be impaired in several tasks of semantic memory. Furthermore, Hodges, Salmon and Butters (1990) demonstrated impaired performance of their early AD patients (mean MMSE score: 24.4) on a category fluency test and some tests of semantic knowledge (e.g., Boston Naming Test, WAIS-R subtests Vocabulary and Similarities).

Few studies can be selected that investigated implicit memory performance in early AD. Koivisto, Portin and Rinne (1996) found intact perceptual priming effects in their early AD patients (mean MMSE score: 22.9). Monti et al. (1996) found
impaired priming for category exemplars in their early AD patients (mean MMSE score: 23.3), while normal elderly controls showed equivalent priming effects as younger normal controls. The AD patients did not benefit from deep encoding in either an explicit memory measure (i.e., cued recall) or an implicit memory measure (i.e., the conceptual priming task). Monti et al. argue that AD, contrary to normal ageing, is characterised by impaired conceptual processing. On the other hand, Fleischman et al. (1996), using the word stem completion task, found similar priming effects for their early AD patients (mean MMSE score: 23.3) and their controls, while usually impaired priming effects are found in AD. As was mentioned above, the age of the subjects tested may be an important factor, since priming effects in normal control subjects aged 70 or above may be as variable as in early AD.

Since hardly any studies are available that examine priming effects in very early AD patients, it is difficult to discuss the value of priming tasks for early assessment. Conceptually based priming tasks might lead to differences between very early or preclinical AD patients and their controls, since these tasks may be dependent on the functioning of semantic memory (which is most likely impaired in a very early stage of the disease). However, much more research must be done before reliable conclusions can be drawn.

4. Conclusion

From the review of experimental findings (section 2), it may be concluded that, in addition to episodic memory problems, there are also major differences in memory functioning between normal ageing and AD in the field of their semantic capacities (i.e., the structure of semantic knowledge). AD patients exhibit, relative to normal elderly controls, poor semantic encoding of to be learned information. This will also affect episodic memory performance, especially in the case of material with an inherent semantic structure or in semantic cueing tasks: AD patients cannot benefit from such cues, contrary to normal elderly controls. AD patients might not be able to discriminate between two related concepts, because the attribute knowledge that distinguishes the two concepts is lost (e.g., Martin & Fedio, 1983; Sailor et al., 1998). In addition, AD patients' deficits are evident in recognition tasks, particularly when semantically related distractors are used – their responses consist of numerous false positive errors. AD patients seem unable to inhibit irrelevant associations.

Many studies examining the impaired category fluency performance in AD, report a ‘degraded structure of semantic knowledge’. In AD, the qualitative characteristics of performance are striking: they do not simply name few correct exemplars in general, but they also show hardly any exploration or awareness of subcategorical information (i.e., the bottom-up breakdown of the semantic knowledge network). Furthermore, they name many subcategory labels and they show many perseverations, relative to their total production of exemplars.

In addition, priming experiments based on more conceptually (i.e., semantically) based encoding tasks reflect deficits in AD patients' performance as well, once again due to their impaired semantic capacities. Also their poor visuospatial span, relative to their auditory/verbal span, has been reported frequently.

Contrasted to AD, poor episodic memory performance in normal ageing mainly concerns deficient initiation of retrieval strategies, rather than poor encoding processes. This may also relate to their compromised performance on semantic memory, as in category fluency, but primarily results in a slow retrieval of relevant
exemplars without a typical profile of responses. Normal elderly controls show relatively intact implicit memory and short-term memory, at least until the age of 70.

However, most patients used in the studies reviewed in section 2 were in an advanced stage of their disease. Symptoms have been reported for several years and the scores on cognitive screening tests are generally low. From the tentative review of preclinical and very early AD patients (section 3), it may be concluded that tests sensitive to semantic knowledge are crucial for detecting AD at the earliest possible stage. These tests may include memorising material with an inherent semantic structure (e.g., story recall), semantic cueing, or category fluency. Possibly, reliable priming tasks that call upon semantic processing may also be useful.

As was discussed in previous sections, clinical memory tests focus on measuring episodic memory performance, although these tests do not use the label 'episodic memory'. The DSM-IV features of dementia regarding memory impairments (superficially) describe episodic memory disorders, again without labelling them as such. The global DSM-IV features fail to describe the more specific memory performance characteristics of AD patients and they even fail to describe the precise episodic memory performance characteristics, reported by numerous experimental studies. More specific information for assessment is gathered from the available clinical memory tests that provide redundant measures of episodic memory performance, although their interpretation is not completely consistent with current theoretical concepts (as was concluded in part I of this chapter). Other neuropsychological tests may provide additional information on memory functioning (such as the category fluency test on semantic memory), but scoring procedures should be modified to obtain more specific information (e.g., scoring the use of different subcategories instead of merely considering the sum of correct exemplars generated within one minute). Nevertheless, the available clinical memory tests fail to offer a broad view of memory functioning of normal elderly subjects and dementia patients: tests other than of purely episodic memory are needed in order to observe the essential differences. At least, the interpretation of test results should focus on aspects of performance that are sensitive to semantic processing capacities rather than simply interpreting results in terms of 'memory deficits'. Particularly in early assessment, tests calling upon semantic knowledge may aid to an earlier and more efficient assessment of AD.

