Université de Montréal

Neuropsychological and Neuroendocrine Correlates of Subjective Memory Complaints in Seniors With or Without Depressive Symptoms

Par Marc Lemay

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiae Doctor (Ph.D.)
en Psychologie

novembre 2001

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée:

Neuropsychological and Neuroendocrine Correlates of Subjective Memory Complaints in Seniors With or Without Depressive Symptoms

présenté par:

Marc Lemay

a été évaluée par un jury composé des personnes suivantes:

Sylvie Belleville

Serge Gauthier

Marie-Jeanne Kergoat

Sonia Lupien

Maurice Ptito

Thèse acceptée le: 21 décembre 2001

Abstract

In the search for earliest evidence of Alzheimer's Disease (AD), the prognostic and diagnostic value of subjective memory complaints in seniors has been much studied, with inconsistent results. While most of the literature shows that memory complaints are a poor predictor of memory performance, being rather related to depressive or other dysphoric symptoms, a few reports do describe memory complaints as at least partly reflecting accurate self-appraisal of memory abilities. Given that memory complaints in seniors are ubiquitous, easily solicited, and of potential informative value, this inconsistency is worth investigating.

Many psychometric tests relating to AD are informative but too laborious to be incorporated into medical practice. In the spirit of the discipline of neuropsychology, which bridges psychology and medicine, this project combined a quick and easy memory complaints and depressive symptom assessment procedure, with exhaustive memory testing and biological measures, in order to characterize a self-selected sample of seniors with memory complaints, depressive complaints, or both.

Subjects were recruited from the community. Exclusion criteria were light and mainly centered on serious psychiatric conditions or inability to understand French or English. Fifty-two subjects, three quarters of them women, were evaluated.

The specific predictions were that depressed subjects would recall fewer negative- than neutral-valence words in a paragraph recall task; that subjects reporting both memory problems and depressive symptoms would be highest in daily life stress; that depressed subjects would have the highest cortisol levels; that high cortisol levels would be inversely correlate to declarative memory

scores; and that subjects with both complaints and depression would have the highest cortisol levels.

Results supported the first two of these hypotheses. Depressed subjects showed a cognitive bias against recall of negative-valence information. Subjects with memory complaints and depressive symptoms had the highest levels of stress, but most of the group differences could be accounted for by traumatic symptoms, which were highest in the complaints-depressed group. The hypotheses about cortisol and cognition or cognition and depression were not supported, possibly because of too-great individual variability, inadquate adherence to study instructions for the collection of saliva samples for cortisol assays, or inadequate sampling rate or length. However, subjects with memory complaints above median showed an abnormal morning cortisol response.

Subjective memory complaints appear to be an uncertain marker of cognitive impairment, but a more reliable and important marker of "psychoneuroendocrine frailty" characterized by subtle cognitive impairment, slight abnormalities in cortisol regulation, and—when accompanied by depressive symptoms—high levels of daily stress and possibly, traumatic symptoms as well.

Résumé en français

Des considérations cliniques et éthiques nous conduisent à souhaiter intervenir le plut tôt possible dans le cours de la maladie d'Alzheimer. L'identification précoce d'un cas putatif de cette maladie présente des difficultés qui augmentent d'autant plus que l'on fixe le seuil qui déclenche une intervention près de la normalité cognitive. Des découvertes récentes concernant la perte de neurones dans le cortex entorhinal chez des sujets ayant des déficits cognitifs à la limite du détectable nous obligent à une réévaluation des idées reçues concernant les paramètres du vieillissement cognitif normal. En effet, ces informations apparaissent inquiétantes si l'on considère par exemple que les banques de données normatives actuelles peuvent être contaminées par des cas insoupçonnés d'Alzheimer. De plus, nous ne possédons toujours pas de marqueur biologique ante-mortem valide de cette maladie chez des gens relativement en bonne santé. Tous ces facteurs nous orientent vers une stratégie de recherche d'indicateurs précoces de l'Alzheimer, qui utiliseraient de l'information obtenue par une approche multidimensionnelle. Pour être applicable à la réalité clinique, des tests de dépistage se doivent d'être simple d'application et sans lourdeur technique.

Cette étude est conçue selon une approche neuropsychoendocriniennne basée sur les plaintes subjectives de mémoire couplée à la mesure du cortisol. La décision de retenir la mesure du cortisol est fondée sur la prémisse que cette hormone de stress est reconnue pour pouvoir prédire le déclin cognitif. Ainsi, on observe chez les sujets démontrant des concentrations élevées de cortisol qui continuent à augmenter au fil des années, des déficits de mémoire déclarative en comparaison à des sujets appariés dont les concentrations cortisolémiques sont basses ou stables. Cette étude vise à élucider les rapports

existants entres les plaintes subjectives de mémoire, les symptômes dépressifs, et la cognition chez des sujets âgés. Au cours de cette étude, 52 volontaires vivant dans la communauté, ont été recrutés et soumis à une évaluation neuropsychoendocrinienne incluant des tests cognitifs, des questionnaires portant sur des dimensions sociales et la personnalité, ainsi qu'une analyse cortisolémique provenant d'échantillons salivaires. Une fois l'évaluation de base complétée, les sujets ont été classifiés selon leur appartenance à l'un de ces trois groupes : plaintes subjectives de mémoire seulement (groupe Mem) ; plaintes subjectives de mémoire accompagnées de symptômes dépressifs (MemDep) ; sans plainte de mémoire ni symptôme dépressif (groupe Témoin).

Le groupe Mem avait des performances inférieures à celles des deux autres groupes sur une mesure de mémoire visuelle. Les sujets dont les plaintes subjectives de mémoire étaient supérieures à la médiane, ont obtenu un taux de rappel inférieur dans une tâche de rappel d'une courte histoire, et démontrait ausi un déclin abnormal de cortisol salivaire dans les 30 minutes suivant l'éveil. Le groupe MemDep rapportait beaucoup plus de symptômes de stress quotidien et de stress post-traumatique que les deux autres groupes. Cette étude a démontré qu'en présence de plaintes subjectives de la mémoire, les symptômes dépressifs concomitants avaient très peu d'effet sur les performances cognitives sauf pour diminuer le rappel libre et augmenter la reconnaissance des éléments de valence négative dans une tâche de rappel d'une courte histoire. L'importance des plaintes subjectives de mémoire en tant que signe précurseur de l'Alzheimer apparaît donc douteuse. Il demeure toutefois que ces plaintes subjectives de mémoire reflétaient souvent soit : des symptômes dépressifs, un niveau excessif de stress quotidien ou la présence d'un syndrome de stress post-traumatique. En conclusion, la plainte

subjective de mémoire serait davantage un reflet d'une "fragilité psychoneuroendocrinienne." Elle associerait dans sa présentation : un dérèglement cortisolémique léger, un trouble léger de mémoire déclarative, et si elle s'accompagne de symptômes dépressifs, un niveau élevé de stress quotidien ou de stress post-traumatique.

Summary

The ethical and rational goal of medical care for AD is obviously to intervene as early as possible in the course of the disease, but the difficulty of identifying patients increases the closer to normality the line is drawn. Recent discoveries about the loss of entorhinal cortex neurons in subjects with borderline-detectable cognitive impairment are forcing a re-evaluation of the parameters of normal cognitive aging, and by the same token render suspect existing normative data sets, which are almost certainly contaminated by cases of unsuspected incipient AD. Depression, which is sometimes seen as a confound in the diagnosis of AD, may rather constitute one of its prodromal symptoms. Finally, a reliable biological marker of AD in relatively healthy people has yet to be found. All these factors have inspired the search for a possible combined marker of incipient AD, one that would be simple in application in order to be integrated into clinical practice. For this study a neuropsychological and neuroendocrine approach was chosen, centered on subjective memory complaints and salivary cortisol measures. Cortisol, a stress hormone, is known to predict cognitive impairment compared in seniors; subjects with high cortisol at two time points separated by years, or whose cortisol concentrations increase from one time point to the next, have slightly impaired declarative memory compared to seniors whose concentrations are consistently low or dropping. In order to clarify the link between subjective memory complaints, depressive symptoms and cognition in non-demented seniors, as well as study the relation of neuroendocrine status to cognitive performance in seniors with subjective memory complaints, 52 volunteers were recruited from the community and subjected to a psychoneuroendocrine evaluation comprising cognitive, social

and personality tests and questionnaires along with salivary cortisol sampling. Once testing had been completed, volunteers were classified in 3 groups: people with subjective complaints only (Mem group); people with both subjective memory complaints and depressive symptoms (MemDep); and people with neither memory complaints nor depressive symptoms (Control). The Mem group was impaired relative to the two other groups on a measure of visual memory. Subjects with memory complaint scores above the median recalled fewer items in a paragraph recall task. The MemDep group had impaired recall but enhanced recognition of negative-valence story elements in that same task, and also reported higher levels of stress and traumatic symptoms than the other two groups.

This study demonstrated that given the presence of memory complaints, depressive symptoms were shown to have very little effect on cognitive performance except to decrease recall and increase recognition of negative-valence story elements in a paragraph recall task. The clinical significance of subjective memory complaints as a sign of cognitive at-risk status (AD or other dementing process) is doubtful. What is less doubtful is that memory complaints are often associated with depression and high levels of stress and possibly post-traumatic stress disorder. Thus subjective memory complaints may perhaps best be conceptualized not only as more or less accurate indicator of cognitive functioning, but more particularly as a marker of "psychoneuroendocrine frailty." This type of mental and physical frailty comprises slightly elevated cortisol secretion and slightly lowered cognitive performance, and—when combined with depression—greatly increased stress levels and traumatic symptomatology.

Table of Tables

Table of Figures

Abbreviations:

The three study groups:

Mem: Subjects with subjective memory complaints

MemDep: Subjects with both subjective memory complaints and depressive symptoms

Control: Subjects with neither subjective memory complaints nor depressive symptoms

Other abbreviatons:

3MS: Modified Mini Mental State Exam

ACTH: Adrenocorticotropic hormone

AD: Alzheimer Disease

ADAS-cog: Alzheimer Disease Assessment Scale—cognitive subscale

CRH: Corticotrophin-releasing hormone

CDR: Clinical Dementa Rating

CIND: Cognitive Impairment, No Dementia

CSHA: Canadian Study of Health and Aging

DEX: Dexamethasone

DSM-III-R, DSM-IV: Diagnostic and Statistical Manual, third revised or fourth edition

EC: Entorhinal cortex

GDS: Geriatric Depression Scale

IADL: instrumental activities of daily living

LTP: Long-term potentiation

MAC-Q: Memory Assessment Clinics Questionirere

MCI: Mild Cognitive Impairment

MEMs: Memory complaints

MMSE: Mini-Mental State Exam

MRI: Magnetic Resonance Imaging

MUML: Medically unexplained memory loss

NFT: neurofibrillary tangles

NPV: Negative Predictive Value

OCD: Obssessive-compulsive disorder

OCS: Obsesessive-compulsive subscale from Symptom Checklist-90

PPV: Positive Predictive Value

RT: Reaction Time

ROC: Receiver Operating Curves

SADAS-cog: Standardized Alzheimer Disease Assessment Scale—cognitive subscale

SCL-90: Symptom Checklist-90

SPs: senile plaques

Acknowledgments

My debts are vast. Thanks to André Delorme, whose kindness and intelligence were a first inspiration; to Marie-Jeanne Kergoat, who trusted me and offered a peer relationship that made me strive to live up to it; to Sonia Lupien, who guided me and encouraged me when I most needed it; to the students, professors and staff at the Department of Psychology of the *Université de Montréal* and the research centre of the *Institut universitaire de gériatrie de Montréal*, who together made for a stimulating and fun place in which to work, learn and grow. Most of all, thanks to Nadine, who endured much and loved more than words can tell. You are the love of my life, my guiding star.

For Nadine and Liam.

1. Introduction

Mild cognitive impairment (MCI) refers to a postulated intermediate stage between normal aging and dementia. It is the subject of intense research because of the potential it offers for early intervention. Although the nosology, neuropsychology and operational definition of MCI remain the subject of debate, the common goal has been to develop a definition that would as include as much as possible progressive, presumed incipient Alzheimer Disease (AD) cases while excluding benign, non progressive cases. Most MCI definitions include a requirement for memory complaints. Memory complaints are a reported judgment on the part of the patient, or an informant who has known the patient for a number of years, that the patient's memory is not as good as it once was. There is converging evidence that medically unexplained memory loss without dementia is an early symptom of Alzheimer Disease. But what is less clear is the value of subjective memory complaints. They certainly have the virtue of simplicity, with the resulting potential for cost savings in an overburdened health care system. On the other hand, the very ease with which memory complaints can be elicited and expressed in the context of the fear of Alzheimer Disease risks creating an avalanche of new patients, due to the pathologizing of normal aging; it also could divert limited resources from other non-memory-related conditions really at the root of those complaints. Memory complaints can be a signal that the patient has, in fact, experienced a certain degree of cognitive

deterioration compared to her or his young adult years. Such deterioration can represent normal aging, or incipient AD; another possibility is that the complaints do not signal memory impairment per se but are rather secondary to some other non-cognitive problem. Indeed, the chief controversy in the literature with regards to memory complaints is their relation to depression as opposed to, or in addition to, their relation to MCI.

This project will therefore concern itself with the relationship of memory complaints, objective memory performance, and depressive symptoms. The specific goal of this project is to determine whether the presence of concomitant depressive symptoms alters the validity of memory complaints—that is, their relation to objective performance—relative to memory complaints unaccompanied by depressive symptoms. But this goal must be seen within the broader context alluded to above, of the challenge of differentiating benign, non-progressive MCI from MCI as incipient Alzheimer Disease. In this project I have not studied subjects with MCI (although memory complaints are part of most definitions of MCI), first because there is no consensus on an operational definition (which creates an opportunity for useful research), and second because existing definitions are too restrictive (excluding the great majority of potential subjects) and would have made this project unfeasible. I will explore the various definitions of MCI definitions at some length in the first sections of this thesis, because the originality of this thesis consists in testing the

validity of memory complaints as a marker of psychoneuroendocrine abnormality, which could in turn be incorporated into a new definition of MCI. Memory complaints are common and can be elicited with varying degrees of seriousness from a good proportion of adults middle aged and older; on the other hand most MCI definitions are so stringent as to produce prevalence rates well below that of AD itself. After surveying the field of MCI research and the maneuvers used to detect incipient AD, and presenting my own methods and results, I will discuss their implications and argue that both these approaches—the inclusive and the exclusive—have their uses and important applications in different milieus.

2. The Nature of the Problem

In the search for earliest evidence of Alzheimer's Disease (AD), the prognostic and diagnostic value of subjective memory complaints in seniors has been much studied, with inconsistent results. Some researchers have found that subjective memory complaints (MEMs) relate more to depressive or other dysphoric symptoms than to objective memory or cognitive performance; but others have found that MEMs do reflect accurate self-appraisal of memory abilities. Given that memory complaints in seniors are ubiquitous, easily solicited, and of potential informative value, this inconsistency is worth investigating. The goal of this project is therefore to test and elucidate the effect that the presence of concomitant depressive symptoms have, if any, on the validity of subjective memory complaints, and in so doing to generate

hypotheses that would later be verified during the subjects' follow-up evaluations.

Neuropsychological research projects on cognition in seniors often make use of stringent inclusion and exclusion criteria intended to make the interpretation of results possible, but which can also make recruitment of adequate sample sizes difficult or impossible, and cause problems with the generalization of results, since strict selection criteria vastly complicates the study of a given disorder as it actually appears in the population. Therefore, in this project we sought to study and describe a self-selected volunteer sample of seniors with subjective memory complaints, using light exclusion criteria mainly limited to serious psychiatric conditions, and inability to understand basic French or English.

Many psychometric tests relating to AD are informative but too laborious to be incorporated into medical practice. In the spirit of the discipline of neuropsychology, which bridges psychology and medicine, this project will combine a quick and easy memory complaints and depressive symptoms assessment procedure, with exhaustive memory testing as well as a stress hormone measure, in order to characterize a self-selected sample of seniors with memory complaints, depressive complaints, or both. This project is also designed to contribute to an ongoing longitudinal study (whose principal investigator is my thesis co-advisor, Dr. Sonia J. Lupien) of the effects of stress and aging on memory and other neuropsychological functions, by orienting

future subject evaluations towards outcome measures and dimensions which will have proved most informative in the present investigation.

3. Theoretical and Empirical Context

By 2030 the Canadian population is projected to increase by a factor of 1.4, while the proportion of those aged 65 years or more is expected to double, and the number of cases of Alzheimer's Disease to triple. The burden of care facing individuals and society will obviously increase as the number of AD cases and AD prevalence rates increase (Small, Rabins et al 1997). Cases of mild cognitive impairment *without* dementia, which at present affects about 17% of seniors in Canada, more than all the dementias combined (8%) (Graham, Rockwood, Beattie, Eastwood, Gauthier, Tuokko and McDowell, 1997) will also increase as the population undergoes demographic aging. The mild cognitive impairment of many of these seniors may be treatable or preventable (Riedel and Jolles, 1996). However, much of the difficulty and the opportunity of research in this domain stems from the unresolved nosology of mild cognitive impairment, and the lack of consensus on a workable operational definition including specificiation of neuropsychological tests (Ritchie and Touchon, 2000). Emblematic of most MCI definitions extant, the definition proposed by Petersen and colleagues (Petersen, Smith, Waring, Ivnik, Tanalos and Kokmen, 1999) has been criticized as too stringent: its application in one sample of seniors with observable cognitive deterioration over 3 years led to the counter-intuitive finding of an MCI

prevalence lower than that of AD itself (Ritchie, Artero and Touchon, 2001). In this thesis the term MCI will be used without reference to any specific operational definition (unless otherwise specified), and will be used to mean *unexplained* (i.e. not clearly attributable to alcoholism, cultural differences, mental retardation) mild impairment of mental function (whether of memory or other cognitive functions) associated with advancing age, and which may or may not represent a prodromal stage of Alzheimer's Disease.

The ethical and rational goal of medical care for AD is obviously to intervene as early as possible in the course of the disease; but the difficulty of identifying patients increases the closer to normality the line is drawn. Existing age-based norms for cognitive tests are almost certainly contaminated by cases of unsuspected incipient AD (Morris, McKeel, Storandt, Rubin, Price, Grant, Ball and Berg, 1991); depression, which is sometimes seen as a confound in the diagnosis of AD, may rather constitute one of its prodromal symptoms (Wetherell, Gatz, Johansson and Pedersen, 1999); and finally a reliable biological marker of AD in relatively healthy people has yet to be found. The difficulties of mapping out the grey zone between incipient AD and "normal aging" begin with the problematic nature of AD itself.