In sum, one may conclude that current theoretical knowledge about memory functioning is not well reflected in clinical assessment, which may, at least in this case, lead to less systematic investigation of differential performance of cognitively healthy and impaired elderly subjects. As a result, important information could be missed and the interpretation of test results may be more hazardous than necessary. Using a broader set of memory tasks might lead to an improved differential diagnosis, especially for the early detection of dementia.

III Research purpose and questions

1. Aim of present research

Erickson and Scott already suggested in 1977 that "... the technology of the research psychologists could be employed in preparing an extensive battery of tasks that would enable us to analyse the complex process of collection, storage and retrieval" (p.
The current research is aimed at developing a memory test battery that can distinguish between purely age-related memory problems and memory disturbances that indicate future dementia.

As has been described before, an abundance of experimental neuropsychological research has been conducted to investigate differences in memory performance between normal and pathological ageing. Simultaneously, many tests have been developed for application in clinical practice with the same purposes. However, few studies are aimed at measuring differences in memory performance when the experimental group is still in a preclinical stage (see section 3 of part II). Furthermore, the neglect of preclinically demented subjects also affects the control group of normal elderly subjects. For example, the longitudinal study by Sliwinski, Lipton, Buschke and Stewart (1996) led to the conclusion that the level of cognitive functioning in normal elderly subjects tended to be an underestimation because of failing to exclude preclinical dementia patients: it resulted in underestimation of the mean, overestimation of the variance and overestimation of the effect of age on cognitive measures. Sliwinski et al. 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). Thus, investigating preclinical dementia subjects will shed light on cognitive functioning in both normal ageing and early stage dementia.

Other important factors that may influence patterns of performance are the age and level of education of the subjects under investigation. The usual sample of healthy elderly subjects and dementia patients in experimental studies is restricted to a relatively young age (around 68 years old) and high educational level (e.g., college education; around 12 years of attained education), which both are not representative for the average elderly population. Backman, Small, Wahlin and Larsson (2000) reported that little research is done on cognitive functioning in ‘very old age’, i.e., 75 years and above, although this age group is becoming more and more relevant because life expectancy continues to increase. In addition, prevalence of dementia is greatest in this age group: it increases exponentially with increasing age. Thus, differentiation of effects of normal ageing versus early dementia is a central issue in subjects of very old age. Furthermore, this differentiation may be more difficult when the subjects are of ‘very old age’, rather than of ‘young-old’ age, because normal ageing effects on cognitive performance may be more pronounced at this age, thereby complicating the differential diagnosis. In addition, low educational background complicates the investigation of very old persons: their level of cognitive functioning is easily mistaken for pathological ageing effects.

In addition, clinical as well as experimental studies predominantly focus on merely one aspect of memory functioning, instead of a broad range of aspects—which would enable us to observe how these different memory aspects interact with one another. Therefore, the present research investigates a battery of various memory tasks, based on recent conceptualisations of memory, involving explicit (i.e., episodic and semantic), implicit (i.e., priming and procedural), and working memory subtests. Episodic memory tests are free recall and recognition (with semantically related distractors) of a ten word list-learning test. Also a paired-associate learning test is included, with five semantically related and five neutral word pairs. Retrieval speed from semantic memory is measured in a category fluency test, in which as many animals and occupations as possible are to be produced in 60 seconds each. In addition, a word identification test measures speed of access to semantic memory, in which low-, middle- and high-frequency words must be read aloud as soon as they can.
be identified. A verbal measure of *short-term memory* (i.e., the phonological loop) is obtained by a digit span task and its visuospatial component (i.e., the visuospatial sketchpad) is measured by a block-positions span task. Conceptual *priming* effects can be generated in a word stem completion task, consisting of ten experimental stems (which can be completed with the words from the ten word list-learning test) and ten control stems. Perceptual priming effects are measured by means of the repetitive presentation of items from the word identification task. Verbal-perceptual skill learning (or *procedural memory*) is implemented in a mirror-reading task.

This test battery is administered to a broad, for age and education representative, sample of (initially) nondemented community dwelling elderly subjects. Two years after the first administration of the battery, a clinical (DSM-IV) diagnosis concerning dementia is made, in order to be able to determine the predictive value of the test battery. In addition, the measurements are repeated by means of a second administration of the battery in order to investigate patterns of development of memory performance in these demented and nondemented elderly subjects.

**2. Research questions**

The intended benefits of the current research, described in the previous section, lead to the following research questions.

**Chapter III:**

- Chapter III is aimed at studying the range of normal ageing effects by contrasting a group performing on a high global level of cognitive functioning (measured by the MMSE) versus a group performing on a lower global level of cognitive functioning, yet is not demented (i.e., a ‘cognitively healthy’ (NC) group and a ‘cognitively impaired’ (CI) group). The main question is how this difference in global cognitive functioning is best characterised in terms of various memory components. In other words, which specific memory aspects show greatest degree of variance within normal ageing? Furthermore, as a preliminary discussion: does the profile found for the CI subjects, as opposed to the profile described for the NC subjects, show resemblance with, for example, the profile typically described for AD patients? Chapter III gives important information on the question whether it is useful to differentiate between the various memory components (i.e., episodic memory, semantic memory, short-term or working memory, priming, procedural memory) when characterising the deficits of the CI group, relative to the NC group. See Chapter III, sections 4 and 5.

- In addition, Chapter III will address some questions concerning the general characteristics of the memory test battery: e.g., the norm scores for a broad sample of (initially) nondemented community dwelling elderly subjects, the level of reliability (internal consistency) of the various memory measures, and the influence of age, education and sex on memory performance. See Chapter III, sections 1 through 3.

- Finally, in Chapter III, it will be investigated how the different memory measures relate with one another. A related question will be: ‘Does implicit memory represent a different memory system to episodic memory, or can it be considered as part of the same memory system?’ Another question will be whether each subtest contributes independently to the profile of memory performance of the elderly subjects, or whether the value of some subtests is negligible for characterising the performance. See Chapter III, section 6.
Chapter IV:

- Chapter IV will discuss the most important question of the current research: how can preclinically demented elderly subjects (best) be differentiated from nondemented elderly subjects? In other words, which profile, based on the different memory measures, is indicative for dementia, approximately two years before diagnosis? Does this profile lead to an improvement in the prediction of dementia, above what is possible using the current clinical assessment methods? The data collected in the first administration of the memory test battery will be analysed by comparing performance of the subjects who were vs. were not diagnosed as being demented, two years after the first administration of the battery. See Chapter IV, section 5.

- In addition, these data will also provide information on the level and nature of memory performance of truly cognitively healthy elderly subjects, who were not diagnosed as demented two years later. The question will be answered whether the level of memory functioning in normal elderly subjects was originally underestimated by not excluding preclinical dementia patients, as Sliwinski, Lipton, Buschke and Stewart (1996) found. 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). See Chapter IV, section 3.

- By means of the second administration of the memory test battery (T2), data are provided regarding patterns of development of memory performance over two years of time, for each clinical group of elderly subjects separately. It will be investigated whether the various memory measures show decline or a stable pattern of performance once subjects are diagnosed as demented at T2, relative to the nondemented subjects. Thus, how may the development of memory performance of elderly subjects (best) be characterised once they are diagnosed as demented? In addition, it will be examined whether patterns differ for different stages of dementia (i.e., 'minimal', 'mild'/moderate'). These data will also be used to describe the sensitivity to decline in the nondemented subjects, and whether the patterns vary with different global levels of cognitive functioning, indicated by the MMSE. See Chapter IV, section 4.

- In addition, the memory profile of the clinically demented subjects will be described by means of the memory subtest data collected at T2. Cross-sectional analyses will investigate which T2 measures 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. See Chapter IV, section 6.

- As a final and preliminary question, Chapter IV will focus on the identification of latent subgroups within the sample of clinically nondemented subjects. Is there a group of subjects who were not diagnosed as demented at T2, but do demonstrate greater than average decline on the memory test battery from T1 to T2? If possible, the characteristics of decline and the profile of memory measures best predicting the decline will be examined. Similarities and/or differences with patterns found for clinically demented subjects will be discussed. See Chapter IV, section 7.

Chapter V:

Chapter V will focus attention on one of the subtests of the battery: the 'Category fluency test'. This subtest may be considered as the most obvious example of a measure reflecting multiple aspects of cognitive functioning (i.e., language, executive functioning, semantic memory). Correspondingly, as was discussed in the literature
Review of clinical and experimental memory testing...

Review presented in part II of this chapter, category fluency performance in AD patients reflects various deficits, more than simply indicated by a ‘sum of correct exemplars’-score. Therefore, Chapter V explores the diagnostic value of the additional scores that can be derived from a category fluency paradigm.

- First, two scores are composed that examine semantic search problems: the number of subcategory clusters (as a measure for ‘switching’), and the number of items per subcategory (as a measure for ‘clustering’).

- Secondly, the degree of integrity of semantic store, which may be characterised by a bottom-up breakdown of semantic knowledge, is investigated by the frequency distribution of the responses produced. This is indicated by the proportion of uncommon (middle- or low-frequency) relative to common (high-frequency) exemplars of a category.

- These aspects of category fluency performance will, first, be investigated in the cognitively impaired (CI) subjects and their matched controls (mNC), known from the first administration of the ‘Category fluency test’ (T₁) (see Chapter V, section 3). Secondly, the development of the various scores over time will be discussed when analysing the data derived from the second administration, two years later (T₂-T₁) (see Chapter V, section 4). It will be investigated whether the various scores show decline or a stable pattern of performance once subjects are diagnosed as demented, relative to the non-demented subjects. In addition, in section 5 of Chapter V, the predictive value of the various scores with respect to dementia will be discussed (i.e., sum of correct responses, number and nature of incorrect responses, scores indicating semantic search problems, and scores indicating the integrity of semantic store).
Appendix: Examples and description of clinical memory tests

Table 1: Different types of clinical memory tests: verbal memory (after Lezak (1995) and *Erickson & Scott (1977)).