3.1 Histopathology

A diagnosis of Definite AD requires confirmation by biopsy or post-mortem autopsy

(Cummings and Khachaturian, 1996). Histopathologically, Alzheimer's disease (AD) is characterized by **neuronal loss**, intracellular **neurofibrillary tangles** (NFTs), and extracellular **senile plaques** (SPs) containing amyloid β protein (Aβ) (Goldman and Côté, 1991). Aβ is thought to disturb calcium homeostasis and cause oxidative damage by free radical formation, resulting in synaptic loss and neurotransmitter deficits (Selkoe, 1997).

Neuronal loss was once thought to be a feature of normal aging, but is now seen as pathological (Gómez-Isla, Price, McKeel, Morris, Growdon and Hyman, 1996). In mild AD, there is heavy loss of neurons in entorhinal cortex layer II, a change that does not occur in normal aging (Gómez-Isla et al, 1997; Gazzaley, Thakker, Hof and Morrison, 1997). The entorhinal cortex layer II sends excitatory input to hippocampus via the perforant pathway; compromise of this circuit is thought to contribute to the memory deficits of AD.

With increasing AD severity, **neurofibrillary tangles** (NFTs) progress from hippocampus and the rest of the limbic system to neocortex, a post-mortem histopathological finding which corresponds to ante-mortem clinically detectable dementia. It is hypothesized that only as NFTs spread from hippocampus to temporal neocortex do more serious memory deficits appear (Hodges and Patterson, 1995). However, NFTs, which are formed especially by the pyramidal cells of hippocampus

and neocortex, are a feature not only of AD but also of Down's syndrome, dementia pugilistica, and post-encephalitic Parkinsonism, as well as normal aging (Morris, Storandt, McKeel, Rubin, Price, Grant and Berg, 1996).

In contrast to NFTs, **Senile Plaques** (SPs) do not increase with normal aging, and do not correspond well to AD severity (Goldman et al, 1991). However, both types of SPs—*neuritic*, containing filamentous Aβ and twisted pieces of axons and dendrites; and *diffuse*, containing nonfilamentous Aβ and no neurites—are now believed to constitute an obligatory first step in the cell-damaging AD cascade (Selkoe, 1997). Senile plaque densities are greater in neocortex than in hippocampus of patients with threshold memory impairments, and it is hypothesized that neocortical SPs signal the histopathological presence of AD, while other events downstream from these result in clinically detectable dementia (Morris et al, 1996). The recent discovery of a genetic mutation leading to an early onset form of familial AD accompanied by lower than expected Aβ plasma levels—the 'Arctic' mutation, found in a family from northern Sweden—has led to the suggestion that protofibrils, an intermediary step in the formation of Aβ fibrils—which accumulate to form senile plaques—may be an important causative factor in the development of more common (sporadic) forms of the disease (Nilsberth, Westlind-Danielsson, Eckman, Condron, Axelman, Forsell, Stenh, Luthman, Teplow, Younkin, Näslund and Lannfelt, 2001). Nilsberth et al.

(2001) speculate that the abnormal overproduction of protofibrils, which not only form neurotoxic Aβ assemblies but are themselves neurotoxic, may be at the root of AD and other neurodegenerative disorders. Such a molecular mechanism would explain the lack of correspondence between senile plaque density and AD severity.

What this brief review highlights is that by the time it is diagnosable, AD has long been established histopathologically in the hippocampus (Morris, McKeel, Storandt, Rubin, Price, Grant, Ball and Berg, 1991), entorhinal cortex (Gómez-Isla et al, 1996) and basal forebrain (Geula 1998). Given the lag between the inception of pathological changes in the brain and clear behavioural deficits, and the difficulties this causes with medical treatment (Geda and Petersen, 2001), much research has therefore gone into attempting to delineate some rational category for aging-associated cognitive decline that would signal incipient AD.

3.2 Neuropsychology

3.2.1 Cognitive Markers of AD

Distinguishing the effects of early-stage dementia from those of depression or normal aging—which are only two of many possible confounds—is one of the thorniest problems of neuropsychology (Habib and Poncet, 1991). In the early stages of AD, disease expression is both heterogeneous and subtle, making genuine AD symptoms very difficult to pick out against an extremely "noisy" background (Rockwood and Morris, 1996). In later stage AD, the clinical picture becomes more homogeneous

(Graham et al, 1996) and in the end, terribly simple.

In early-stage AD, not only is the background of individual variability extremely noisy, but the identity of the cognitive marker to search for remains a matter of some controversy. One current in the literature concerns a putative more rapid rate of forgetting in AD patients than in normal elderly controls (Welsh, Butters, Mohs, Beekly, Edland, Fillenbaum and Heyman, 1994). This rate of forgetting is often reported as a savings score, i.e. the percentage of material initially recalled that is also recalled after a delay (Petersen, Smith, Kokmen, Ivnik and Tangalos, 1992; Welsh, Butters, Hughes, Mohs, and Heyman, 1991). In the typical Immediate Recall paradigm, subjects are asked to remember material (verbal or nonverbal) immediately after its presentation, while in the Delayed Recall paradigm they are asked to recall the same material after a delay, often filled with other material they must consciously attend to. In an information-processing model of memory (Estes, 1978), three main stages of memory processing have been inferred from subjects' responses to experimental situations, namely encoding, storage or consolidation, and retrieval (Kaszniak, 1986). Together, these three processing stages are hypothesized to be involved in transforming Short-term (or Primary) memory into Long-term (or Secondary) memory (Waugh and Norman, 1965; Kaszniak, 1986). Short-term memory is of limited capacity and constrained to what the subject is consciously attending to at the moment, while Long-term memory is of theoretically unlimited

capacity. The transformation of Short term into Long term memory is gradual; once material has been encoded, consolidation, and thus resistance to disruption, can continue for decades (Squire, Haist and Shimamamura, 1989; Haist, Gore and Mao, 2001). Nevertheless, the Delayed recall paradigm—and with it savings scores—are theoretically linked more to the processes of storage and consolidation or retrieval rather than to the process of encoding. Lower savings scores may suggest a deficit of consolidation or retrieval rather than encoding. The interpretation of savings scores is problematic however because they are correlated with and dependent on initial immediate recall scores (Cohen and Cohen, 1983). Because of this, Robinson-Whelen and Storandt (1992) used a multiple-regression strategy to partial out immediate recall scores from a story recall task and found that diagnostic status (AD vs. normal) then failed to account for a significant proportion of variance in Delayed Recall scores. Thus, to focus on Delayed Recall or savings score would, in this view (Robinson-Whelen and Storandt, 1992) miss the characteristic encoding (as opposed to consolidation or retrieval) deficit of incipient AD. However, Reed and coworkers (Reed, Paller and Mundas, 1998) equalized baseline performance in a visual learning task in AD and normal subjects by manipulating stimulus exposure times, and found that AD was associated with lower savings scores. Such evidence suggests that AD affects both aspects of memory, and that studying the performance at the level of multiple memory components is probably a wiser choice at this present state of the art

than searching for a weakness in a unique isolated component.

Another approach is to test and examine a large number of people, let several years pass, and retest and re-examine everybody to determine who has progressed to AD, deriving thereby a constellation of symptoms shared by the latent AD sufferers and not the normally aging at baseline. Tierney and collaborators adopted this method (Tierney, Szalai, Snow, Fisher, Nores, Nadon, Dunn and St. George-Hyslop, 1996). They tested 123 seniors twice over two years, and used the diagnosis at follow-up (AD or normal) to pick out from among all the baseline examinations, those tests which best differentiated future AD cases from future non-cases. At baseline the inclusion criteria included age above 70, an MMSE (Folstein et al, 1975) score greater than 23, and a 3-month history of memory problems which interfered with daily functioning (corresponding to Stage 2 or 3 of the Global Deterioration Scale [Reisberg, Ferris, DeLeon and Crook, 1982]). Volunteers underwent an extensive medical workup and those with dementia or any other known cause of memory impairment were excluded from the study (Tierney et al, 1996). Two years later, 29 subjects, or 24% of the baseline sample, met DMS-III-R criteria for Probable AD (American Psychiatric Association, 1987). Tierney and colleagues (1996) then conducted a linear regression analysis to determine which combination of baseline neuropsychological test scores predicted AD diagnosis at two-year follow up. Using age and education as covariates, and based on scores from just two neuropsychological tests, namely the delayed recall

portion of the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964; Lezak, 1982), and the Mental Control subtest from the Wechsler Memory Scale (Wechsler and Stone, 1973) the logistic regression equation elaborated by Tierney and coworkers (1996) provided the degree of risk a case had of progressing to dementia within 2 years. The equation calculated risk as a percentage of statistical certainty; this percentage was then dichotomized (those at 50% risk or above were classified as future cases), a classification procedure which had a sensitivity of 76%, a specificity of 94%, and an overall accuracy of 89%. These results have been considered promising, worthy of prospective replication (Reekum et al, 1999). However, there are three methodological problems with the approach of Tierney and colleagues (1996), which would weigh against the need for such a replication:

First, the RAVLT (Rey, 1964; Lezak, 1982) is not accepted well by test subjects, which could hinder the collection of relevant and accurate data on behavioral symptoms of AD. The test consists of a 15-word list to be read aloud to the subject fives times, with a recall trial after each exposure, followed by an exposure to a distractor list of 15 words before the final, delayed recall trial. More seniors refused to complete this test than any other in the Canadian Study of Health and Aging (CSHA); 19 of 21 psychometricians working on the CSHA reported that there were "often" or "almost always problems" with the administration of this test (Tuokko, Kristjansson and Miller, 1995). Seniors will often balk at performing laboratory tasks with low face

validity, which they consider trivial or ridiculous; Larrabee and Crook (1988) point out that face validity, and thus a greater degree of subject cooperation, is particularly important when what is being evaluated are behavioral symptoms, rather than theoretical constructs. Because of the possibly confounding factor of subject motivation (or lack thereof), the use of the RAVLT is therefore questionable in this population.

Second, Tierney et al's (1996) approach emphasizes delayed rather than immediate recall; and the interpretation of delayed recall scores—representing the efficiency of consolidation and retrieval processes—is, as we have seen above, confounded with the effect of initial encoding.

Third and most important, Tierney et al's (1996) procedure categorizes subjects as at risk based on an even or greater chance of developing diagnosable AD within 2 years. Given, however, that the preclinical phase of AD (during which symptoms can be detected but before a diagnosis can be made) can stretch to ten years or more (Linn, Wolf, Bachman, Knoefel, Cobb, Belanger, Kaplan, D'Agostino, 1995; Elias, Beiser, Wolf, Au, White and D'Agostino, 2000), a procedure that would allow detection of "at risk" cases earlier would be preferable.

The approach of Tierney and colleagues is thus beset with methodological difficulties that limit its usefulness. Nevertheless, their approach, based on screening, is common-

sensical and worth exploring further.

3.2.2 Screening Measures

The ultimate goal of screening is to provide disease-modifying or curative treatment to patients earlier rather than later in the course of a disease. In order to do this the disease must of course be detected in its earliest manifestations: screening is meant to detect disease before it is diagnosable *as* disease. In practice, patients who screen "positive" on a screening examination are then examined more extensively in order to definitely rule out disease—or confirm the screening result. In medicine the term "screening" is used in two ways: one can screen for risk factors, such as hyperlipidemia as a risk factor for the subsequent development of heart disease; and in this present usage, one can screen for the early stages of disease (Pablos-Mendez and Cimino, 1997). Although it may appear self-evident that screening for a terrible disease in order to detect it at its earliest possible stages would be more beneficial, or less harmful, than doing nothing, in the case of AD at least the balance of risks and benefits has been judged unfavorably by both Canadian (Patterson, Gauthier, Bergman, Cohen, Feightner, Feldman, Grek and Hogan, 2001) and U.S. (Petersen, Stevens, Ganguli, Tangalos, Cummings and DeKosky, 2001) academic consensus groups, who have both recommended against systematic screening of asymptomatic individuals. In order to put this negative recommendation into context, it is necessary to go into some detail about the difficulties and unresolved problems with screening

programs for AD. The following will serve to introduce the notion that no single, unifactorial approach to the problem of screening is possible, and that an ideal program would probably involve a multi-disciplinary search for combined biological and neuropsychological (Almkvist and Winblad, 1999) as well as psychosocial markers of early AD disease.

Beginning with the most general level of analysis, the first difficulty with screening programs is that they have at least two built-in biases which tend to produce inflated estimates of their benefits: *lead-time bias* and *length bias*. Let us suppose that on average AD eventuates in death 8 years after diagnosis. If, by means of screening for AD, the average time between diagnosis and death is extended to 20 years, screening would then appear to be beneficial, since in the first instance AD patients live 8, and in the second 20, years post-diagnosis. This is lead-time bias. The cause-specific mortality rate is not reduced, and patients are simply followed for a longer time before death without the disease course being affected in any way. Length bias also produces inflated estimates of the worth of screening programs, because by their nature they tend to capture more mild cases and miss rapidly deteriorating cases (i.e., people who are too sick to be screened). Length bias is particularly germane to the field of AD research, since those uncertain, borderline cases which screening programs are meant to detect are just the ones which are likely to live long anyway, or respond to non-specific treatment, while cases of rapidly progressing AD are more likely to be missed

simply because they degenerate so fast there are fewer cases are available to pass through screening programs. In fact drug treatment of AD may itself be subject to a type of length bias, if all it accomplishes is extend the period of decline without reducing the personal and societal burden caused by the disease (Dresser, 2000).

In the clinic, there are four main approaches to screening (Patterson and Gass, 2001): (1) asking the patient whether memory complaints are present; (2) querying an informant; (3) verifying the ability to perform instrumental activities of daily living (IADL; e.g. using the telephone, handling finances); and (4) performing a mental status exam. The first of these approaches is the subject of the present thesis. The second has the weakness that seniors living in isolation, without close and lasting relationships, could not be screened. The third approach is problematic with regards to detection of MCI as preclinical AD, because a patient with mild cognitive impairment and difficulties performing IADL along would be more likely to be diagnosed as already having AD rather than MCI. The fourth method of screening concerns a mental status exam, using a testing instrument and procedure simple enough to be integrated into primary care medical practice.

Often the person's score on a screening instrument will be reduced to a binary datum, representing the person's performance as either above or below a given threshold. Thus, the first problem consists in selecting the measure on which the screening

threshold is based, and the second in selecting the threshold. Such a threshold or cut-off value cannot, of course, encompass all the wealth of clinical material that an experienced clinician may perceive in assigning a diagnosis (O'Connor, Pollitt, Hyde, Miller and Fellowes, 1991). (This is an important point I will return to in my discussion.) Nevertheless, taking into account the seriousness of the disease in question along with the time, effort and cost involved in ruling out potentially rectifiable causes of dementia, it is clear that there is a great need for tests which detect the early stages of AD with "good enough" sensitivity and specificity (Patterson, 1994).

These latter two factors are, along with ease of use, what define a good test. The sensitivity of a test is its ability to correctly classify a case (i.e., an instance of the disease in question) against a "gold standard", in most cases a complete clinical examination followed by a diagnosis; whereas the specificity of a test refers to its ability to correctly classify a non-case (the absence of the disease) against the same gold standard. In theory a test with perfect sensitivity (missing no cases of disease) and specificity (misidentifying no non-cases as cases) is possible, but in reality due to the variability in disease expression, sensitivity and specificity can only be increased at each other's expense. For a clinician, however, it is more useful to know the positive and negative predictive value (PPV and NPV) of a test than its sensitivity and specificity. The PPV give you the odds that a person actually does, and the NPV that a

person actually does not, have the disease in question, given a positive or negative test result respectively. These odds depend on the prevalence of the disease in the particular population to which that person belongs. Calculating and utilizing the PPV and NPV are only possible in situations where such information is available or can be estimated with a reasonable degree of accuracy.

In the primary care setting, where the AD prevalence is low, increasing sensitivity at the expense of specificity leads to a decrease in the already low positive predictive value, with no compensatory increase in negative predictive value, since this latter is already almost at ceiling levels. In a high prevalence, Memory Clinic population on the other hand, the same trade-off leads to a steeper drop in positive predictive value, which is, one can argue, more than compensated for by the corresponding steeper *rise* in negative predictive value. The argument here is qualitative, not quantitative. What is the goal of the procedure? Arguably, in a primary care setting, in a family practice physician's office, for example, where the prevalence of AD is expected to be relatively low, it is better to have a test—or a test threshold—which maximizes specificity at the cost of sensitivity, because to do otherwise—to choose a more sensitive test—would not buy you much more negative predictive value compared to what your test would lose in positive predictive value. In a secondary care setting, however, such as in a memory clinic or geriatric hospital, where AD prevalence is expected to be relatively high, one could argue that it is better, within limits, to give

more emphasis to sensitivity than to specificity. This has the effect of slightly increasing the NPV, and of decreasing the PPV to a greater degree than in the low-prevalence population. Thus, such a trade-off "costs" relatively more in positive predictive value and consequently such a sensitivity/specificity ratio will produce more false positives cases. But that perhaps is no drawback. Since the raison d'être of the memory clinic is to make sure (more sure than in the initial evaluation which produced the referral to this secondary- or tertiary-care setting) that no patients with incipient AD are going without treatment, care and follow-up, it is better to mistakenly evaluate in-depth more actually healthy people than to turn away more actually diseased but undiagnosed (or unscreened) people who will then have no other recourse for specialist treatment. Thus the choice of a screening procedure has an ethical aspect. And like all ethical questions this one has an eminently practical application, which returns us to our earlier specific challenge: which test(s) to use, using what cut-points?

As mentioned above, of the many available cognitive assessment tools, undoubtedly the one most frequently used in general, non-specialized settings is the **MMSE** (Folstein et al, 1975). The MMSE is quick and easy to administer, for which reason it has gained wide acceptance. It features six categories of items: orientation to time and place (which are scored on 10 points); registration of three words (3 points), along with their recall (3 points); attention and calculation (5 points); language (8 points); and visual construction (for a total of 30 points). Over the years and across the many

centres utilizing the MMSE, minor variations have developed in its administration, thereby undermining one of the test's major advantages, its cross-disciplinary intelligibility. These variations have been addressed by the development of a standardized form of the MMSE (Molloy, Alemayehu and Roberts, 1991).