<table>
<thead>
<tr>
<th>test category:</th>
<th>test name (authors):</th>
<th>test description:</th>
<th>comment on test:</th>
</tr>
</thead>
<tbody>
<tr>
<td>word span/word list learning</td>
<td>*1. New Word Learning and Retention Test (Shapiro &amp; Nelson, 1955); *2. Modified Word Learning Test (Walton &amp; Black, 1957); *3. Synonym Learning Test (Kendrick, Parboosingh &amp; Post, 1965);</td>
<td>1. learning and retaining (over 24-hour period) the meaning of 5 new words; 2. learning and retaining (over 24-hour period) the meaning of 6 (out of 10) new words; 3. added a recognition procedure to 2;</td>
<td>1. no differentiation between brain-damaged and non-brain-damaged patients; dependent upon level of intelligence; short-term memory factor; 2. little performance overlap between organic and inorganic patients; no correlation with age and intelligence; too tiring and difficult for some elderly patients; 3. high correlation with psychiatric diagnosis; more appropriate for elderly people with dysphasic problems;</td>
</tr>
<tr>
<td></td>
<td>Auditory-Verbal Learning Test (AVLT; Rey, 1964; Taylor, 1959)</td>
<td>5 presentations with recall of a 15-word list, 1 presentation of a second 15-word list and a sixth recall trial; retention after 30 minutes; recognition: identification of words of the first list when shown a list of 50 words containing all the items from both the lists and words that are semantically associated or phonemically similar to words on both lists, or alternate word sets.</td>
<td>immediate memory span; learning curve; learning strategies; retroactive and proactive interference tendencies and tendencies to confusion or confabulation; short-term and longer-term retention; comparison between retrieval efficiency and learning;</td>
</tr>
<tr>
<td></td>
<td>Selective Reminding (Buschke &amp; Fuld, 1974)</td>
<td>presentation of a list of words for immediate recall; on all subsequent stimulus presentations, subjects are only told those words they omitted on the previous trial.</td>
<td>differentiates initial storage, retention and retrieval; scores: long-term retrieval, short-term recall, consistent long-term retrieval, random long-term retrieval, reminders, intrusions;</td>
</tr>
<tr>
<td></td>
<td>California Verbal Learning Test (Delis, Kramer, Kaplan &amp; Ober, 1987)</td>
<td>differs from the AVLT that each of the 16 items in each CVLT list belongs to one of four categories of “shopping list” items; free recall trial and cued recall trial (item categories are used as cues); recognition format;</td>
<td>many different scores; use of learning strategies and their effectiveness (semantic clustering, serial-order clustering); capacity for concept formation; scores comparing free and cued recall and recall and recognition;</td>
</tr>
<tr>
<td></td>
<td>Hopkins Verbal Learning Test (Brandt, 1991)</td>
<td>3 learning trials of 12-word lists, each list containing 4 words from 3 semantic categories which differ for each of the lists; subsequently a 24-word recognition list containing all 12 target words, 6 semantically related foils and 6 unrelated words;</td>
<td>a true positive rate and a false positive rate are obtained from the recognition trial – these provide an accuracy (or discrimination) score and a bias score (tendency to say ‘yes’);</td>
</tr>
<tr>
<td>Test Category: paired-associate word learning</td>
<td>Test Name (Authors): *Inglis Paired-associate tasks (Inglis, 1957)</td>
<td>Test Description: short test battery composed of paired-associate learning tasks—given under 2 learning modalities (auditory/visual) and 2 different modes of reproduction (recognition/recall)—and 4 digit-span tests, employing different learning modalities.</td>
<td>Comment on Test: significant differences between matched memory-impaired and unimpaired elderly patients on the paired-associate tasks (also on the auditory recall alone) but not on the digit span tasks; no correlation with intelligence or age; short-term memory factor; easy and hard pairs are counted separately (both receiving 1 point for each correct response); comparison of recall of well-learned verbal associations and retention of new, unfamiliar verbal material; small age decrements; no ideal verbal memory test: words lend themselves to visual imagery; correct responses for easy pairs receive half points and hard pairs receive whole points; no gender and education effects;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Story Recall: measures both the amount of information that is retained when more is presented than most people can remember on one hearing and the contribution of meaning to retention and recall.</td>
<td>Babcock Story Recall Test (Babcock, 1930; Babcock &amp; Levy, 1940)</td>
<td>21-unit story is used for immediate and delayed recall; after the first recall trial, the story is read again with the message that it has to be reproduced 'in a little while' (20 minutes); when spontaneous recall is not productive, directed recall is allowed by questioning the s.;</td>
<td>May elicit proactive inhibition and/or interference effects and is more a learning test (in comparison with Logical Memory);</td>
</tr>
<tr>
<td>Logical Memory (WMS; WMS-R)</td>
<td>WMS: auditory presentation of 2 stories (A and B); immediate free recall after each presentation; the stories contain 24 and 22 memory units respectively; delayed recall trials after 20 to 60 minutes; WMS-R: addition of a 30-minute delayed recall of the stories; each story contains 25 scoring units; more precise scoring guide; stories changed in more contemporary content and language;</td>
<td>WMS: unstandardised test because of many administration variations and scoring methods; delayed recall more sensitive to variables that affect verbal learning than immediate recall; story A typically elicits better recall than story B, despite the presented order of the stories;</td>
<td></td>
</tr>
<tr>
<td>California Discourse Memory Test (Kramer, Delis &amp; Kaplan, 1988)</td>
<td>2 stories, each 29 units for verbatim scoring and 5 main ideas for gist scoring; immediate and delayed recall trials; cued and multiple-choice recognition trial; structured format for error analysis;</td>
<td>Groups may differ in recall of verbatim units, recall of gist, amount of material lost on delay and recognition.</td>
<td></td>
</tr>
<tr>
<td>Test Category:</td>
<td>Test Name (Authors):</td>
<td>Test Description:</td>
<td>Comment on Test:</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Design Learning</td>
<td>Rey's Visual Design Learning Test (Rey, 1968)</td>
<td>Like the AVLT, ss. are shown (in 5 trials) a series of 15 items, each item containing 2 elements that are very verbalisable (e.g., a dot in a circle, a triangle above a horizontal line); after each trial ss. must draw all the designs that they can remember regardless of the order; recognition test immediately after the learning trials; no interference trial;</td>
<td>Decline in performance as from the 30s, dropping between .75 to 1.75 points in each subsequent decade; slight drop of 1.15 points between youngest group’s recognition score and that of 70+ group implicates retrieval instead of learning problem for the elderly.</td>
</tr>
<tr>
<td></td>
<td>Corsi Blocks: Supraspan techniques (Milner, 1971)</td>
<td>24 test trials of tapping sequences, 1 (or 2) tap(s) greater than the patient’s immediate span (as measured by the Corsi Block-tapping Test); the same sequence is repeated every third trial;</td>
<td>Not sensitive to effects of normal aging (ss. aged 20-75-years); pts. with right temporal lesions (including hippocampus) showed no learning;</td>
</tr>
<tr>
<td></td>
<td>7/24 Spatial Recall Test (Rao et al., 1984)</td>
<td>7 poker chips are randomly placed on a 6x4 checkerboard; 5 learning trials of a array, each with a 10 second exposure; single learning trial of a different pattern; first free recall of first array; 30 minutes later delayed recall trial;</td>
<td>Learning score: total number of correctly placed chips; keen eyesight and good motor control not required; too verbalisable for highly performing ss.;</td>
</tr>
<tr>
<td></td>
<td>Form Sequence Learning (Hamsher, Roberts &amp; Benton, 1987)</td>
<td>Ss. are shown a horizontal array of geometric designs for 30 seconds; following exposure of the stimuli, a double horizontal array of figures is presented, containing as many foils as targets; ss. must select the items just seen in the order they had appeared; the array increases from 2 or 3 items to 4 to 6 item sequences;</td>
<td>2 points for a perfect score, 1 point for a single substitution or reversal of a figure; scores can be derived for sequencing and perceptual accuracy; test appeared not sensitive to side of temporal lobe damage;</td>
</tr>
<tr>
<td></td>
<td>Biber Figure Learning Test (Glosser, Goodglass &amp; Biber, 1989)</td>
<td>10 test items (each composed of 2 geometric figures) and 30 distractors (3 for each target figure, of which one differs in orientation, one in its shape and one quite different); after all 10 test items are shown (for 2 seconds each), ss. are asked to draw them from memory, for 5 learning and recall trials; recognition is tested immediately after these trials; 20 minutes later delayed free recall and delayed recognition;</td>
<td>Age effects found on free recall but not on recognition; significant verbal learning component (because of low retention score by left brain damaged patients and high correlation with a verbal learning test);</td>
</tr>
<tr>
<td></td>
<td>Visual Spatial Learning Test (Malec, Ivanik &amp; Hinkeldey, 1991)</td>
<td>Use of a 6x4 grid and 7 stimulus items placed on the grid, providing different nonsense designs that are difficult to verbalise; after presentation of each pattern of designs, ss. must reproduce it with an empty grid and 15 designs; selecting the correct 7 designs and placing them as they were when seen on the grid; 5 learning trials, 30-minute delayed recall;</td>
<td>Performance scored for recognition learning of the designs, recall of the target positions on the grid and recall of designs in their proper places; age effects found for learning and delay trials; worse performance by right (vs. left) temporal lobe pts.;</td>
</tr>
<tr>
<td>test category</td>
<td>test name (authors)</td>
<td>test description</td>
<td>comment on test</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------</td>
<td>------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>paired-associate learning</td>
<td>Non-Language Paired Associate Learning (Fowler, 1969)</td>
<td>stimuli are objects, which are arranged into 6 sets of easily associated pairs and 4 hard-to-associate pairs. After being shown all pairs, ss. must pick out from a set of 22 items those that had been shown with a paired item when the first item of the pair is presented; presentation order varies from trial to trial;</td>
<td>lower scores for “hard” words (Associate Learning, WMS) than for “hard” object pairs for all age groups; no difference for “easy” combinations;</td>
</tr>
<tr>
<td></td>
<td>Visual Paired Associates (WMS-R):</td>
<td>8 nonsense line drawings are shown, each paired with a square of a different colour; on recall, the designs are shown in a different order and ss. must name the colour that goes with each design; 3 learning trials and a 30-minute delay trial;</td>
<td>several items are readily verbalised (several pairs of colour and design are easily associated); visual-verbal learning test instead of visual learning test; biggest age decline between 35-44 and 55-64 ranges;</td>
</tr>
<tr>
<td>design reproduction</td>
<td>Complex Figure Test (Rey, 1941; Osterrieth, 1944)</td>
<td>immediate and delayed recall of a complex line drawing; ss. are not forewarned at the copy trial that they will be asked to draw the figure from memory.</td>
<td>susceptible to verbalisation; significant age effects; high education delays decline until after age 69; men scored better than women; left-lesioned pts show preserved recall of overall structure with loss of details; right-lesioned pts show even more problems with recall than with copying;</td>
</tr>
<tr>
<td></td>
<td>Visual Reproduction (WMS; WMS-R):</td>
<td>3 (or 4) cards with printed designs are shown for 5 seconds each; ss. draw what they remember of the design after every presentation; immediate recall; sometimes followed by a 30 minutes delay trial;</td>
<td>better scoring guidelines in revised version; largest age effects of all WMS(-R) tests; verbal encoding possible; no pronounced differences between performances of pts. with right- vs. left-sided lesions</td>
</tr>
<tr>
<td></td>
<td>Memory for Designs Test (Graham &amp; Kendall, 1960)</td>
<td>15 geometric designs are shown for 5 seconds each; after every presentation, ss. draw what they remember of the design.</td>
<td>correlated with age and intelligence; possible confounding by perceptual and motor deficits;</td>
</tr>
<tr>
<td></td>
<td>Benton Visual Retention Test (Benton, 1974)</td>
<td>10 cards containing 1 to 3 geometric figures are shown for 10 seconds each; subjects draw after every card presentation what they remember of the design.</td>
<td>widely used; sensitive to unilateral spatial neglect and spatial organisation problems; poor discriminant of brain damage; test of visuomotor or perceptuomotor performance instead of immediate visual memory; possible correlation with intelligence and age;</td>
</tr>
<tr>
<td>test category:</td>
<td>test name (authors):</td>
<td>test description:</td>
<td>comment on test:</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>recognition</td>
<td>Recurring Figures Test (Kimura, 1963)</td>
<td>uses 20 drawings of geometric or irregular nonsense figures; after presentation of all these cards, ss. must recognise them from 140 cards, each shown for 3 seconds; false positives are subtracted from right positives to correct for guessing;</td>
<td>no difference in scores of right and left temporal lobe pts. (but right lesioned pts. had twice as many false positives); both groups remembered geometric figures better than nonsense figures;</td>
</tr>
<tr>
<td></td>
<td>Visual Retention Test (Metric Figures) (Warrington &amp; James, 1967a)</td>
<td>stimuli are 5x5 inch squares each divided into 25 squares of which four are blackened and variously positioned so that no 2 stimulus figures are alike; 20 target figures are shown for 2 seconds each; after each presentation ss. must identify the target from among 3 other, similar figures; a second administration provides a 10 seconds exposure with the figures rotated 180°.</td>
<td>error scores are counted; total error scores did not differentiate between LH and RH pts; on 2 sec. administration, pts with right parietal lesions made more errors than pts with left parietal lesions: role for unilateral visuospatial inattention;</td>
</tr>
<tr>
<td></td>
<td>Continuous Recognition Memory Test (Hannay, Levin &amp; Grossman, 1979)</td>
<td>presentation of 120 line drawings of various flora and fauna for 3 seconds each; ss. must decide whether it is identical to one already seen or a different drawing; scores: hits, false alarms, missings, perceptual discrimination;</td>
<td>age and education effects in normal ss.; discriminates moderate and severe head injury (not mild injured pts.);</td>
</tr>
<tr>
<td></td>
<td>Figural Memory (WMS-R)</td>
<td>4 trials of 1 to 3 abstract rectangular designs in white and two shades of grey, fitted into squares of equal sizes; after each trial ss. must identify the target design(s) from an array of similar designs; no delayed trials;</td>
<td>non-verbalisable stimuli; small but consistent age effects; 65+ ss. show little variability in scores (leading to limited discriminability); did not discriminate between pts. with left- and right-lateralised lesions;</td>
</tr>
<tr>
<td></td>
<td>Continuous Visual Memory Test (Trahan &amp; Larrabee, 1988)</td>
<td>same as Continuous Recognition Memory Test, but uses more abstract stimulus designs; delayed recognition trial after 30 minutes;</td>
<td>age effects characterised by increase in false alarms; sensitive to brain damage, especially to right-sided lesions;</td>
</tr>
</tbody>
</table>
Table 3: Different types of clinical memory tests: tactile memory (after Lezak (1995)).