Significant cognitive impairment is most often defined as a score of 23 or less out of 30 (Tombaugh and McIntyre, 1992; Anthony, LeResche, Niaz, von Korff & Folstein, 1982). In their comprehensive review of the MMSE, Tombaugh and McIntyre (1992) listed all the studies then available which reported on its criterion validity; they selected only those studies which used 23/24 cut-point to differentiate AD and other serious cognitive impairments from cognitive normality. Of the 12 studies which listed the MMSE values for both the cognitively impaired and normal groups, the mean MMSE score of those deemed normal was 27.2 (SD = 2.5), with a range of 20 to 29. There is thus a wide range of cognitive normality as indexed by the MMSE across different centres. The cut-off of > 23 for "normality" is a matter of rough consensus from years of use in clinical settings, where the prevalence and severity of AD cases is higher than that of the general population. Moreover, scores in the range of 24-27 in persons above 80 years of age, have been associated with increased mortality rates (Gussekloo, Westendorp, Remarque, Lagaay, Heeren, and Knook, 1997) and cognitive deterioration over time (Izaks, Gussekloo, Dermout, Heeren, and Ligthart, 1995) compared to those with scores in the range of 28-30. Persons with near perfect scores

(26-29) have been found to have neuropathologically confirmed AD within 5 years or so (Morris, Storandt, Miller, McKeel, Price, Rubin and Berg, 2001). The MMSE is thus considered to be just as insensitive to the manifestations of mild cognitive impairment (Teng, Chui and Gong, 1989; Ihl, Frolich, Dierks, Martin and Maurer, 1993) as it is to the severe cognitive impairments of late stage AD (Auer, Sclan, Yaffee and Reisberg, 1994).

The Alzheimer's Disease Assessment Scale, cognitive subscale (**ADAS-cog**) (Rosen, Mohs and Davis, 1984) has become the gold standard in dementia drug research. It is an 11-item test, which measures memory and orientation, attention, language and praxis. Of the commonly used dementia psychometric tests, the ADAS-cog has the best discriminative validity across the whole course of the disease (Ihl et al, 1993); standardization of administration procedures has further increased its sensitivity (Standish, Molloy, Bédard, Layne, Murray and Strand, 1996). However, the ADAS-cog was not designed to aid diagnosis but rather to track progression, and is relatively insensitive to mild AD (Wolfson et al, 1997). We decided to include the ADAS-cog, in Standish et al's (1996) standardized version, because it is the only cognitive instrument extant which was designed specifically for AD, and because of preliminary data that it is more sensitive than two other commonly used instruments in its correlation to the volume of brain areas associated with memory.

The Modified Mini Mental State (**3MS**) has an expanded scoring range of 0-100 rather than 0-30; this and other changes increase its discriminating power compared to the MMSE (Teng and Chui, 1987; McDowell, Kristjansson, Hill and Hébert, 1997). Using data from the Canadian Study of Health and Aging (Canadian Study of Health and Aging Workgroup, 1994), Kristjansson and colleagues (1996) calculated that for the 3MS a cut-point of 80 was optimal for differentiating Cognitive Impairment, No Dementia (CIND) from dementia. The Canadian Study of Health and Aging (CSHA), however, was a probability sample study of the prevalence of AD in the Canadian population, and was not designed to detect mild cognitive impairment. Accordingly, this cut-point is not a recommendation for future clinical practice, but only an example of the process of setting thresholds on clinical parameters between various treatment options (Kristjansson et al, 1996). The simplest example of this process is deciding on the relative importance of false positives and false negatives in choosing a cut-point that defines a positive test, i.e. indicating the presence of a disease or the need to go ahead for further in-depth testing, as the case may be. The MMSE has often been used for this latter purpose. This task becomes more complex when the clinician is confronted with more complex diagnostic possibilities (Feinstein, 1990). Kristjansson and coworkers (1996) took up the challenge of distinguishing, not simply between AD and cognitive normality—which is no really easy task—but between these first two categories and a third, intermediary category, potentially equivalent to the prodromal

stages of AD but also possibly representing suboptimal cognitive performance due to any number of non-AD causes. For this category Kristjansson et al (1996) used the term Cognitively Impaired, Not Demented (CIND). The usefulness and limitations of such a term become clear in a model of incipient-AD detection.

Consider a patient at a walk-in clinic who presents with subjective memory complaints; the person is tested with a standardized screening battery, and classified into one of three categories defined by some given combination of cut-off scores. Someone whose scores suggests cognitive normality would (for example) be reassured and followed-up in six months; someone whose scores (in the context of a whole evaluation) suggest AD, would be referred to a specialist; and someone whose scores fall somewhere between these two would be also referred to a specialist for an in-depth evaluation, or offered participation in an experimental prevention trial. The choice of between various test cut-off points—or tests of varying sensitivity and specificity—must be informed both by the risk of missing an early symptom of a serious disease (e.g. of AD or depression) *and* by the risk of wasting limited resources through unnecessary specialist referrals.

For example, one test which has recently been proposed as suitable for determining the need for more detailed assessment of cognitive and functional status in seniors, is the clock drawing test (Esteban-Santillan, Praditsuwan, Ueda and Geldmacher, 1998).

This test consists in having the subject draw, on a blank sheet of paper, a clock indicating "10 past 11." Using simple scoring criteria which take into account only the placement of the clock hands, Esteban and collaborators found that, in a group of 41 patients with mild probable AD and 39 age- and sex-matched normal controls, a score of less than 3 out of a possible 4 on the clock drawing test was associated with a PV+ of 100% and a PV- of 51%. The specificity was 100% but the sensitivity was only 12%. Translated into our fictitious primary-care, low-prevalence population, this works out to a PV+ of also 100% (since the test's specificity was 100%, the PV+ will of course be 100% in all populations regardless of disease prevalence) and a PV- of 98%. Thus, using such a test and cut-point with one thousand consecutive patients would produce exactly zero unnecessary referrals. These numbers may seem impressive, but one must also consider that because of the test's low sensitivity, in the long run 88% of true positive cases would be misdiagnosed as normal. The optimal weighing of test sensitivity and specificity, or of positive and negative predictive values, is not at issue here so much as the difficulty of ever settling the question of screening for AD once and for all with any one test.

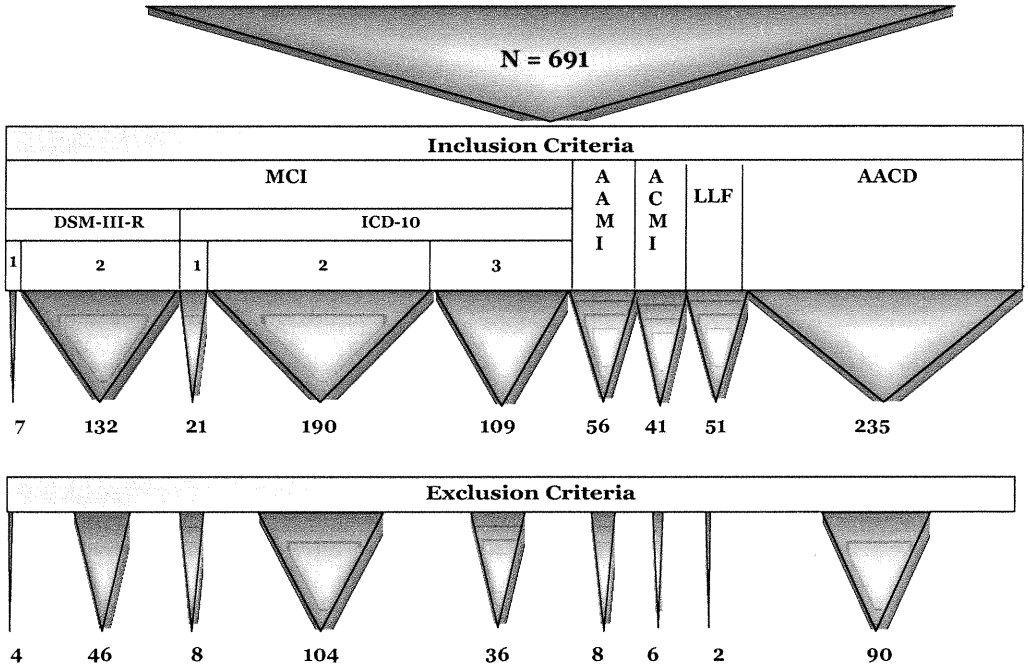
The weighing of the relative importance of these risks is thus, to reiterate, an ethical and practical matter. Kristjansson et al's (1996) approach holds promise, because it makes possible the relative weighing of those risks on empirical grounds. Their proposed cut-point of 80 on the 3MS for identifying Cognitively Impaired, Not

Demented (CIND) cases, for example, gives twice as much weight to sensitivity as to specificity. That is, such a threshold is grounded in a qualitative judgment that an error of commission (i.e., mistakenly classifying a true Normal case as CIND) is only half as objectionable as an error of omission (i.e., mistakenly classifying a true CIND case as Normal). Such a relative weighing of sensitivity and specificity would be appropriate, for instance, in a memory clinic (as in the example above) or, arguably, at the beginning of a prevention study, where a random selection of those classified as CIND would receive a putative AD-preventive agent. Before such a trial could be designed, however, a better definition of an at-risk population would have to be formulated, since the CIND category has empirical and theoretical weaknesses which curtail its usefulness in such an endeavor. The sampling methodology used in the CSHA limited the number of participants with milder rather than more serious cognitive impairments, and consequently the Cognitively Impaired, Not Demented (CIND) category is the least stable in terms of internal coherence and external validity (Graham et al, 1997). The main drawback of the CIND category, however, is that it is not a nosological entity, but was simply meant to be a catch-all category for all those cases of subnormal cognitive capacity in the CSHA sample that were not attributable to an underlying AD pathology. The CSHA needed such a category because of the many uncertainties concerning the validity of existing age-related cognitive impairment categories.

3.3 Diagnostic Uncertainties

The Canadian Study of Health and Aging revealed that fully two thirds of seniors with clinically demonstrable cognitive impairment did not have dementia, and were therefore categorized as **Cognitively Impaired, Not Demented (CIND)** (Ebly et al, 1995). CIND affects 16.8% of the population, more than all the dementias combined (8%) (Graham et al, 1997). Graham and her coworkers (1997) reported that CIND was associated with an increased rate of functional incapacity and need for institutional care compared to cognitively intact seniors; this is important from a public health perspective, since institutional care represents the heaviest demands on health spending. However, the association of CIND with increased rates of functional impairment is mostly due not to aging-associated cognitive impairments, but rather to chronic conditions such as aphasia, chronic schizophrenia or longstanding mental retardation. In fact, fully two thirds of CIND cases can be classified by etiology, e.g. drug or alcohol addiction, psychiatric diseases, multiple sclerosis, brain tumors, and other known causes of subnormal cognitive performances, as well as by confounding factors such as cultural differences (Ebly et al, 1995). There is thus a need for a category for those whose cognitive impairments are of unknown etiology, while not meeting the criteria for an AD diagnosis.

Figure 1
Mild Cognitive Impairment Categories in the Canadian Study of Health and Aging



Nine definitions of aging-associated cognitive deficits, in five distinct categories, were examined in the course of the Canadian Study of Health and Aging (CSHA). The different cognitive deficit categories examined by the CSHA have strict and mostly mutually exclusive inclusion and exclusion criteria.. The large triangle at the top of **Figure 1** represents the 691 clinically confirmed CIND cases from the CSHA. The base and area of each triangle in Figure 1 are proportional to the number of cases each MCI category "captures" of that N = 691 CIND sample. The bar underneath the main triangle, and the other bar lower down, represent the inclusion and exclusion criteria respectively for the five categories of aging-associated cognitive deficits examined by the CSHA: Mild Cognitive Impairment, which has two types under the DSM-III-R

(American Psychiatric Association, 1987) and three in the International Classification of Diseases nomenclature of the World Health Organization (WHO, 1993) Age-Associated Memory Impairment (Crook, Bartus, Ferris, Whitehouse, Cohen and Gershon, 1986); Blackford and Larue, 1989) Age-Consistent Memory Impairment (Blackford and Larue, 1989); Late-Life Forgetfulness (Blackford and Larue, 1989); Aging-Associated Cognitive Decline (Levy, 1994). As can be seen from the sizes of the triangle in the bottom row, few CIND cases could be classified into either of the five categories and nine types of MCI. The 3 types of MCI derived from World Health Organization criteria (World Health Organization, 1993) taken together captured the greatest proportion of clinically confirmed CIND cases (21%). A tenth, additional type of MCI, not shown on the graph, consisting of proposed modifications of Age Associated Memory Impairment criteria (Blackford and Larue, 1989), captured the smallest proportion of CIND cases, i.e. none. CIND is not a "pure" category, so it is not surprising that so few CIND cases would pass through the two layers of inclusion and exclusion criteria of tentatively "pure" MCI categories.

At first, the CSHA reserved the category of Age Associated Memory Impairment (AAMI) (Crook et al, 1986) to describe those with aging-associated mild cognitive impairment (MCI) of unknown cause; but it was found that the original AAMI criteria were so inclusive that nearly one third of the Canadian population over 65 could be so categorized (Graham et al, 1997). Indeed, AAMI is a controversial category, and the

literature is rife with contradictory findings. There are reports that AAMI, or memory loss of unknown etiology, is progressive (Tierney et al, 1996; Bowen, Teri, Kukull, McCormick, McCurry and Larson, 1997) or non-progressive (Hänninen, Hallikainen, Koivisto, Helkala, Reinikainen, Soininen, Mykkänen, Laakso, Pyörälä and Riekkinen, 1995; Snowdown and Lane, 1994; Youngjohn and Crook, 1993); that it is associated with hippocampal atrophy (Parnetti, Lowenthal, Presciutti, Pelliccioli, Palumbo, Gobbi, Chiarini, Palumbo, Tarducci and Senin, 1996) or hippocampal preservation (but with loss of asymmetry) (Laakso 1996); and that AAMI rates increase (Larrabee and Crook, 1994) or decrease (Koivisto et al, 1995) with age. Notwithstanding these inconsistencies, some intriguing findings have been reported on the AAMI category. Hänninen and colleagues studied the frontal function capacities of AAMI subjects, defined as persons with subjective memory complaints as indexed by a score of ≥ 25 on the Memory Assessment Clinics Questionnaire (MAC-Q; Crook, Feher, and Larrabee, 1992), objective impairment on the Paired Associate Learning subtest from the Wechsler Memory Scale (Wechsler 1945) and normal cognitive capacity as indexed by a score of ≥ 24 on the MMSE. AAMI subjects showed impaired performance on three out of four tests assessing frontal lobe function: the Wisconsin Card Sorting Test, the Stroop test, and the Trail Making test (Hänninen, Hallikainen, Koivisto, Partanen, Laakso, Reinikainen and Soininen, 1997). This is interesting in the light of other reports that normal aging is associated with slight but significant

decrements in frontal lobe function, similar to that produced by dorsolateral prefrontal cortical lesions (Levine, Stuss and Milberg, 1997). However, the AAMI criterion of performance 1 standard deviation below young adult means on standard tests of cognitive functioning has been considered over-inclusive (Levy, 1994; O'Brian and Levy, 1992) because it makes AAMI roughly equivalent to normal aging.

The criteria for Aging-Associated Cognitive Decline (AACD) (Levy, 1994) were developed partly to counter this criticism; here the reference is to age-appropriate norms, where the cognitive performance of a senior aged 63, for example, is compared to previously gathered data on seniors aged 60 to 65 and who are considered cognitively normal. However, as mentioned above, the validity of age-appropriate norms for cognitive performance depends on there being no unsuspected incipient AD cases within the normative sample, which is uncertain and even unlikely (Morris et al, 1996).

That Mild Cognitive Impairment is a controversial category with shifting boundaries might be expected, given that the conditions it describes are subtle and difficult to distinguish from normal aging (Jolles, Verhey, Riedel and Houx, 1995). These controversies must be addressed in any research project on cognitive aging, because there are important consequences to selecting one or the other MCI definition. Prevalence rates for the various categories of aging-associated Mild Cognitive

Impairment examined by the CSHA varied by factor of 50 according to the diagnostic criteria used (Ebly et al, 1995). The various criteria were also mostly mutually exclusive, with the majority of CIND cases fitting into only one of the 5 MCI categories; 19% of CIND cases fit into two categories; 4% into three, and 0.8% into four (Ebly et al, 1995).

That the MCI category were mutually exclusive arises from imposition of stringent operational definition of research groups. This has caused numerous difficulties with research in this domain. Tristmans and coworkers (Tristmans, Clincke and Peelsmans, 1990) reported that so few patients met all inclusion and exclusion criteria for Age-Associated Memory Impairment (AAMI) that there was doubt as to whether AAMI was a real disease entity. Similarly, in a report on the characterization of recruitment of mild AD patients, Berg et al (Berg, Hughes, Coben, Danziger, Martin and Knesevich, 1982) reported that stringent operational defintions had made their study unusually difficult to realize. A recent definition of MCI (Petersen, Smith, Waring, Ivnik, Tangalos and Kokmen, 1999) has already been applied by another group, who found that it was not only unfeasibly restrictive, but it had poor predictive validity for the development of AD (Ritchie et al, 2001). If restrictive inclusion and exclusion recruitment criteria make research so difficult, why are they used? The most obvious reason is that only by using operationalized definitions can researchers meaningfully communicate their findings. Another reason strict sample selection criteria are used

relates to the attempt to control for confounding factors. This control, applied in the form of requiring that research samples be homogeneous, is considered necessary for two reasons. One of these has to do with the concept of external validity, that is the grounds for generalizing research results to a wider population than the actual sample studied. The other has to do specifically with increasing statistical power: it is sometimes thought that homogeneous groups increase the power of statistical tests to reveal a true intergroup difference. (Yastrubetskaya, Chiu and O'Connell, 1997). Both of these reasons would appear to be based on erroneous reasoning.

According to Oakes (1972), any sample is bound to be "atypical" in some way compared to the population it was drawn from; but this does not render the results of a study of such a sample invalid. Rather, says Oakes, the external validity of the findings is a function of the population to which they are generalized. In his argument Oakes stated that the impossibility of selecting a sample that was wholly representative in every respect to the population it was drawn from is not a problem that would spoil the interpretability of results. Thus, following the suggestions of Oakes, rather than refuse many potential volunteers, which would increase the time and effort needed to reach an adequate sample size and perhaps even make the research project unfeasible (as in Yastrubetskaya et al's [1997] study), one could accept a greater degree of heterogeneity of entry characteristics, with due caution in the subsequent generalization of results towards other populations.