<table>
<thead>
<tr>
<th>test category:</th>
<th>test name (authors):</th>
<th>test description:</th>
<th>comment on test:</th>
</tr>
</thead>
<tbody>
<tr>
<td>tactile memory</td>
<td>Tactual Performance Test (Halstead, 1947; Reitan &amp; Wolfson, 1993)</td>
<td>blindfolded ss. complete 3 trials on a formboard (with preferred hand, the non-preferred hand and with both hands respectively); after the formboard trials, ss. must draw the board from memory (the shapes and their placement relative to one another); score on each formboard trial is the time to completion; drawing trial: memory score (number of correct shapes reproduced) and location score;</td>
<td>significant age effects; LH-pts tend to perform better than RH-pts; drawbacks: experienced psychological distress when blindfolded, amount of time (especially for older and brain injured pts) and equivocal nature of data obtained;</td>
</tr>
<tr>
<td></td>
<td>Tactile Pattern Recognition (Milner, 1971)</td>
<td>uses 4 pieces of wire, each twisted into a distinctly different nonsense shape; several training trials on matching the figures (without seeing them); matching after delay period (with distraction task);</td>
<td>commissurotomised pts performed better with their left than with their right hand (learning can take place without words and is mediated by the RH);</td>
</tr>
</tbody>
</table>