The final reason that a high degree of sample homogeneity is considered necessary has to do with a misapplication of the statistical concept of power. According to Kraemer (1981), the concept of homogeneity of research samples is confused by the lack of clarity as to *which* homogeneity is meant. The implicit, and mistaken identification that researchers often make, according to Kraemer (1981), is between homogeneity of entry characteristics and homogeneity of treatment response. Persons suffering from both depression and anxiety, say, could react differently to an experimental anti-depressant treatment than persons who are only depressed. Note however that comorbid anxiety in such a study would not be necessarily associated with *less*, but simply with a possibly, and unpredictably, *different* response. Anxious depression could well be just the type of depression that would have best responded to the investigational product (Kraemer 1981). There is no necessary relation between heterogeneity of entry criteria and heterogeneity of treatment response (or intergroup difference). Nor does homogeneity of entry characteristics ensure homogeneity of treatment response; instead, what it does ensure is something that falls into the category of unforeseen consequences. That is, the process of selecting subjects based on stringent inclusion and exclusion criteria introduces a bias such that the resulting "pure" sample may end up bearing little relation to the population it was drawn from and meant to represent. This eventuality was demonstrated in the field of MCI research by Ebly and her colleagues (1995) and in the field of AD research by

Schneider and his colleagues (Schneider et al, 1997). Thus, not only do srict inclusion and exclusion criteria make recruitment of adequate sample sizes in cognitive aging research difficult or impossible (Tristmans et al, 1990; Yastrubetskaya et al, 1997; Berg et al, 1982; Kraemer 1981) but it can also create such counter-productive operational definitions of MCI as one that is both unstable and unpredictive of AD (Ritchie et al, 2001).

Faced with this issue of problematic heterogeneity of presenting symptomatology in the cognitively impaired, The Canadian Study of Health and Aging (CSHA) invented an atheoretical category to classify—or rather, to preserve from premature classification—those subjects whose cognitive deficits could not be associated with any known disease or confounding factor (e.g. cultural difference): *Circumscribed Memory Impairment* (CMI), which affects an estimated 5.3% of Canadian seniors (Graham et al, 1997). However, CMI is an exploratory concept and not a diagnostic category. It designates aging-associated cognitive decline, without any implications as to etiology, nosology or treatment. CMI is a parsimonious and neutral label, intended to preserve doubtful cases for eventual differential diagnosis. But since CMI is a concept rather than a diagnosis, the problem remains of a suitable operational definition of a target population for research purposes.

The procedure used by Tierney et al (1996) to classify subjects as likely or not to

progress to AD within 2 years is not useful if the goal is to detect incipient AD at far earlier stages in order to improve the efficacy of preventive or disease-modifying treatments. There has thus been much research in neuropsychology on identifying the type of mild deficits that would be pathognomic of incipient AD; and most of that work has been on declarative memory.

3.4 Declarative Memory

Declarative or explicit memory is one of the two main types of long or medium term memory. The other type, variously called non-declarative or implicit memory, refers to memories which can only be accessed through performance, or which can affect behaviour without the subject's awareness, as in priming and conditioned responses (Graf and Schacter, 1985). Declarative (or explicit) memory, on the other hand, *can* be consciously recollected and declared to others (Graf and Schacter, 1985). Declarative memory includes semantic and episodic memory. Semantic memory is culturally shared, and consists of temporally non-specific facts and general concepts (e.g. "lions are more ferocious than antelopes") as well as words and their meanings. Episodic memory refers to verbal and non-verbal events that happened at a specific time and place in the person's experience. The hippocampus proper is considered to be the most important brain structure for declarative memory in humans (Squire 1992).

3.5 Hippocampal Function

A bilateral structure extending along the lateral ventricle, and named after its resemblance to the sea-horse, the hippocampus forms part of the limbic system which was itself named (from the Latin *limbus*, border) by the French neurologist Paul Broca, who first characterized the phylogenetically primitive gyri that border round the brain stem (Broca 1878). In an information processing model (Velmans 1991), the hippocampus enables the formation of adaptive associations between stimuli and the contexts in which they occur (Gray 1995). In a neurobiological "scaffolding – storage" model, the hippocampus can be considered as the site of the initial erection of the scaffolding of memory: after an initial period spanning minutes to hours, the scaffolding work continues in the entorhinal cortex, where the memory trace is gradually build up within extensive entorhinal – neocortical connections; eventually, long term storage or consolidation takes place within neocortical connections, without hippocampal or entorhinal involvement (Haist et al, 2001). The patterns of neural circuitry would seem to provide support for such a model: virtually all information enters and leaves the hippocampus through the entorhinal cortex, which is itself richly interconnected with the perirhinal and posterior parahippocampal cortices; these, in turn, receive input from multi-modal association areas in the posterior parietal, temporal, and prefrontal cortices (Geula 1998). The first brain structure to sustain detectable damage in AD pathology is the entorhinal cortex, the major input area to

the hippocampus (Gómez-Isla et al, 1996); next affected is the CA1 sector of the hippocampus, one of its major output areas (Braak and Braak, 1991). The resulting isolation of the hippocampus is thought to constitute one of the two main causes of the memory impairments of AD (Geula 1998). The other cause stems from disturbance of a variety of subcortical inputs, notably from cholinergic basal forebrain neurons (Goldman and Coté, 1991). The projections from these neurons spread diffusely to almost the entire cortex, and are thought to subserve learning and memory via effects on the organism's capacity to attend to relevant stimuli (Voytko, Olton, Richardson, Gorman, Tobin & Price, 1994). Thus, given the complexities of neural circuitry and its complex and selective vulnerability to AD pathology, the disease can affect consciousness in a variety of ways.

The most studied memory component in AD has certainly been long-term (or secondary) memory, but other cognitive functions are also affected. For example, immediate (or primary) memory impairments, elicited by digit span tasks, and visuospatial deficits, detected by a Block Design task, were found in the AD samples from the Canadian Study of Health and Aging (Steenhuis and Østbye, 1995). Digit span, however, differed by less than half a point between normal and AD groups, casting some doubt on the usefulness of immediate memory measures. As well, using the different visuospatial abilities of the normal and impaired groups (Steenhuis and Østbye, 1995) to define an early-stage AD group would require reference to age- and

education-matched norms; and such norms are riddled with conceptual and empirical difficulties. To begin with, demographic test score adjustments treat age and education solely as causes of psychometric bias rather than as possible risk factors for AD (Tombaugh and McIntyre, 1992). Next, whether adjusted scores are preferable to unadjusted scores is an empirical question, not always confirmed in reality (Kraemer, Moritz and Yesavage, 1998). Last and most important, the norms to which a person's performance is compared might have been contaminated by the presence of unsuspected mild AD cases in the normative samples (Morris et al, 1996). Gómez-Isla et al (1996) found that if at any point before death there had been a diagnosis of even Very Mild Alzheimer's Disease (Berg 1988), a category corresponding to the earliest clinically detectable cognitive decline, a postmortem examination invariably revealed a profound loss of entorhinal cortex (EC) neurons. Normal controls carefully screened to exclude Very Mild AD (Berg 1988) showed some neurofibrillary tangles and some senile plaques postmortem, but never such a loss of EC neurons. Such findings are forcing a re-evaluation of the parameters of normal cognitive aging, and by the same token render suspect existing normative data sets (Morris et al, 1996; Gómez-Isla et al, 1996).

Congruent with the role of the hippocampus in memory processes and its selective vulnerability to the disease process, some researchers have reported that AD patients show particular weakness on tests of delayed episodic recall (Kopelman 1985), an

ability which, however, also declines with age. It has been suggested that the age effect on delayed recall scores can be eliminated by using the savings score instead, that is the percentage of immediate-recall material that is retained over a delay interval (Welsh et al, 1994) But as discussed above, savings scores have the fundamental problem of being correlated with and dependent on initial scores (Robinson-Whelen and Storandt, 1992).

That AD patients do not forget more, but rather encode less well, was also reported by Wilson et al (1983), using the following paradigm: in a word recognition task, healthy adults have a higher hit rate for rare words than for common words (so long as the rare words are distributed randomly among more common words), presumably because greater attentional resources were allotted to their encoding; AD patients fail to show this rare word advantage. More tellingly, when the intervals between the reading and recognition trials were adjusted so that total hit rates were equal for both AD patients and normal controls, these later still showed a rare word advantage while the AD patients did not (Wilson et al, 1983). Here, by carefully controlling for possibly confounding differences on one level of information processing, Wilson et al (1983) were able to isolate and identify a real difference on another level. Using a similar approach, Belleville et al (1998) were able to differentiate the effects of aging on working memory from its effects on memory storage. Previous investigators had evaluated processing and storage capacity separately and found that seniors were

impaired on both. Belleville and her coworkers (1998) equalized the storage demands by adjusting the list length in a word-list alphabetical recall task to individual capacities. Thus, by controlling for slight differences in storage capacity, Belleville et al (1998) demonstrated that normal aging is not associated with a manipulation deficit, but only with a slight storage deficit compared to younger adults. Such a finding holds promise for AD research, since manipulation of information, considered as part of working memory (Baddely 1986), is thought to be impaired very early in the course of AD (Kertesz and Mohs, 1996).

Although the cognitive deficits of AD are certainly the most salient ones both in the popular imagination and in the middle stages of the disease, it is affective symptoms which are believed to be the first sign that something is wrong in early-stage AD (Gauthier, Thaland Rossor, 1996; Royall 1997; Jost and Grossberg, 1996). In practice, such signs are often ignored or ascribed to the normal effects of aging (Royall 1997). But hippocampal damage early in the course of AD has been linked to affective disturbances such as the "predementia syndrome" reported in some patients, characterized by a reduced ability to deal with complex social situations (Jolles et al, 1995). This results in depressive symptoms, passivity, withdrawal, and increased dependence on, and preference for, a stable environment (Persson , Berg and Nilsson, 1991). Interestingly, a similar symptom constellation (although not directly related to AD) is also found in monkeys, where it is associated with elevated levels of

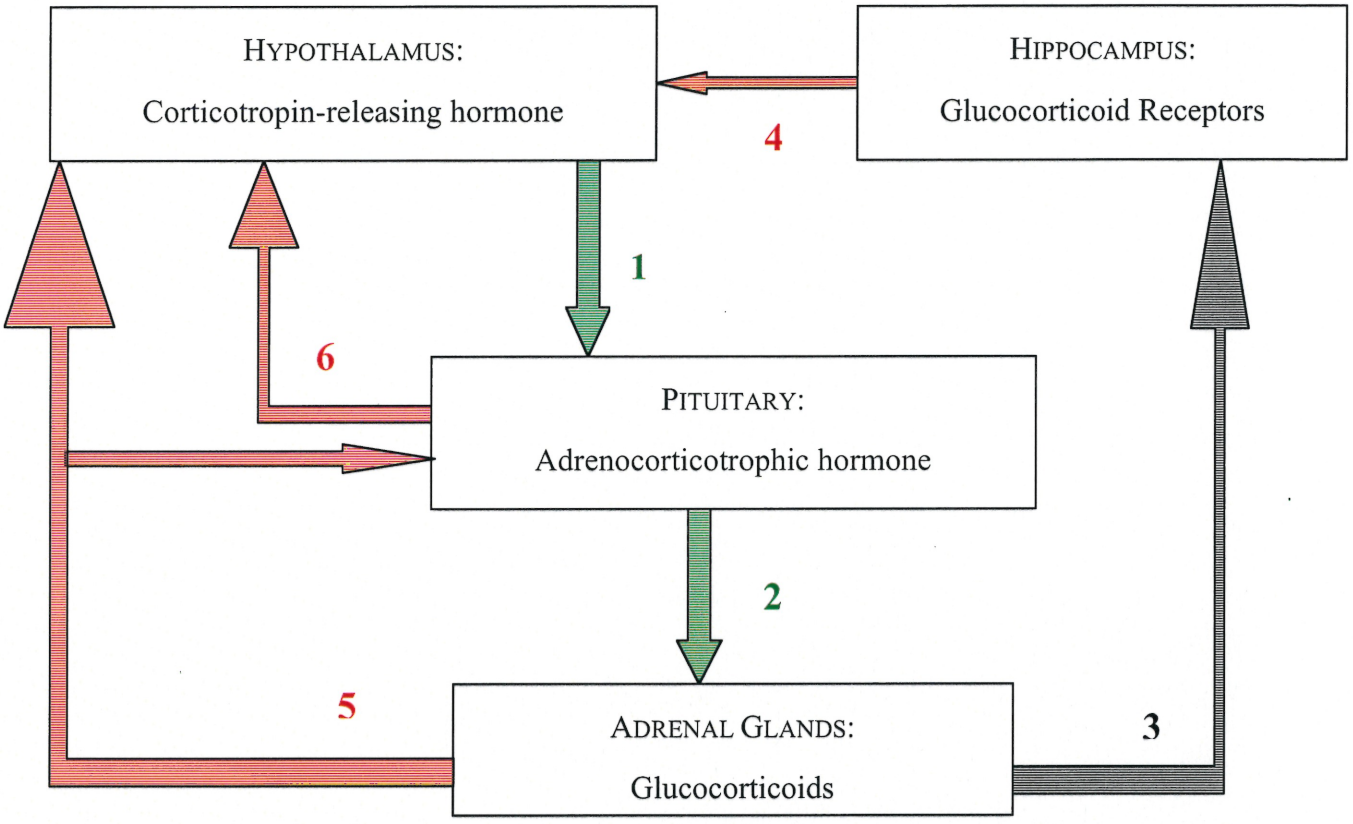
glucocorticoids, the stress hormones (Gust, Gordon, Hambright and Wilson, 1993; Veenema, Spruijt, Gispen and Van Hoof, 1997).

3.6 Neuroendocrinology

3.6.1 Allostasis

Glucocorticoids, so named because they promote energy mobilization via carbohydrates, are catabolic hormones that are essential for adaptation to stress but which in chronic excess become neurotoxic, with particular affinity to the hippocampus and deleterious effects on learning and memory in humans (Lupien et al, 1994, 1998). In response to stress, glucocorticoids (primarily cortisol) are secreted by the adrenal glands under control of adrenocorticotropic hormone (ACTH) from the pituitary, which is itself controlled by corticotrophin-releasing hormone (CRH) from the hypothalamus (McEwen 1987). In **Figure 2** green arrows represent agonist, and red arrows antagonist effects on subsequent hormone or cortisol release. The control of glucocorticoid hormones—in humans, primarily cortisol—is accomplished by the hypothalamic-pituitary-adrenal (HPA) axis. A given stimulus increases the liberation of CRH from the hypothalamus (1), which leads to increased release of ACTH from the pituitary (2), which in turn causes the adrenal glands to release glucocorticoids (3 and 5).

Figure 2



Stress may be defined as any change in the internal or external environment that demands, or seems to demand, that an organism change in order to adapt. The essence of the physiologic response to stress is the ability to achieve stability through change, a concept termed allostasis (Sterling and Eyer, 1988; McEwen, 1998). The related concept of *allostatic load* refers to the potentially deleterious effects on the organism of a less than perfectly self-regulating HPA axis (Sterling and Eyer, 1988). Sooner or later the stress response must cease, and when the system works this can be achieved through three modes: Increased circulating cortisol (3) is sensed by glucocorticoid

receptors in the hippocampus which then send a signal to the hypothalamus (4) that inhibits further CRH release. Cortisol regulation is also effected via the so-called long-feedback loop (5), whereby circulating cortisol inhibits ACTH release from the pituitary and CRH release from the hypothalamus. Finally, the pituitary can also directly inhibit the hypothalamus from releasing further CRH via the short feedback loop (6) (McEwen 1987; Kupferman 1991; Lupien and Forget, 1995). This whole system is under exquisite balance, and small defects in its operation, though initially negligible, can in the long run contribute to pathophysiological processes.

The main psychological factors defining stress are novelty, unpredictability, lack of control over the situation, lack of outlets for frustration, and a perception that things are getting worse (Mason, 1968; Sapolsky 1992) Confronted with these challenges to its equilibrium the organism uses its metabolic system, the HPA axis, and the cardiovascular, metabolic and immune systems in order to protect itself. But as stated above, with allostatic load the HPA axis can be imperfectly self-regulating, or imperfectly able to adapt to stress, with deleterious consequences related to chronic ovesecretion of glucocorticoids. The perfectly adaptive stress response is one stimulated by a stressor, sustained for an appropriate time, and then turned off. Maladaptive responses, leading to allostatic load, come in four main varieties (McEwen 1998): The first consists in the sheer excess of stressful events; each "emergency response" of the HPA axis comes at a certain cost, for example the risks

associated with a sudden rise in blood pressure, where over repeated "hits" the costs can accumulate into an ischemic catastrophe. The second type of allostatic load is a lack of habituation to repeated stressors; the third is the failure of the HPA axis to recover after a reasonable time; and the last type consists of an inadequate or flat HPA system activation: not reacting at all to a stressor can be itself extremely damaging, not only behaviourally but physiologically, as for example when inflammatory cytokines are not counterbalanced by the anti-inflammatory glucocorticoids (McEwen 1998). This last type of allostatic load concerns the deleterious effects of subpar cortisol concentrations, but the first three types are concerned with the risks at the other extreme, of moving too far towards the other end of the inverted-U shape, or "healthy window" of adapted HPA system reactivity.

Psychologically perceived stress activates the HPA axis, and too frequent stress, or too inadequate habituation or recovery of the stress response, may be deleterious both behaviourally and physiologically. Stress is associated with a functional disconnection of frontal lobe functions from an organism's behavioural repertoire (Arnstern and Goldman-Rakic, 1998). Frontal lobe functions are associated with slower, conscious information processing, which under stress may be selectively abated in favor of quicker posterior cortical and subcortical functions; this may have survival value, but the chronic activation of this process can hinder adaptation to social situations demanding careful consideration rather than rapid reactions (Arnstern et al, 1998).

Physiologically, elevated cortisol (in animals, corticosterone) concentrations are not directly damaging to hippocampal neurons, and acute (short term) mild stress often has no effect on, and in certain conditions can even improve, hippocampal-dependent memory (Cahill and McGaugh, 1996; Sandi, Loscertales and Guaza, 1997; Stansbury, Haley and Koeneker, 2000). This is not surprising when one considers the role of cortisol in attention, that is in selecting for relevant stimuli while filtering irrelevant stimuli. But chronic stress and chronic hypercortisolism disturbs memory and adaptation through several mechanisms. First, cortisol causes energetic disturbances by inhibiting glucose transport and by activating a cascade of excitatory amino acid neurotransmitters, causing excessive mobilization of postsynaptic calcium, which in excess becomes damaging to cellular cytoplasm. This renders hippocampal neurons more vulnerable to other insults, for example hypoxia (such as results from sleep apnea, a condition more common in seniors than in young adults) or hypoglycemia. Over years, the indirect and initially trivial, but nonetheless deleterious effects of hypercortisolism on hippocampal neurons can accumulate and contribute to the development of neurotoxic effects on hippocampal neurons (Sapolsky 1992; Deshmukh and Deshmuck, 1990).

Lastly, allostasis—that is, the self-regulating process by which the HPA axis turns the stress response on and off—is itself associated with wear and tear on the organism (Sterling and Eyer, 1988). Thus the HPA axis protects the body from stress; but if the

stress response is elicited too often by external or internal stressors; if the stress response fails to habituate to repeated stressors; if the stress response fails to shut off at an appropriate time; or if the stress response is absent or of inadequate magnitude, then a system designed to protect the body from stress becomes itself a significant stressor to many body systems.

3.6.2 Periodicity

Corticosteroid concentrations vary in a circadian cycle, defined by the duration and frequency of secretory episodes rather than secretory rate; in healthy adults, the mean plasma cortisol concentration reaches its zenith shortly after awakening, falling irregularly thence until bedtime, when it begins rising again during the hours of darkness (Weitzman, Fukushima, Nogeire, Roffwarg, Gallagher and Hellman, 1971). There are many minor plasma cortisol concentration peaks during the day, mostly related to mealtimes or the anticipation of mealtimes (Krieger, Allen, Rizzo and Krieger, 1970), or some concatenation of physical, physiological, or psychological stress (Sapolsky 1992). In the presence or anticipation of a stressor, peak physiological concentrations of cortisol of up to 4000% (or 40-fold) above baseline are reached within 15 to 30 min, and return to normal within another 60 to 90 min (Goya, Rivero and Pascual-Leone, 1995). More modest but robust rises of 50-75% are reached within 30 min after awakening in children, adults and seniors of both sexes (Pruessner, Wolf, Hellhammer, Buske-Kirschbaum, von Auer, Jobst, Kaspers and Kirschbaum, 1997).

The higher basal cortisol levels found in both healthy seniors and AD patients, compared to young adult controls, is more pronounced in the evening and night time (Dori, Casale, Solerte, Fioravanti, Migliorati, Cuzzoni and Ferrari, 1994). Dori and coworkers (1994) also reported that the nadir (23:00) values of plasma cortisol was significantly and directly correlated to age. This finding was recently confirmed and extended by Raff and coworkers (Raff, Raff, Duthie, Wilson, Sasse, Rudman and Mattson, 1999), who found that in healthy elderly men and women cortisol levels were significantly higher than those of young healthy controls at 23:00 hours but not at 7:00 hours.

Knowing the boundaries of normal periodicity of cortisol secretion can thus help identify incipient HPA-axis abnormalities in the absence of frank hypercortisolism. For example, a subject could have a normal range of cortisol concentrations at any given time, but abnormal pattern of circadian peaks and troughs across time.

3.6.3 Cognition

The HPA axis is closely involved in the formation of memory. Specifically, three stages in the process of memory formation depend upon adequate glucocorticoid (primarily cortisol) levels: arousal, selective attention, and memory consolidation. Impairments on these three levels directly influence declarative memory. Arousal means a general state of readiness to attend to stimuli. Selective Attention (also called sensory integration in the cortisol literature) refers to the ability to focus on causal (or

relevant) factors in the environment to the exclusion of non-causal (irrelevant) factors, in order to select an appropriate response (Lupien and McEwen, 1997). Cortisol has an inverted-U shaped function with both arousal and selective attention (Henkin, McClone, Daly and Bartter, 1967; Wolkowitz, Reus, Weingartner, Thompson, Breier, Doran, Rubinowand and Pickar, 1990; Wolkowitz, Weingartner, Rubinow, Jimerson, Kling, Berretini, Thompson, Breier, Doran, Reus VI and Pickar, 1993; Lupien and McEwen, 1997). The consolidation of memory refers to the process of establishing long-term, possibly permanent memories. One model of the cellular substrate of learning consists in the formation of networks of Long-Term Potentiation (LTP). LTP is defined as the increased susceptibility of one neuron to be stimulated by another after a period of intense stimulation by that other (Kandel, 1991). LTP networks are formed predominantly under glucocorticoid modulation in the hippocampus, the brain structure containing the highest density of glucocorticoid receptors (Lupien and McEwen, 1997; McEwen 1987). The hippocampal formation (including the entorhinal cortex) is the first brain region affected in mild AD, but increasing AD severity seems to shift further damage from hippocampal formation to temporal neocortex (Detoledo-Morrell, Sullivan, Morrell, Wilson, Bennett and Spencer, 1997), which transition, it is hypothesized, also marks the appearance of the more serious semantic, as opposed to only episodic, memory impairments more consistently across patients than earlier in the course of the disease (Hodges and Patterson, 1995).

The brain has two type of cortisol receptors, Types I and II, in different regions of predilection. In the hippocampus, which contains both types of receptors, the process of memory formation is aided or hindered in characteristic ways corresponding to the two receptor types. Agonist and antagonist studies of these receptors have shown that Type I receptor activation is involved in the process of selective attention, that is in evaluating a situation and selecting an appropriate response; while Type II receptor activation is particularly involved in the process of memory consolidation (for review, see Lupien and McEwen, 1997). Elevated cortisol levels are thus associated with decrements in three components of information processing which are necessary for declarative memory formation: arousal (Born, Hitzler, Pietrowsky, Pauschinger and Fehm, 1989), selective attention, and consolidation.

The evidence is that acute excess cortisol causes reversible hippocampal atrophy and reversible declarative memory loss, and that chronic excess cortisol causes irreversible hippocampal atrophy and irreversible declarative memory loss (Martignoni, Costa, Sinforiani, Liuzzi, Chiodini, Mauri, Bono and Nappi, 1992; Issa, Rowe, Gauthier and Meaney, 1990; Sapolsky, Krey and McEwen, 1986; Ling, Perry and Tsuant, 1981; Landfield, Waymire and Lynch, 1978). This finding also holds in humans. Significant declarative memory impairments and a 14% reduction of hippocampal volume have been found in seniors showing significant basal cortisol increase over 5 y (Lupien, de Leon, De Santi, Convit, Tarshish, Nair, Thakur, McEwen, Hauger and Meaney, 1998).

Excess or rising basal cortisol levels have also been linked to the cognitive deficits observed in Cushing's disease (Starkman, Gebarski, Berent and Schteingart, 1992), depression (Sachar, Hellman, Rofwarg, Halpern, Fukushima and Gallagher, 1973), healthy seniors (Lupien et al, 1998; Lupien, Gaudreau, Tchiteya, Maheu, Sharma, Nair, Hauger, McEwen and Meaney, 1997) and AD (Davis, Davis, Greenwald, Mohs, Mathe, Johns and Horvath, 1986). What is interesting is that in all cases of disturbed HPA activity, high cortisol levels were also associated with subjective memory complaints (Starkman et al, 1992), depressive symptomatology (Lupien et al, 1995), or both.

3.7 Depression

Commonality Between Depression and AD

Depression has often been seen as a condition that complicates AD diagnosis, but whose differentiation from AD is relatively straightforward when supported by a neuropsychological evaluation (Lamberty and Bieliauskas, 1993). Of late, however, more attention has been paid to the possibility that in certain cases depression and AD may represent two different stages of the same disease, rather than two independent diseases (Lupien, Nair, Brière, Maheu, Tu, Lemay, McEwen and Meaney, 1999; Wetherell, Gatz, Johansson and Pedersen, 1999). Supporting this possibility are three important factors shared by AD and depression. Two of these have to do with the dexamethasone resistance test. In this test, the synthetic steroid dexamethasone is

administered at 11 pm, and the next day basal cortisol levels are measured at 8 am; in healthy subjects, these levels should have dropped because of the inhibitory effect of the HPA axis once it detects, via cortisol receptors in the pituitary gland and hippocampus, acutely elevated cortisol levels. In some patients, cortisol levels do not drop, a phenomenon called dexamethasone (DEX) resistance. DEX resistance is often (but not always) associated with hypercortisolism, because chronic high levels of cortisol damage cortisol receptors in the hippocampus, one of the most important links in the negative feedback loop of the HPA axis. The first factor that AD and depression have in common is that in both diseases increasing age interacts with HPA activity. Late-onset AD patients are more likely to become DEX resistant than early-onset AD patients, and AD patients who *are* DEX resistant become more so with time. In depression, cortisol levels increase with age, which accounts for approximately 20% of the variance in cortisol levels (Halbreich, Asnis, Zumoff, Nathan and Shindledecker, 1984). Second, the incongruent findings of DEX resistance along with normal cortisol levels, or conversely, of no DEX resistance along with hypercortisolism, have been reported for both AD (Miller, Sastry, Speranza, Lawlor, Mohs, Ryan, Gabriel, Serby, Schmeidler and Davis, 1994) and depressed patients (Wolkowitz et al, 1990). Third, hypercortisolism is associated with decrements in cognitive performance in both AD (de Leon, McRae, Tsai, George, Marcus, Freedman, Wolf and McEwen, 1988) and depression (Van Londen, Goekoop,

Zwinderman, Lanser, Wiegant & De Wied, 1998). Another important factor that may mediate the relationship between AD and depressive symptomatology is stress: although the relation between past and present stressful life events and the first or subsequent onset of depression is complex (for review see Hammen, 1992), there is evidence that either current stressors or early life stress could be a contributing factor to depression (Pianta and Egeland, 1994; Lewinson, Allen, Seeley and Gotlib, 1999; De Marco, 2000). In rats, in whom stress can also cause depression stress leads to hippocampal cortisol receptor down-regulation (Moreau, 1997); and it is believed that the DEX resistance observed in some depressives may stem from a similar stress-induced down-regulation of cortisol receptors. In effect, their hippocampal receptors would be adapted to such high levels of cortisol that the additional exposure to the synthetic steroid would fail to trigger the shut-off mechanism of the HPA axis (Sapolsky et al, 1986). The psychological activity of negative anticipation or worrying can also lead to HPA axis activation (Schulkin, McEwen and Gold, 1994; McEwen, 1998). This suggests a possible neuroendocrine explanation for the controversy in the literature as to the association between subjective memory complaints and depression.

Cognitive Consequences of Depression or Alzheimer Disease

Depression has disturbing effects on typical frontal lobe functions (George, Ketter and Post, 1994), with correspondingly disturbed patterns of frontal lobe regional cerebral blood flow, which normalize upon recovery (Dolan, Bench, Brown, Scott and

Frackowiak,1994; Bench, Frackowiak and Dolan, 1995). Frontal lobe functions involve planning and executing appropriate goal-directed behaviour, initiating and monitoring voluntary responses and inhibiting inappropriate responses, as well as abstract thought and shifting attention and cognitive sets (for review see Miller, 2000). The cognitive consequences of depression have been probed using (for example) the Wisconsin Card Sorting Test, a commonly used neuropsychological index of frontal cortex function; results however have been inconsistent and difficult to relate to purely frontal functions (Merriam, Thase, Haas, Keshava and Sweeney, 1999). In contrast to this "localisationist" approach, other investigators have tried to probe depression's effects on cognition using behavioural, performance-based measures that are less theoretically linked to any given brain region. One research has concerned the dissociation between implicit and explicit memory performances in depressed persons. The dissociation consists in depressed patients' having impaired explicit but preserved implicit memory compared to normal controls (Danion, Willard-Schroeder, Zimmermann, Grangé, Schlienger and & Singer, 1991). These findings were not confirmed by Rohling and Scogin (1993), who tested explicit and implicit memory in depressed patients, psychiatric in-patient controls, and normal controls; they found no effect of depression on explicit (effortful) memory. They postulate that their results differ from others in the literature in that they accounted for the effect of hospitalization, with its attendant medication and isolation (Rohling and Scogin,

1993).

Other researchers have examined other components of memory, and found that the key impairment of depressed persons is that they encode (rather than retrieve) less well than normal controls, because they elaborate less or transform less the to-be-remembered material, with the impairment increasing as its "obviousness" or coherence of the relations between its elements decreases (Weingartner, Cohen, Murphy, Martello & Gerdt, 1981).

Depression is also associated with impairment in some measures of verbal fluency, particularly in tests of fluency which require the respondent to produce exemplars belonging to a given category (e.g., animals). On the other hand, depressed persons show *less* impairment in categorical fluency than AD patients, who themselves fail to benefit from the semantic cueing of a categorical fluency task as compared to a purely verbal (i.e., generating words beginning with a given letter) fluency task (Hart, Kwentus, Taylor and Hamer, 1988; Geffen, Bate, Wright, Rozenbild and Geffen, 1993). Although categorical fluency is considered less demanding than letter fluency (as reduced regional cerebral blood flow would suggest), AD patients may be particularly impaired in this task because of a possibly pathognomic loss of functional (i.e., necessary for optimal performance) asymmetry in the frontal lobes (Falgatter, Roesler, Sitzmann, Heidrich, Mueller and Strik, 1997). This has led some researchers

to suggest that a less-demanding task such as categorical fluency could be useful to differentiate depressed from demented patients (Geffen et al, 1993).

Depression is associated with preferential remembering, or, depending on the experimental design, preferential forgetting of negative-valence material. This phenomenon manifests in two characteristic ways in depressed persons: as the "State-Dependent Memory Effect" and as the "Mood Congruence Memory Effect." With the State-Dependent effect, material memorized in a given emotional state is more easily recalled in the same emotional state. In the case of the Mood Congruence Effect, of particular interest to our purpose, the degree of fit between the emotional tone of the material to be remembered and the mood of the subject affects that subject's ability to remember the material. Although this latter effect is widely reported in the literature, the direction of the effect is variable and seems to be sensitive to experimental conditions (Bazin, 1991).

As a example of the varying results in the literature, one could cite Danion, Kauffmann-Muller, Grange, Zimmermann and Greth (1995), who studied subjects with major depression. Subjects were presented with positive-, negative- or neutral-valence material. For recall, depressed subjects had worse perfomance with neutral than with positive or negative valence material; for recognition, depressed subjects' performances were worse for negative than for neutral valence material. There was no

valence effect on implicit recall.

Bradley, Mogg and Williams (1994), by contrast, studied high and low negative-affect subjects (a non-clinical sample who differed in the degree of depressive and anxiety symptoms); presented depression- or anxiety-relevant or neutral material; and found that high negative-affect subjects recalled more depression-relevant than neutral material, but that this valence effect was only apparent if anxiety scores were partialled out (Bradley et al, 1994). They also found a valence effect on implicit recall, with high negative affect subjects demonstrating a greater priming effect for the depression-relevant material than for the neutral material.

There have been several attempts to clarify the issue or to show how it can not be resolved within a simplistic conceptual model of mood-congruence. Another researcher's findings, that depressed subjects encode and remember more positive- than negative-valence material, were interpreted by distinguishing personal (experiential) memory from non-personal (non-experiential) memory (Calev, 1996). In this view, depressed subjects do recall greater mood-congruent (i.e. negative-valence) material from their personal lives, but the effect is reversed for new non-personal material; Calev (1996) speculates that this preferential encoding and retrieval of positive-valence non-experiential material could contribute to recovery from depression. More recently, Derouesne (2000) summarized this research by

emphasizing the exquisite dependence of mood-congruence outcomes on the methodologies and theoretical frameworks employed, a view which agrees with that of Bazin (1991). Derouesne (2000) also commented that while strongly emotional experiential memories can be especially vivid and persistent (as in "flashbulb memories"), strong emotions during recall can produce limited memory impairments or "lacunar amnesia."

Finally, another methodology employed in research on mood-congruence effects is that of information-processing time. Stip and Lecours (1992) examined the effects of the valence of a word in a lexical decision task. Depressed subjects and normal controls were presented with words of negative or neutral valence and non-words; subjects had to respond by pressing a key to signal whether or not the presented word was known to them. Depressed subjects had significantly slower reaction times to words of negative valence than to neutral words. Stip and Lecours interpreted their findings of the light of Ellis and collaborators' (1984, 1985) resource allocation model of attention in depression. Thus, for depressed subjects only, a list of depressing words acts to induce a further drop in attentional resources, in a sense exacerbating an already existing impairment of cognitive capacities in depressed subjects (Stip and Lecours, 1992). This impairment is postulated to consist of cognitive resources tied up in rumination—mulling over sad thoughts—and also, in the case of negative-valence items, of some type of interference between the negative-valence ideation within the

subjects and the negative-valence material to be remembered. It is an empirical question, however, whether subjects with depressive symptomatology would show impaired recall of negative-valence story elements, rather than merely delayed response, as in the lexical decision task used by Stip and Lecours (1992). Whether negative mood at recall improves or impairs recall of similarly negative-valence items, however, is unclear in the literature; Teasdale and Russell (1983) suggested that for valence effects to become apparent it may be necessary that the items to be recalled have some relevance to the subject, i.e. from having lived them or by association.

3.8 Subjective Memory Complaints

The literature on memory complaints (MEMs) reflects one basic inconsistency: some authors reports that MEMs correlate more with depression than with objective memory performance, while others report that MEMs *do* at least partly reflect accurate self-appraisal of objective memory abilities. Certainly much of the variation in study results can be attributed to variations in the MEM questionnaires used (Wilson and Evans, 1996). For example, O'Boyle et al (1990) found that MEMs were not correlated with memory problems, but rather with depression and affective status. Their study made use of a 24-item, 4-point, ad hoc questionnaire of unknown psychometric properties. O'Hara et al (1986) also used an in-house developed questionnaire, which revealed no differences on objective memory performance by MEM or depression diagnosis group; however, interestingly, they *did* find that MEMs were correlated with

the severity of depressive symptoms (rather than with diagnosis of depression). This finding concurs with Kendler and Gardner's (1998) report that the severity of depressive symptoms do not seem to vary discretely by depression diagnosis, but rather continuously. A recent review also emphasized that in neurodegenerative diseases such as AD, depressive symptoms are found to vary as in a continuum, rather than by distinct clinical entities (Kumar and Cummings, 2001). Thus, in this project we did not study subjects with a diagnosis of depression, but rather subjects with a continuum of depressive symptoms.

Smith et al (1996) utilized the Memory Functioning Questionnaire (MFQ) (Gilewski, Zelinski and Schaie, 1990), a 64-item, 7-point Likert-scale, and reported that those who deteriorated over a test interval did have more MEMs initially, but that their predictive power was very weak. Possible problems with the MFQ is that it is very long, and provides only one direction for the answers to go, i.e. there is no way for subjects to indicate that their memories have actually improved in some respects compared to their younger years. Using a single-item, 3-point scale for MEMs and a single-item, 4-point scale for depressive symptoms, Grut and coworkers (Grut, Jorm, Fratiglioni, Forsell, Viitaen and Winblad, 1993) determined that MEMs reflected objective memory performance only from subjects with mild AD (more impaired AD patients tending to underestimate, and depressed patients to overestimate, their memory problems). Despite the crude measures, this finding is significant because

mild AD is the very stage where MEMs would have the most information value. Schmand and coworker (1996, 1997) also found that MEMs accurately reflect memory problems, but only to a small degree compared to the predictive power of age or affective status. The scale they used was the Subjective Memory Scale, a 10-item, 2- to 4-point scale. The Subjective Memory Scale is brief, with high internal consistency (Cronbach's alpha: 0.72); however, being unipolar like the MFQ, it offers no possibility of indicating memory improvement, only deterioration.