Table 4: Different types of clinical memory tests: remote memory (after Lezak (1995)).

<table>
<thead>
<tr>
<th>test category:</th>
<th>test name (authors):</th>
<th>test description:</th>
<th>comment on test:</th>
</tr>
</thead>
<tbody>
<tr>
<td>remote memory</td>
<td>Famous People Tests (Albert, Butters &amp; Levin, 1979; Albert, Butters &amp; Brandt, 1981a,b; Butters &amp; Albert, 1982)</td>
<td>tests of recall or recognition of famous persons: Facial Recognition Test (25 photographs of famous persons from each of six decades (1920s-1970s)), Old-Young Test (old and young versions of the same famous people, questionnaires about famous people from these decades (recall and recognition);</td>
<td>Korsakoff pts showed a marked gradient: low scores for recent material and (nearly) normal scores for material from early decades;</td>
</tr>
<tr>
<td></td>
<td>Presidents Test (Hamsher &amp; Roberts, 1985)</td>
<td>4 different administrations: verbal naming (VN; free recall of the current president and his 5 immediate predecessors), verbal sequencing (VS; chronologically arranging 6 cards with the president’s names), photo naming (PN; naming of 6 pictures, presented in non-chronological order) and photo sequencing (PS; chronologically arranging the 6 pictures);</td>
<td>VN: age effect for ss. with &lt;12 years of schooling; PN/PS: only educational effects; VS: no age or educational effects; more pts with RH lesions failed the sequencing tests than pts with LH lesions; significant relation between general cognitive deterioration and number of task failures;</td>
</tr>
<tr>
<td></td>
<td>Remote Memory Battery (Beatty et al., 1988)</td>
<td>identification of pictures of 15 persons famous in each decade from the 1940s to the 1980s and answering to 15 public events questions from each of these decades; items that are failed are readministered with a card giving 4 alternative responses;</td>
<td>cognitively deteriorated pts (according to MMSE-scores) tended to recall less information from the 1980s: represents genuine time gradient or just decreased mental efficiency or diminished energy for news items (events in the outside world);</td>
</tr>
<tr>
<td></td>
<td>Autobiographical Memory Interview (Kopelman, Wilson &amp; Baddeley, 1989)</td>
<td>2 sections: autobiographical incidents schedule and personal semantic memory schedule; 3 questions from 3 time blocks: childhood, early adult life and recent events; prompts are given when pts cannot respond; answers are graded on a 0-3 scale;</td>
<td>amnesic pts performed significantly worse than control ss. on all variables (greatest difference on recent memory);</td>
</tr>
</tbody>
</table>
Table 5: Memory batteries, containing several types of memory tests (after Lezak (1995) and *Erickson & Scott (1977)).

<table>
<thead>
<tr>
<th>test category:</th>
<th>test name (authors):</th>
<th>test description:</th>
<th>comment on test:</th>
</tr>
</thead>
<tbody>
<tr>
<td>memory batteries</td>
<td>*Babcock Test of Mental Efficiency (Babcock &amp; Levy, 1940)</td>
<td>includes several tasks involving recall of previously learnt material.</td>
<td>correlated with intelligence; learning and retention of paired associates most impaired by aging; complex and time consuming test;</td>
</tr>
<tr>
<td></td>
<td>*Wechsler Memory Scale (WMS; Wechsler, 1945)</td>
<td>7 subtests: Personal and Current Information, Orientation, Mental Control, Logical Memory, Memory Span, Visual Reproduction and Associate Learning; computation of ‘Memory Quotient’ (MQ), based on the sum of all tests (as a representation of memory abilities);</td>
<td>using MQ incorrect: it assumes that memory is a unidimensional function, it incorporates an overly inclusive concept of memory (including subtests as personal information or orientation) and it is insensitive for disease specific patterns of certain memory function deficits; high correlation of MQ with IQ; relies on verbal skills;</td>
</tr>
<tr>
<td></td>
<td>*Wechsler Memory Scale-Revised (WMS-R; Wechsler, 1987)</td>
<td>successor of the WMS; extended orientation/information section from the WMS; addition of delayed-recall versions of the subtests Logical Memory, Paired-Associate Learning and Visual Reproduction and of tests of spatial memory span and visual paired-associate learning and a test of recognition for geometric visual patterns;</td>
<td>improvements: correction of unitary MQ score (different indexes), visual vs. verbal memory tests, delayed recall measures and norming procedures; only one form (unlike WMS); the indexes reflect some disease-associated patterns of memory impairment;</td>
</tr>
<tr>
<td></td>
<td>*Cronholm Battery (Cronholm &amp; Ottosson, 1963)</td>
<td>30 word-pair test (conceptually related words, not frequently associated) and 20 (or 30) figure test (20 drawings: identification in an array of 50 drawings; 30 personal data test: drawings of 6 fictitious persons with 5 items of information associated with each person – testing for immediate recall and recall after several hours).</td>
<td>measures learning and retention; 20 figure test allows a verbally impaired patient to respond; correlation with intelligence was not investigated;</td>
</tr>
<tr>
<td></td>
<td>*Williams Scale for the Measurement of Memory (Williams, 1968)</td>
<td>immediate recall (digit span), nonverbal learning (pegboard task), verbal learning (word-definition learning from the Modified Word Learning Test), delayed recall (of 9 pictures, after 7-10 minutes, followed by cued recall and recognition dependent on the ss.’s level of performance) and memory for past personal events (features of ss.’s first school).</td>
<td>delayed picture recall best at discriminating between psychiatric and organic patients and was least affected by age and intelligence; the pegboard task was poorest at predicting cerebral pathology and was most affected by age and intelligence; user-friendly;</td>
</tr>
<tr>
<td></td>
<td>Guild Memory Test (Gilbert, Levee &amp; Catalano, 1968)</td>
<td>immediate and delayed recall of 2 stories and of 10 word pairs, digits forward and reversed, visual recognition; 2 forms; norms for 5 age ranges; first American memory battery that maintains scoring and interpretation differences between the memory functions assessed;</td>
<td>all tests (except Digit Span) normal decline with age; scores differentiate normal young, normal old and impaired (mild to moderate) elderly ss;</td>
</tr>
</tbody>
</table>
Table 5 (continued)