Jonker and coworkers (1996) utilized a binary yes/no scale for MEMs and reported that, at least in non-AD, non-depressed subjects, MEMs did correlate with memory problems. However, depression as an exclusion factor may not be indicated for a study of MEMs, given on the one hand, the data showing that depressive symptomatology—but not diagnosis—correlates with MEMs (O'Hara et al, 1986), and on the other hand, the data on affective disturbance as an early indicator of AD (Jost and Grossberg, 1996; Royall 1997). There is clearly something here that needs disentangling.

Crook et al (1992) used the Memory Assessment Clinics Questionnaire (MAC-Q), which has six 5-point items that asks the respondant to rate his or her memory compared to how it was in young adulthood. Five items ask about memory abilities of daily life, and a sixth item asks for a general rating of memory ability. Since it allows

the scoring of subjective improvement as well as deterioration in memory abilities, the MAC-Q is the only true (because bipolar) Likert scale reported in the MEM literature. Using this instrument, Crook et al (1992) reported that MEMs were uncorrelated to age, depressive symptoms or education, but were significantly higher in women than men. The MAC-Q score also accounted for a small but significant proportion of variance in objective memory test scores.

Altogether, these data show that the literature is inconsistent on the relation of memory *complaints* to objective memory *performance*; the data further suggest that this inconsistency may be due to overly simplistic (yes/no), complex (64 questions with five possible answers each), or limited (no "improvement" answer possible) questionnaires. Lastly, it is also possible that the failure of some investigators to find significant correlation between memory complaints and performance may be due to the use of insufficiently sensitive neuropsychological tests (Lupien et al, 1994).

3.9 Rationale for Combined Markers

3.9.1 Summary of Previous Sections

HISTOPATHOLOGY: By the time it is diagnosable, AD is already well established neuropathologically.

NEUROPSYCHOLOGY: Retrospective identification of neuropsychological markers has relied on markers that are not sensitive to the very mild forms of cognitive impairment. Screening measures hold promise as a means for selecting patients with

MCI at risk for AD, but only further longitudinal studies will permit a clearer characterization of the utility of several commonly used cognitive assessment instruments.

DIAGNOSTIC UNCERTAINTIES: Antemortem diagnosis of AD remains highly variable across diagnostic criteria, with prevalence rates varying by a factor of 10 according to the criteria used (Erkinjuntti et al, 1991). Differences in diagnostic criteria are not the only potential problem; for example, the NINCDS-ADRDA diagnostic definition of AD (McKann, Drachman, Folstein, Katzman, Price and Stadlan, 1984), which was recently recommended as sufficiently reliable and valid by the American Academy of Neurology (Knopman, DeKosky, Cummings, Chui, Corey-Bloom, Relkin, Small, Miller and Stevens, 2001), has been criticized for an arbitrary age limit for onset of between 40 and 90 years (Leach and Levy, 1994), which goes counter to recent work on detecting signs of incipient AD in ever younger patients—and neglects the oldest old. Identification of MCI cases is beset with still more difficulties; prevalence rates have been found to vary by a factor of 50 according to the criteria used. Diagnosis of depression is also beset with conceptual difficulties, since the intensity of depressive symptoms seem to vary continuously, rather than discretely as per diagnostic (e.g., DSM-IV) criteria. Moreover, certain depression-like symptoms, namely apathy and social withdrawal, may be more closely associated with a subsequent dementing process than diagnosable depression itself. The imposition of strict diagnostic criteria in MCI research may not be suited for all

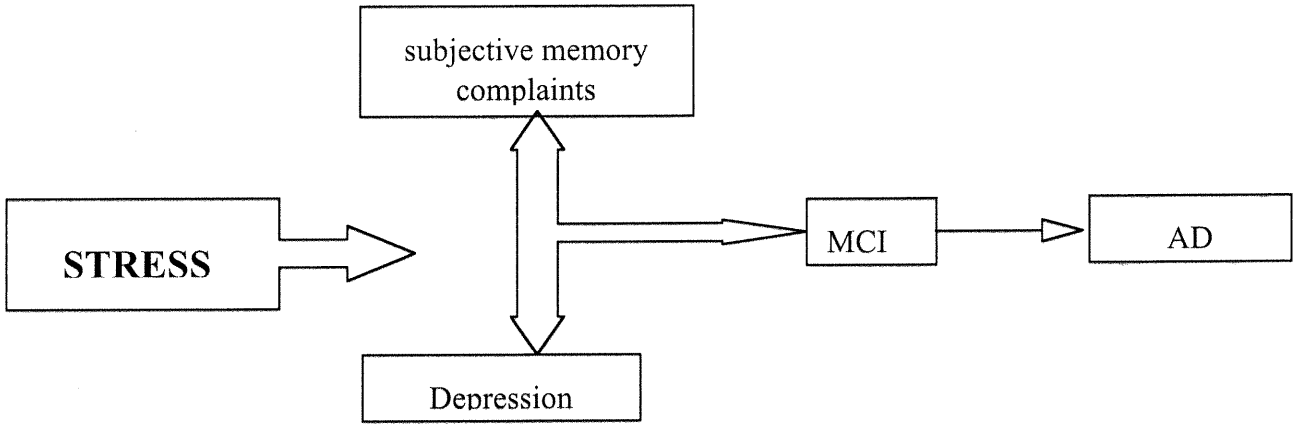
contexts; indeed, given the rapidly evolving state of neurobiological knowledge, and the continuing debate as to the optimal definition of MCI (Ritchie et al, 2001), it could be argued that experimental neuropsychology has a responsibility to systematically evaluate alternative definitions.

DECLARATIVE MEMORY is the first cognitive function to be affected by AD. Though the issue is controversial, it seems as if AD is associated with an impairment at the encoding or consolidation stages of immediate memory. When controlling for performances on immediate recall, delayed (or retrieval) memory does not reliably differentiate AD from normal elderly controls. Normal aging is associated with a slight impairment in encoding capacity, whereas AD is associated with impairment in both encoding and in working memory. Thus, MCI as a prodrome of AD should be associated with both types of impairments, whereas MCI as a benign, non-progressive condition should be associated only with the first type of impairment.

The relevance of NEUROENDOCRINOLOGY, DEPRESSIVE SYMPTOMATOLOGY AND SUBJECTIVE MEMORY COMPLAINTS to the present attempt at elucidating some of the concomitants of MCI is that all these phenomena share some important factors in common. Disturbed neuroendocrine function is associated with increased cognitive dysfunction in depression, and subjective memory complaints have been associated with depression and other mood disorders as well as with objective cognitive

impairment. Depression has been seen as a prodromal stage of AD. Given that subjective memory complaints (MEMs) are theoretically linked to stress and concomitant hypercortisolism, which in turn is associated with selective damage to hippocampal neurons and functional frontal lobe disconnection, one can draw up a model of disease progression from MEMs and depression, through MCI to AD, linked at every step by chronic high stress and its attendant hypercortisolism (**Figure 3**). In this model, instead of trying to differentiate MEMs from depression, the two are seen as intimately linked; however, only some patients who present with either MEMs or both MEMs and depressive symptoms have objectively verifiable mild cognitive

Figure 3



impairment, and of those, only some go on to develop AD. The challenge then, in dealing with subjects with memory complaints, or memory complaints and depressive symptoms, is to discover the characteristics that differentiate those who in whom such

complaints signal the presence of benign, non-progressive MCI—or some other cause of their memory complaints—from those in whom such complaints signal the presence of incipient AD.

Rowe and collaborators (Rowe, Steverman, Walker, Sharma, Barden, Seckl and Meaney, 1997) have reported that antidepressant treatment can restore HPA function in aged, cognitively impaired rats by up-regulating glucocorticoid receptors. This suggests that a similar early intervention in humans with mild cognitive impairment and hypercortisolism caused by blunted HPA hippocampal negative feedback can possibly prevent progression to AD. Of course, an even better window of opportunity would be still earlier, before the onset of MCI. In this light, taking subjective memory complaints as a symptom of depression, and both MEMs and depression as symptoms of neuroendocrine dysfunction, some particular constellation of all three factors, still to be discovered, could serve as an inexpensive and practical early warning sign of hippocampi at risk.

3.9.2 The Need for Multiple Markers

As the previous sections have made clear, researchers looking for the earliest possible sign of AD have grappled with the problems of insensitive instruments, premature homogeneity of research samples defined by criteria that are to some degree arbitrary, and variability in disease expression. It seems safe to assume that no single discipline—psychology, neuropsychology, endocrinology, and genetics—can or will

supply by itself the universally valid marker of incipient AD. What is becoming clearer, however, is that HPA axis dysfunction is associated with many of the pathologies of aging, specifically late-onset depression and late-onset AD; and that subjective memory complaints, which have been linked to depression, should not for all that be discounted, since depression can itself represent the prodromal stage of AD. The hope that this project is putting to the test is that of increased informational value garnered by the use of a multidisciplinary approach to attempt to identify a particularly ominous combination of markers on multiple levels, social, psychological, neuroendocrine and neuropsychological. Obviously this question can only be firmly answered with the help of longitudinal studies. This project is a first step in that direction. By helping to characterize the interrelationships of subjective memory complaints, mild cognitive impairment, neuroendocrine status, and social factors in subjects drawn from the population and already fearful about their memories, it is hoped that some of those fears can be put to rest—or directed to productive action, as need be: it has been shown that patients with many medical conditions can improve their lot by developing coping skills and cultivating a supportive social milieu (for review, see McEwen, 1998).

4. Goal of the Project

At a general level this project is concerned with the search for the earliest possible signs of Alzheimer Disease. One of these possible first signs is a subjective memory

complaint: a person reports memory problems. An obvious question about subjective memory complaints is accuracy: Do they accurately reflect impaired memory compared to an earlier time in that person's life, or do they relate rather to depression, or to a surfeit of stress in daily life? The controversy in the literature is that some researchers have found that subjective memory complaints relate more to depression or other mood states than to present or future (in longitudinal studies) cognitive performance; while other researchers have found that memory complaints *do* reflect with at least some degree of validity present or future cognitive performance. Because subjective memory complaints in seniors are ubiquitous, easy to elicit and possibly informative, this question is worth resolving.

Specifically, the goal of this project is to answer the question: Is the relation between subjective memory complaint and objective cognitive performance the same in both subjects with and without depressive symptoms? In other words, is the relation between subjective memory complaint and objective performance, such as it is in non-depressed persons, altered by the presence of concomitant depressive symptoms? There are three possible answers to this question. Let us call a group of persons with subjective memory complaints but no depression "Mem", and a group with both subjective memory complaints and depressive symptoms, "MemDep". Thus, if one were to submit these two groups to a neuropsychological test battery in order to compare their cognitive abilities, the three possible outcomes are:

(1) Mem = MemDep;

(2) MemDep > Mem; and

(3), MemDep < Mem.

The first outcome (1) would beg the question, when two subjects have equivalent memory complaints and equivalent memory abilities, why does one subject have depressive symptomatology and the other not? A possible answer to this question would have to do with the subjects' stress and stress hormone status. This forms one of the hypotheses which this project will test.

The second outcome (2), with the MemDep group having cognitive performance superior to that of the Mem group while having the same degree of memory complaint, would suggest the presence of a cognitive bias in the MemDep group that would lead them to underestimate their cognitive capacities. This forms another of the hypotheses which this project will submit to the test.

The third possible outcome (3), with the MemDep group showing worse objective performance than the Mem group, would suggest that the depressive symptoms in the MemDep group could be another reflection of a presumed underlying pathophysiological process that also gave rise to their objective and self-perceived memory problems (an objective, "bottom-up" explanation). Or one could posit that the depressive symptoms in the MemDep individuals are due to the depressing awareness

of mental decline, of not being able to accomplish what had once been cognitive second nature (a subjective, "top-down" explanation). Finally one could posit a synthesis of both subjective and objective angles of explanation (i.e. by invoking metacognitive, biological and environmental factors).

This project makes use of a multi-modal assessment procedure to provide a *thick description* of a group of self-selected seniors with memory complaints, with or without concomitant depressive symptomatology. By "thick description" is meant a multi-disciplinary approach combining the points of view, not only of neuroendocrinology, neuropsychology, and social psychology but also the points of view of those we are studying: persons with memory complaints. This taking into account of a subject's own interpretation of his or her behaviour is what Geertz (1973/1977) meant by engaging in "thick description" in anthropological research. This approach was intended to help bridge cultures. It is hoped that the present research project, by using a stress hormone measure, combined with an intensive neuropsychological, psychosocial and demographic evaluation, will permit a more complete assessment than can be afforded by either of these measures taken separately, and perhaps provide converging evidence for the existence of subgroups within the subjective memory complaints rubric. Thus this project is also intended to in some small way help bridge two cultures, that of research and clinical care on the one hand, and that of the patient or research volunteer on the other.

5. Hypotheses

The specific hypotheses were as follows:

H1. Since depressed persons have a cognitive bias against stimuli of emotional of negative valence, the MemDep group will recall fewer neutral than emotional elements at both immediate and delayed recall, whereas there will be no difference in recall by valence for the Mem group.

H2. Since stress may cause depression, the MemDep group will have higher self-reported stress than the Mem group.

H3. Since depressed persons have higher cortisol levels than normal controls, the MemDep group will have higher cortisol than the Mem group.

H4. Since high basal cortisol levels can have deleterious effects on declarative memory, declarative memory scores will be inversely correlated with basal cortisol levels in both groups.

H5. Since depression is related to high cortisol in seniors, and high cortisol is related to declarative memory deficits, the MemDep group will have impaired declarative memory function relative to the Mem and Control groups.

6. Material and Procedure

6.1 Subjects

Advertisements were placed in four waves, over four months, in two French and two

English Montreal newspapers. The advertisement was headed Stress and Aging and called for volunteers interested in participating in a research project on stress and aging. Interested parties were invited to call in to volunteer or for more information if they had any combination of:

- worries about memory
- significant stress in last five years
- depressive symptoms (sadness, social isolation).

6.2 Evaluation

6.2.1 Classification Instruments

Subjects were classified as to the presence or absence of subjective memory complaints, depressive symptoms, and anxiolytic or anti-depressant drug usage. For the first two factors, standard questionnaires were used: the Geriatric Depression Scale, and the Memory Clinics Assessment Questionnaire.

Geriatric Depression Scale. Depressive symptoms were evaluated with the Geriatric Depression Scale (GDS) (Yesavage, Brink, Rose, Lum, Huang, Adey and Leirer, 1983). The GDS is a 30-item questionnaire dealing with the subject's experience of depressive symptomatology; in this study it was filled out by the interviewer in a face-to-face interview with the subject and a score of 12 or higher was taken to indicate depressive symptomatology. The GDS has been validated in seniors (Montorio and

Izal, 1996). Although it is not considered valid for demented populations (Stiles and McGarrahan, 1998), in this study it was used because the intended subjects would be not have been diagnosed as demented. The GDS is also the most used questionnaire in, and therefore most relevant to, current geriatric practice.

Memory Assessment Clinics Questionnaire. The Memory Assessment Clinics Questionnaire (MAC-Q) (Crook et al, 1992) is a six-item, 5-point self-scored questionnaire which asks the respondant to rate his or her memory compared to when they were in high school. For this study, all references to "high school" were amended to "high school or about eighteen years old". The first five items of the MAC-Q are scored 1 ("much better"), through 3 ("about the same"), to 5 ("much worse") points. A last general question about the subject's impression of memory change ("In general, how would you describe your memory as compared to when you were in high school?") is scored 2 to 10 points on the same subjective scale. The score range is thus 7-35. The minimum score would correspond to a subject rating his or her own memory as much better now than it was in high school (or at eighteen years of age). A score of 21 would indicate that on average the subject perceives that his or her memory is about the same as it was then. The criterion score of 25 for subjective memory complaints requires that several items be rated as "somewhat" or "much poorer now" (Crook et al, 1992). The MAC-Q has been chosen for four reasons: (1) it is brief and immediately intelligible to respondents, which is important since seniors often balk at

filling out lengthy questionnaires which they perceive as absurd or trivial (Flicker, 1988); (2) it is bipolar, which obviates the problems of dealing with a restricted range of scores because it allows the subject to rate improvement as well as deterioration in memory abilities; (3) it is uncorrelated to depressive symptoms, which reinforces the construct validity of memory complaints (Crook et al, 1992) and prevents spurious correlation between two questionnaires (for memory complaints and depression) that really only measure the same thing; and finally (4) since the MAC-Q seems to be gaining general acceptance in memory and aging research (e.g., Hänninen, Hallikainen, Koivisto, Helkala, Reinikainen, Soininen, Mykkänen, Laakso, Pyörälä and Riekkinen, 1995; Koivisto et al, 1995), its use also helps in the dissemination of research results.

6.2.2 Classification Procedure

Subjects were classified by presenting symptomatology after they had completed all evaluations. Classification (see below, **Results**) was effected based on results from the MAC-Q, the GDS, and the concurrent medication questionnaire. This produced three included groups: a **MEM** group composed of individuals with subjective memory complaints only; a **MemDep** group with memory complaints and depressive symptoms; and a **Control** group with neither significant levels of memory complaints nor depressive symptoms.

Exclusion Criteria

Those who were taking anti-depressants, or anxiolytics more than twice a week, were excluded because of the complexities of accounting for the possible cognitive effects of their medicine. Subjects who could not understand for be understood clearly in French or English were also excluded. Subjects less than 50 years of age were excluded, as were subjects with apparent psychiatric disorders (one subject with grandiose thought processes).

7. Assessment

The assessment procedure was designed to evaluate subjects on two levels. The first assessment level utilized standard tests such as are used in daily practice for screening in primary care and memory clinics. The second level featured a battery of tests, designed to generate in-depth information on multiple cognitive functions.

7.1 Standard

Volunteers accepted into the study took part in two evaluation session split over two days, for a total of about two and a half hours. Subjects first came to the *Institut universitaire de gériatrie de Montréal* for a testing session lasting approximately two hours. After the first evaluation session, subjects then left with a home package containing three days' worth of saliva sampling papers and psychosocial self-report questionnaires (see below). Following this, subjects returned to the *Institut* with their three days' worth of saliva sampling papers and completed the neuropsychological assessment procedure in a session lasting about 30 minutes.