<table>
<thead>
<tr>
<th>test category:</th>
<th>test name (authors):</th>
<th>test description:</th>
<th>comment on test:</th>
</tr>
</thead>
<tbody>
<tr>
<td>memory batteries</td>
<td>Denman Neuropsychology Memory Scale (Denman, 1984, 1987)</td>
<td>4 ‘verbal’ tests (story recall (also delayed trial); paired associate learning (also delayed trial); remote verbal information; digits forward and backward); 4 ‘nonverbal’ tests (figure recall (CFT); tonal matching task; facial recognition task; questions about visual details of familiar objects, signs, etc.); only 1 form;</td>
<td>extensive age norms, but no educational norms (ss. were highly educated); no consistent patterns shown in factor analytic studies; scores available of several neurological pts; drawbacks concern ‘verbal’ vs. ‘nonverbal’ interpretation and scoring;</td>
</tr>
<tr>
<td></td>
<td>Rivermead Behavioural Memory Test (Wilson, 1986; Wilson, Cockburn &amp; Baddeley, 1985)</td>
<td>12 subtests: recalling a name, spontaneously remembering where something has been hidden and asking a specific question in response to a cue (tests of prospective memory), recognition of ten recently shown pictures, immediate and delayed recall of a story, recognition of five unfamiliar faces, immediate and delayed recall of a route, remembering a message and questions about orientation in time, place and person;</td>
<td>tapping everyday memory demands for brain-damaged people; high face validity; 4 parallel forms; norms for 16-69 age range; no age or gender effects; related to mental ability (Raven, NART); norms for 70-94 age range (scores were lower than of 16-69 range); discriminates memory problems of moderately to severely injured head trauma pts;</td>
</tr>
<tr>
<td></td>
<td>Randt Memory Test (Randt &amp; Brown, 1986)</td>
<td>recall of 5 words using selective reminding technique, digits forward and backward, recall of word pairs, recall of a paragraph, recognition and name recall of 7 out of 15 line drawings of common objects; use of telephone interviews to obtain 24-hour recall data; administration takes 20 minutes;</td>
<td>specifically designed for longitudinal studies of pts with mild to moderate impairment of storage and retrieval functions; ceiling effects expected with younger ss. (with memory problems); useful with aging and diffuse brain diseases;</td>
</tr>
<tr>
<td></td>
<td>Memory Assessment Scales (Williams, 1991)</td>
<td>addresses 3 kinds of memory functions: attentional functions and short-term memory (STM), learning and immediate memory, memory following delay; examined in verbal and nonverbal modalities and 1 test involves integration of verbal (names) and nonverbal (faces) material;</td>
<td>nonverbal tests are fairly verbalisable; 3 factors for neurologically impaired pts: nonverbal memory and reasoning, STM and concentration, verbal memory; dementia pts: scored worse than all other groups on all scores except Visual Memory;</td>
</tr>
<tr>
<td></td>
<td>Learning and Memory Battery (Tombaugh &amp; Schmidt, 1992)</td>
<td>7 tests: Paragraph, Word List and Word Pairs (using selective reminding and delayed recall), Digit Span (forward and backward), Supraspan, Simple Figures, Complex Figure;</td>
<td>norms (for 8 age ranges) available for every trial for every test; redundant findings; not likely that distinguishing data will emerge from different tests;</td>
</tr>
<tr>
<td></td>
<td>Doors and People Test (Baddeley, Emslie &amp; Nimmo-Smith, 1994)</td>
<td>4 main subtests: ‘People’ (verbal cued recall), ‘Shapes’ (visual recall), ‘Names’ (verbal recognition) and ‘Doors’ (visual recognition); yields overall episodic memory score; names in verbal tests prevents visual encoding; crosses in Shapes test prevents verbal coding; recognition of doors prevents verbal encoding;</td>
<td>specific retrieval deficit measured by comparing recall and recognition scores; visual and verbal measures allow test for material-specificity; recall measures (3 successive learning trials and delayed recall) provide standardised measures of learning and forgetting; high face validity;</td>
</tr>
</tbody>
</table>