During the first testing session, subjects were evaluated with:

1. The **standardized** (Molloy et al, 1991) **Mini-Mental State Examination** (Folstein et al, 1975), and the **2. Modified Mini-Mental State Examination (3MS)** (Teng and Chui, 1987). Both these instruments are standard cognitive assessment tests which are simple to administer and in wide usage; they are included in order to provide a rough, and generally understandable, estimate of the subjects' cognitive capacities. The 3MS is a modified version of the Mini-Mental, scored on a range of 0-100 rather than 0-30; on both tests higher scores indicate better performance and thus better inferred cognitive ability.

3. The **standardized** (Standish et al, 1996) **Alzheimer's Disease Assessment Scale—cognitive subscale (SADAS-cog)** (Rosen et al, 1984) is the standard assessment instrument for Alzheimer's Disease drug trials. It includes 3 trials of word-list learning to test immediate memory, as well as praxis, comprehension and language production tests. The SADAS-cog was included because of the expectation that it would be more sensitive and provide finer distinctions than either the MMSE or 3MS, while still being widely used in the geriatric and neuropsychological communities. Since the non-standardized form of the ADAS-cog is considered sensitive to very mild Alzheimer's Disease (Zec et al, 1992), one can surmise that the standardized (SADAS-cog) form, which is more sensitive to inter-group differences, would be even

more so. The SADAS-cog is scored on a range of 0-70, with higher scores indicating a greater degree of impairment.

4. The **Comprehension** and **Similarities** scales from the Wechsler Intelligence Scale for Adults—third edition (Wechsler, 1981) was used in order to provide a gross estimation of the subjects' judgment and abstract reasoning abilities.

7.2 Experimental

7.2.1 Declarative Memory

1. The **Experimental Declarative Memory Test** was developed by Lussier and collaborators (Lussier, Peretz, Belleville and Fontaine, 1989). It evaluates immediate and delayed declarative memory, as well as non-declarative memory. Declarative memory is evaluated by a cued recall test; non-declarative memory is evaluated by a word stem completion test which takes place between the immediate and delayed declarative recall trials. Both are derived from a list of 12 imageable and concrete word pairs. Unrelated word pairs are generally considered sensitive measures of declarative memory deficits; thus, the list is composed of six moderately related word pairs and six unrelated word pairs. The related and unrelated pairs are equivalent in word frequency, length, and grammatical category. For the declarative memory task, the subject was presented with the list of words pairs to be read aloud. The word list was presented twice in succession, on a Macintosh laptop computer. As in the Boston revision of the Wechsler Memory test (Milberg, 1986), on the second presentation the

words were presented in reverse order, for example: after an initial presentation of the pair **APPLE—FRUIT**, at the first recall trial the volunteer was shown **APPLE--?** and at the second recall trial, **FRUIT--?**. The Experimental Declarative Memory Test (EDMT) has been proven to be more sensitive than its near-equivalent, the Paired-Associates Learning (PAL) subtest from Wechsler Memory Test-Revised (WMS) (Wechsler, 1987), perhaps because in the EDMT the non-associated pairs are greater in number (6 compared to the PAL's 4) and better controlled for frequency of usage and semantic relatedness (Lupien, Lecours, Lussier, Schwartz, Nair and Meaney, 1994).

2. **Story recall with emotion factor** was similar to the Logical Memory subtest from the Wechsler Memory Scale, and was inspired by the work of Stip and Lecours (1992), who found that in a lexical decision task where subjects had to decide whether a string of letters constituted a word, depressed subjects took significantly longer to respond face to a "moodlist" of negative affect-laden words, than they did to respond to a neutral word list. In normal controls there were no significant difference in response time between the two lists. In this study we sought to extend these findings to a declarative memory task loaded with negative affect and neutral affect words.

The two-paragraph long story was composed to contain 26 story elements, half of which had previously been determined to be of high emotional (negative) value and

half of which were previously determined to be of low emotional (neutral) value (Appendix 1). The sex of the protagonist in the stories was matched to the sex of the subject. Immediate and delayed (25 minutes after Immediate Recall), trials were tested for 26 elements; a 10-element recognition trial (5 negative and 5 neutral-valence elements) followed. The expectation was that subjects with depressive symptoms would recall fewer neutral than emotional elements at both immediate and delayed recall, whereas there would be no difference in recall by valence for the subjects with memory complaints but no depressive symptoms.

3. The **Visual Reproduction Test** is from the Wechsler Memory Scale—Revised (Wechsler, 1987), and was included in order to evaluate nonverbal immediate and delayed recall, which are known to be affected both by aging and Alzheimer's Disease (Flicker, 1988). It consists of four drawings of increasing complexity which are presented one at a time, in order of complexity. The subject has to reproduce each drawing immediately after its presentation (I mmediate nonverbal recall) and, after a delay filled with other (non-memory dependent) tasks, has to try and reproduce all four drawings (delayed nonverbal recall). The subject's drawings are scored following published scoring criteria; the number of elements varies by the complexity of the drawing and ranges from 7 to 18 per drawing. On both the Immediate and Delayed recall trials, the maximum score is 41.

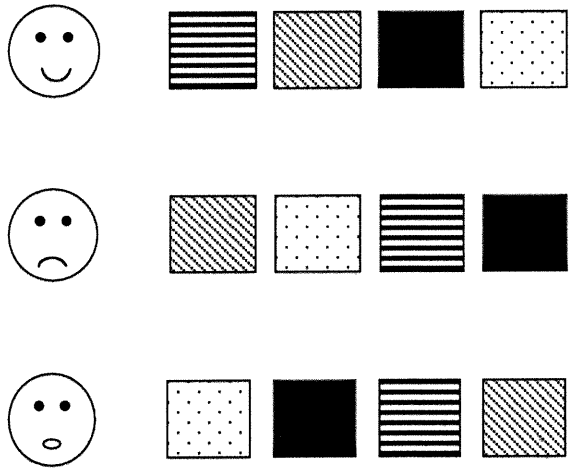
7.2.2 Attention, Psychomotor and Verbal Fluency

1. Simple and Choice Reaction Time. A slowing of cognitive performance with aging is well known (Welford, 1984). The rationale for including a test of Simple and Choice Reaction Time (RT) was to distinguish the non-specific effects of aging from the hippocampal-specific effects of incipient AD. Simple and Choice RT tests were performed on a Macintosh laptop computer. For the **Simple Reaction Time** test, the subject was instructed to hit the space bar as fast as possible every time an "X" appeared on screen; the subject was instructed to hit the space bar to begin the test, and to be ready for the first "X" to appear after that. There were 30 trials. The score was the median reaction time for correct responses. For the **Choice Reaction Time** test, the subject was instructed that either an "X" or an "O" would appear on screen, and every time either of one did, subjects had to hit the corresponding key on a two-key response box. The subject was instructed to fixate the cross ("+") symbol at centre-screen while waiting for either of the target to appear. Subjects were told that they had to give the *correct* answer, by hitting the X or O key, as fast as possible whenever that symbol appeared. In this test there were also 30 trials, 15 each for the two symbols. The score was the median reaction time for the correct answers. Errors were also counted. Another outcome measured was the difference between the median Simple and Choice reaction times: compared to young adults, seniors are known to show an increased difference between these two, which is linked to reduced cognitive

processing speed (Ferris, Crook, Sathananthan and Gershon, 1976).

2. Selective Attention. The information processing capacities of seniors are taxed more than those of young adults when they must monitor the environment for the conjoined appearance of two stimuli (Rabbitt, 1965). In the Selective Attention test the subjects had to determine as quickly as possible whether a target was present; the target was a *black square*. The response was a key press on the computer keyboard; a key on the right was labeled with a large **Y** and one of the left with a large **N**. The target was presented with a varying number (2, 4, 7 or 10) and kind (black or white circles or white squares) of distractors. Reaction times per number of distractors were calculated (Foster, Behrmann and Stuss 1995).

Figure 4



3. Conditioned Associative Learning Test (CAL). Age-related memory deficits are most apparent on tasks that make great demands on the ability of the subject to learn new associations and make fine distinctions between easily confusable elements (Spencer and Raz, 1995), particularly

when they have to both hold onto newly acquired material and attend to further inputs (Craik, 1994). The CAL, modified from Levine, Stuss and Milberg (1997), is sensitive to frontal lobe functions (Levine et al, 1997), which are known to be important in memory formation (Buckner, Kelley and Peterson, 1999) and are impaired early in the course of AD (Patterson, Mack, Geldmacher & Whitehouse, 1996). In this test, the subject was confronted with four face symbols and four differently patterned squares on a page contained in a binder, and had to learn an arbitrary, pre-established linkage between each face and each square through a process of trial and error (**Figure 4**). The subject was shown the first row, and the instructions were given as follows: "Ok, here's another test, this one's really tough. See, here there are four different faces, and four different squares. See, what you've got to accomplish here is learn which face goes with which square. Each face goes with one and only one square, and it never changes. They go in pairs you see; each square has its face and each face its square. The pairs were made up before the test and they never change. What you have to do is to find out what those pairs are; you have to find out which face goes with which square. Let's say that this face {experimenter points to "Oh" face] goes with this [black] square—I'm not saying that it does, this is just an example. So let's say that I point to this ["Oh"] face. So at the beginning you have no choice but to guess, you have no idea. Let's say that you point to this square [diagonal stripes.] I say, "No." So you guess again. Let's say you point here [horizontal stripes]; I say, "No." So you try

again, and let's say you point here [black square], and I say "Yes." Ah! so now you know that this face [pointing to "Oh" face] goes with this [black] square, and it will always go with this square, for the rest of the test. That means that when we go to the next row, if I point to this face ["neutral" face], you're not going to try this [black] square first, because you know that it goes with this other ["oh"] face. Are you all right with this, do you have any questions?"

There were 8 rows per page, and this first, unstructured trial continued until subjects succeeded in producing 8 correct first answers in a row or until the 32nd trial, whichever came first. If the 32nd trial was reached before the criterion of eight correct first answers in a row, the test continued with a structured trial, again comprised of 4 other pages of 8 trials per page (32 trials), with the difference that between each 8 trials they were shown the answer key which showed each face paired to its corresponding square. At each presentation of the key for this structured trial, they were told, "See, this square goes with this face, this one with that one, this one with that one, and this one with that one," with the experimenter slowly pointing with a pencil to each member of each pair as it was named.

The criterion for success in the structured trial was 8 correct first answers, with the proviso that the run of correct first answers had to commence immediately after a presentation of the key.

There are at least ten ways of scoring the CAL but in order to avoid redundancy in this study we concentrated on three: (1) The number of correct first responses in the first unstructured part, separated by runs of 8 trials; (2) the number of prior-response repetitions, defined as giving the answer that was correct for the immediately preceding pair as first choice of answer for the following pair; and (3) the total number of choices to reach criterion in the test, combining the unstructured part and the unstructured part if it was used.

4. Fluency. This test was included to evaluate the differential effect of aging on categorical (e.g. Animal) versus letter fluency. Categorical fluency has been found to be more impaired in AD patients than letter fluency (Rosen, 1980), and to better discriminate AD patients from normal seniors (Monsch, Bondi, Butters, Salmon, Katzman and Thal, 1992).

In this test we asked subjects to produce as many exemplars that belonged to a given category as they could within 1 min. There were two letter categories which entailed producing as many words as possible that start with a given letter (matched for word-frequency in French and English, according to the subject's most fluent tongue; S-A in English, P-F in French); proper nouns or variations on a theme (e.g. black, blacker, blackest) were not allowed. The other categories were First Names (either sex) and Animals.

7.2.3 Neuroendocrine

Infradian Salivary cortisol sampling. Salivary sampling was chosen in lieu of plasma sampling because it is less invasive. The infradian aspect (less than a day) also simplifies data collection, since night-time plasma sampling can be disrupting and cumbersome. In general plasma sampling has allowed accurate measurement of circadian cortisol cycles, but is less than ideal for glucocorticoid measurement in seniors, who may not agree to undergo the blood sampling procedure, or may be confused by the disruption in their routine, a disturbed state which could alter the very cortisol levels being measured. Thus we decided to measure cortisol levels in a natural environment using a non-invasive technique. Salivary cortisol levels have been shown to be highly reliable measure of free cortisol levels. Five times a day for three days at baseline, participants wet a $3^{1/2}$ by 5 cm paper with saliva by placing it in the mouth until the saliva front reaches just beyond the four cm line. The top 1 cm portion was kept dry for recording subject data. The paper was then air-dried and stored in a plastic bag until it was collected for analysis. Salivary free cortisol was ethanol-extracted and labeled with [^{3}H]cortisol and B-63, a cortisol-specific antibody (Endocrine Sciences, Terzana, CA, USA), for radioimmunoassay. Previous use of this method has revealed intra- and inter-assay variability of less than 1%, with good compliance by seniors. This method has been shown to lead to reliable cortisol measures in elderly subjects and to distinguish groups that differ in terms of cortisol secretion (Lupien et al, 1998).

There were five saliva sampling periods: the first was immediately upon awakening; the 2nd was 30 min later; the 3rd was at 2 p.m., the 4th at 4 p.m. and the 5th and final sample was taken upon retiring for the night. Subjects were told to be as precise as possible. They were told that some flexibility was allowed in that the two afternoon samples could be taken at the prescribed time ± 30 min, i.e. the 4 p.m. sample could be taken between 3:30 and 4:30 p.m.

7.2.4 Psychosocial

1. The **Derogatis Stress Profile** (Derogatis et al, 1997) is a stress questionnaire which combines the measurement of both positive and negative stressors; this is important because the absence of positive (rewarding) experiences in life can be as stressful as the presence of negative (aversive) experiences. This questionnaire was included in order to obtain an indication of the levels of stress experienced by the subjects. It is composed of several subscales indexing stress in various domains of experience: **Vocational** is the Vocational Satisfaction dimension, e.g. "I get great pleasure from the people I work with." This item is an example of the items scored negatively in the DSP (i.e. with 0 indicating "not at all true of me" to 4 "extremely true of me," some stress items are scored directly, and others are scored as 4 minus the subject's rating for that item), based on the rationale that the absence of a normally expected reward is itself a punishment. **Subjective** is a subjective rating done by marking a 10 cm line numbered 0 to 10 with a bar indicating the level of stress currently experienced.

Anxiety evaluates anxious thought patterns, e.g. "When I know I have something unpleasant to do, I worry about it for a long time." **Domestic** is for Domestic satisfaction, e.g. "Interacting with my family and friends is a source of great enjoyment for me." **Health** Posture refers to behaviour known to directly affect health, as in "I smoke too much." **Role** Definition refers to the subject's self-definition, both privately and to others, e.g. the stress-inducing belief "I really believe it is lonely at the top" and the stress-*re*ducing "I believe you can get a lot of help from others in getting the job done in life." Finally, **Attitude Posture** refers to the achievement ethic, as in "Most things I do I see as a challenge."

2. The **Perceived Stress Scale** (PSS) (Cohen, Karmarck and Mermelstein, 1983) was included because stress is ultimately a subjective experience. The PSS is designed to measure the degree to which the subject judges his or her present life situation to be stressful. The PSS consists of 14 items designed to measure the degree to which respondents find that their lives are out of their control, unpredictable and overwhelming. Items tap stress-related experiences such as feeling annoyed with things which are out of control, thinking about unmet goals, and feeling nervous and stressed about life in general. Each item is scored by the respondent on a 5-point Likert scale ranging from Never (0) to Very Often (4). The PSS score is obtained by reversing the scores of seven positive items and then summing all items.

3. The **Social Support Scale** (James and Davies, 1987) is a questionnaire which measures the positive social relationships in the interviewee's life (e.g., "Do you feel that you are an important part of your family's (or anyone else's) life?"); one question out of ten asks about relationships which are perceived as overly demanding. The non-redundant negative relationships are subtracted from the positive relationships and the sum constitutes the Social Support score. This instrument was included in order to provide the counterpart of the Derogatis Stress Profile, in that positive affect and depression has been found to be related to a specific combination of weak resources (social support among them) and very high levels of undesirable events (Murrell and Norris, 1984).

3. The **Symptom Checklist–90** (Derogatis, Lipman and Covi, 1973) is a questionnaire that evaluates nine primary symptom dimensions of psychiatric outpatients. This test is included in order to control for the possible presence of traits which could mediate the expression of subjective memory complaints. The nine dimensions rated are: Somatization, Obsessive-Compulsive, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation and Psychoticism.

4. The **Daily Activities Checklist** (Arbuckle, Gold, Chaikelson and Lapidus, 1994) is a self-report checklist of 22 different activities (e.g. gardening, reading, exercising, volunteer work) that the subject has engaged in from never to every day during the

past month. This test was included to measure the activity levels of seniors, which have been shown to account for some variability in intellectual abilities in general (Schaie, 1983) and verbal memory in particular (Arbuckle, Gold and Andres, 1986).

6. The **Trauma Symptom Checklist** (Briere and Runtz, 1989) is a questionnaire which measures symptomatology in adults stemming from traumatic experiences in adulthood or childhood. It was included because of the relationships between post traumatic stress disorder and depression (Tomb, 1994), and key brain regions involved in memory formation and emotional regulation (Bremner, 1999).

7. The **Self-Efficacy Scale** (Schwarzer, 1992) measures the degree of personal agency, i.e. the belief that one's actions are responsible for the successful outcome of any given course of behaviour. It is a ten-item scale (e.g. "When necessary I can be assertive") with each item scored from 1 to 4 by the respondent; the total score thus ranges from 10-40.

7.3 Statistics

Data were analysed with analysis of variance (ANOVA); significant main or interaction effects were followed by contrast analyses if the result was predicted by a hypothesis, or by post-hoc tests if not. The post-hoc test used was the Spjotvoll-Stoline test (Spjotvoll and Stoline, 1973; Winer, Brown and Michels, 1991; Statistica 4.1, Statsoft), which is an extension of the Tukey's Honestly Significant Difference

test for unequal sample sizes. Certain two-group comparisons were carried out with the non-parametric Mann-Whitney test.

8. Results

8.1 General Description of Sample

Seventy one people came to the first testing session, which took place one-on-one with the present author in a private office at the *Institut universitaire de gériatrie de Montréal*. Two declined participation after having been explained the nature of the project in person by the experimenter. The remaining 69 subjects were evaluated with the psychoneuroendocrine test battery.

The total evaluated sample thus comprised 69 subjects, 51 of which (74%) were women; the age of the participants ranged from 50 to 83 for both sexes and for women and men respectively the means were (± standard deviations) 67.4 ± 6.8 and 67.8 ± 6.6; years of education ranged from 5 to 21 (means 13.7 ± 3.3 and 15.0 ± 3.5).

The classification procedure utilized the results from two tests (see below) as well as a concurrent medication use interview. Persons presently taking anti-depressant medications, as well as those taking anxiolytic medicines (e.g. diazepam) more than three times a week were excluded from this study (but will be followed-up as part of Dr. Lupien's longitudinal research). Then, the remaining cases were classified using an algorithm based on the results from two tests, the Geriatric Depression Scale (GDS)

and the Memory Assessment Clinics Questionnaire (MAC-Q). Any person with subjective memory complaints but no significant depressive symptomatology (MAC-Q > 24, GDS < 12), and who was not taking anxiolytics more than occasionally or on anti-depressant treatment, was thus classified into the MEM group. Any person presenting with both subjective memory complaints and significant depressive symptomatology (MAC-Q > 24, GDS > 11) was classified into the MemDep group.

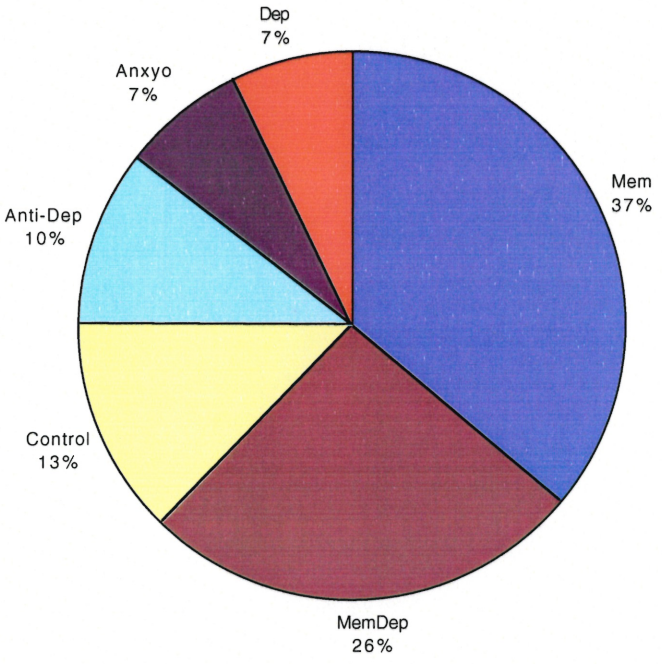
Originally it had been intended to compare and contrast only two groups, the Mem and MemDep groups. But only after everybody had been tested did we realize that among the volunteers who had answered the advertisements, some (9 out of 69, or 13% of the evaluated sample) turned out to have neither subjective memory complaints nor depressive symptoms. We decided to form a group with those who fit that criterion (the criterion, that is, of not matching the criteria for the other two groups), because we sought to study a sample of seniors who would tend to do just that, i.e. answer ads and seek attention and/or professional help because of their worries about their memory capacities, depressive symptoms, or stress levels. The classification procedure was thus amended after the fact to account for these subjects. Therefore, people scoring below the cut-off values for both subjective memory complaints and depressive symptomatology (MAC-Q < 25, GDS < 12) were classified as Controls. Note however that these are not Controls as the term is usually employed, i.e. they are not subjects who are "normal" as far as operational definitions, nor were they matched

to study subjects by demographic characteristics. Rather, this sample was constituted of people who answered an ad asking for volunteers to participate in a research project relating to stress and aging, and who did not have significant levels of either subjective memory complaints or depressive symptoms.

The total recruited sample (including subjects not studied in this project) can be divided into six groups for the sake of clarity: The first three groups, MEM, MemDep and Control, are explained above. The other groups were comprised of the 17 subjects who were excluded from this study. These subjects were excluded for any of the following reasons: they had significant depressive symptoms only and no subjective memory complaints (Dep group); they were taking anti-depressant medicines (Anti-Dep group); or they were taking anxiolytics more than three times a week (Anxio group).

Figure 5

Categories of Evaluated Volunteers



Only after all subjects had been evaluated was it decided to exclude these 17 subjects in order to eliminate factors which would make inter-group comparisons problematic. Thus, all the subjects categorized as Dep, Anti-Dep, or Anxyo were excluded from the present project. (All 69 evaluated subjects will be followed up as part of a longitudinal study conducted by Dr. Sonia Lupien.) These 17 excluded subjects comprised 5 cases of depressive symptoms without subjective memory complaints; 7 cases of current anti-depressive medicine usage; and 5 cases of more than occasional anxiolytic usage (> 3 times a week).The total evaluated sample of 69 evaluated subjects were thus categorized as illustrated in **Figure 5.** The last three groups (Anti-Dep, Anxio and Dep) fall outside the purview of this project. The included and excluded groups did

not differ in age (means ± SD for included and excluded groups: 68.1 ± 6.9; 65.5 ± 6.1), education (14.0 ± 3.5; 14.2 ± 3.4) or MMSE score (28.1 ± 2.3; 28.0 ± 2.1). This study therefore focused on a subset of the 52 subjects, comprising 62.3% of the total evaluated sample.

8.1.1 Research Sample

The present study sample consisted of seniors with memory impairments with or without depressive symptoms. A control group of volunteers who had scored below cut-off on questionnaires of memory complaints or depressive symptoms was also included. The three groups are comparable in age, years of education, and cognitive abilities as indexed by MMSE or 3MS scores **Table 1**.

Table 1
Cognitive Screening Tests
Means and 95% Confidence Intervals

Group	Age	Education	MMSE	3MS
Mem N=25	68.7 (65.7-71.6)	13.8 (12.5-15.3)	27.8 (27.0-29.5)	92.7 (90.1-95.4)
MemDep **N=18**	68.3 (64.8-71.8)	14.0 (12.0-16.0)	28.4 (27.3-29.5)	94.8 (92.7-96.9)
Controls N=9	66 (61.5-70)	14.1 (11.8-16.5)	28.1 (26.1-30.2)	93.9 (89.2-98.5)

8.2 Standard Screening

The Mini-Mental State Exam (MMSE; Folstein et al, 1975) and the Modified Mini-Mental State Exam (3MS; Teng and Chui, 1987) are easy to administer and provide an aperçu of a subject's cognitive capacities; however, they are not sensitive to mild

cognitive impairment (Tombaugh and McIntyre, 1992). The present study confirms this. **Figure 6** shows the individual scores by group and by test for the MMSE. The cut-point for cognitive impairment is below 24 on the MMSE (Tombaugh and McIntyre, 1992); as shown in figure 7, this describes only 4 cases from the study sample. Of course, the 23/24 cut-point is meant to flag possible cases of AD, not mild cognitive impairment. Similarly, using the proposed cut-point of 80 on the 3MS to detect mild cognitive impairment (Kristjansson et al, 1996) captured only 2 cases from the total evaluated sample (**Figure 7**). It could be argued that cut-points should be adjusted to individual combinations of educational level and age range; but as discussed above, such a procedure remains controversial with little empirical data to support it (Kraemer et al, 1998). In addition, there are no generally recognized MMSE or 3MS cut-points for mild cognitive impairment.

Figure 6
MMSE Scores

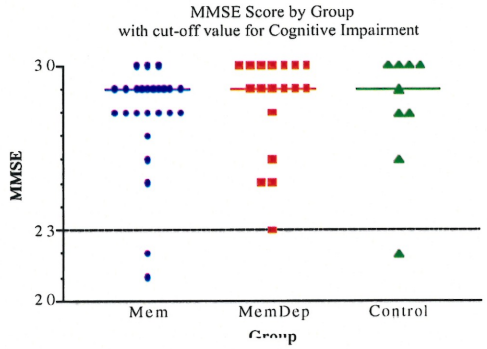
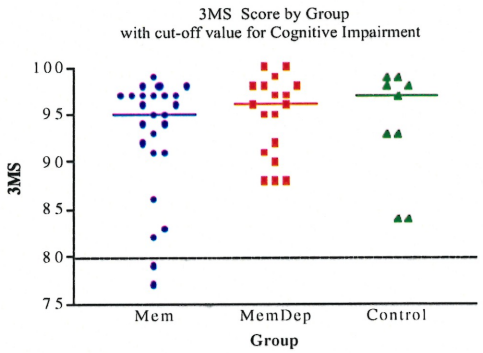


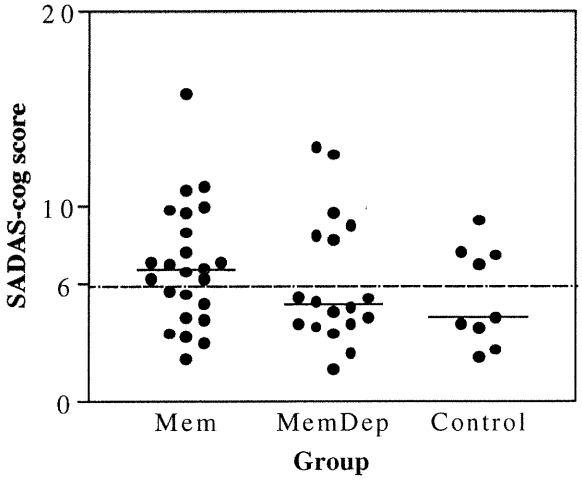
Figure 7
3MS Scores



The 3MS (Teng and Chui, 1987) was designed to provide finer differentiation between cognitive capacity levels than the MMSE; but this did not turn out to be the case with this study sample. One way to quantify differentiation is the ability of a test to confer different scores on different subjects. The standard deviation is not appropriate as a measure of comparative spread of scores since it depends on the units used by each test. Thus, a comparison was made using each test's coefficient of variation (CV), calculated as the test's mean divided by its standard deviation. The means and standard deviations of the CVs for each of the three groups within the two testing instruments was .08 ± .01 for the MMSE and .06 ± .01 for the 3MS (Wilcoxon signed rank test of the difference: p = 0.25).

Figure 8

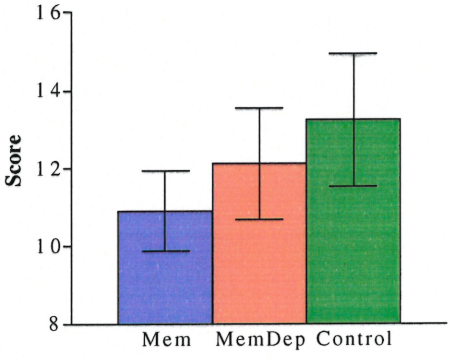
SADAS-cog Scores
with median bars and cut-off value for possible cognitive impairment
(Higher score = greater impairment)



The results from the SADAS-cog (Standish et al., 1996; Rosen et al, 1984) are different (**Figure 8**). First, the dispersion of scores obtained is greater: a Friedman's test of the Coefficients of Variation associated with each instrument for each group were significantly different [Friedman's exact p = 0.028]. Dunn's Multiple Comparison Test showed that the CV of the SADAS-cog was significantly greater than that of the 3MS ($p < 0.05$). Thus,

according to these results, the 3MS and the MMSE do not differ in their dispersion of test scores between subjects, while the SADAS-cog does differentiate more than the 3MS. In addition, the SADAS-cog places 24 of 52 subjects, or 46% of the sample, in the category of suspected mild cognitive impairment. There was no significant difference in SADAS-cog score by group [F (2,48) = 1.12, p = .34.] There was little redundancy between tests. Seventeen of 25 Mem subjects (68%) were classified as "suspected mild cognitive impairment" by any one of the three cognitive screening instruments, similarly classified were 7 of 18 MemDep subjects (39%) and 4 of 9 Control subjects. However, only 4 subjects (2 from the Mem and 1 each from the MemDep and Control groups) were classified as such by any two instruments, and only 1 subject by all three screening instruments. Since the 3 tests were correlated (p = .001 for all bivariate correlation), a Multivariate ANOVA was performed on all three parameters; there were no group differences [Wilkes' Lambda = .92, F (6,88) = .62, p = .71].

Figure 9
Comprehension Scores from Wechsler Adult Intelligence Scale-R means with SE



Comprehension and Abstract Reasoning

A one-way ANOVA was used to analyze the Judgment and Abstract Reasoning data from the Wechsler Adult Intelligence subscales. There was a trend for a group effect on Comprehension [F (2,51) = 2.86, p

= .067], with the Mem, MemDep and Control groups in rank order from worst to best (**Figure 9**); and no significant group effect on the Similarities test, which measures abstract reasoning [F (2,51) = .03, p = .98].

8.3 Experimental

8.3.1 Declarative Memory

Experimental Declarative Memory Test

Immediate and delayed recall in declarative memory, and non-declarative (implicit) memory were assessed by means of the Experimental Declarative Memory Test (EDMT). First, to test immediate recall, a three-way ANOVA with two repeated measures was performed. *Relatedness*, i.e. related and unrelated word pairs was one repeated measure, and *presentation* i.e. first or second presentation of the word list (at the second presentation the order of the pairs was reversed) was the other; *Group* was the between groups factor. There were no significant group differences in the number of words recalled [F (2,49) = .13; p = .882] (**Figure 10**). Then, because they are conceptually different, separate, two-way ANOVAs with one

Figure 10
Correct Recall by Relatedness and Presentation
Group Means with SE

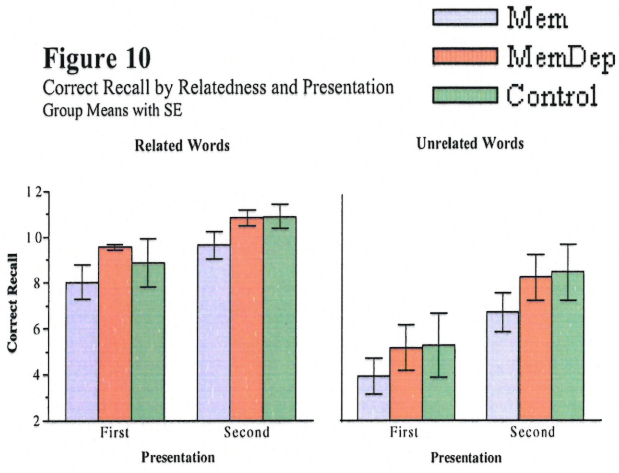
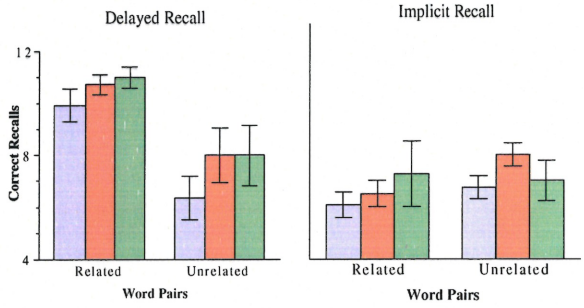


Figure 11
Delayed and Implicit Recall by Relatedness
Means with SE

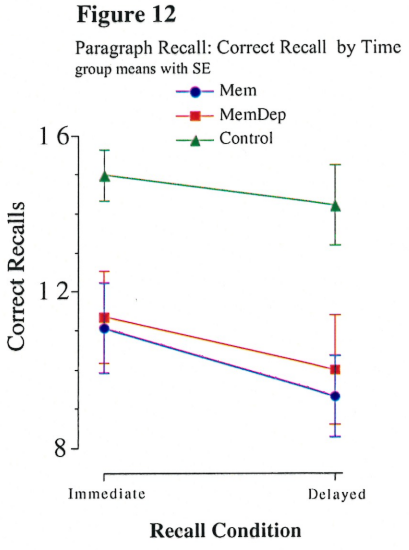


repeated measure were performed on delayed and implicit recall, with relatedness as the repeated-measures factor and Group as the independent variable. There were no significant group differences on either measure [F(Delayed Recall) (2,49) = 0.36, p = .70 and F(Implicit Recall) (2,49) = 1.44, p = .25] (**Figure 11**).

Paragraph Recall with Valence Factor

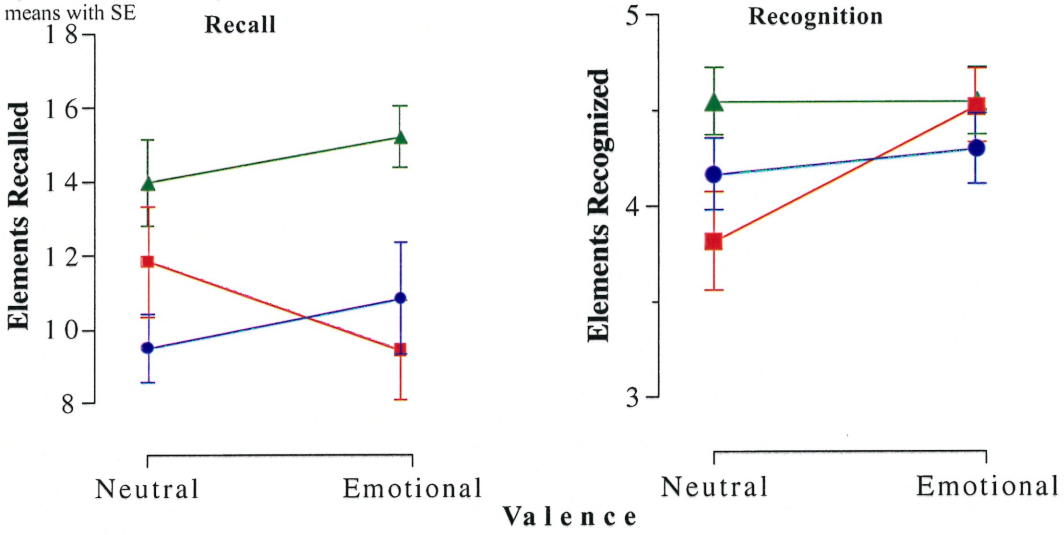
A two-way ANOVA with one repeated-measures factor was performed on the total story elements recalled by group at immediate and delayed recall. Story elements recalled was the repeated-measures factor and group was the independent variable. There was no significant group effect on total story elements recalled [F(Total Recall) (2,47) = 2.85, p = .068] and no significant interaction of group by time [F(Group by Time) (2,47) = .45, p = .64]: thus, the three groups of subjects forgot at roughly the same rate from the immediate to the delayed recall trials (**Figure 12**). In order to test for the effect of the valence on the recall of story elements in the different groups, a two-way ANOVA was then performed, with Valence (Emotional and Neutral story elements) as the repeated-measures factor and Groups as the independent variable. There was no significant effect of group [F(Group) (2,47) = 3.1, p = .057] or Valence [F(Valence) (1, 47) = .01, p = .96], but there was a

Figure 12



significant group-by-valence interaction [F(group-by-valence) (2,47) = 3.78, p = .03] (**Figure 13, Recall**). A contrast analysis of recall by valence was then carried out. The contrast of {Mem and Control} vs. {MemDep} was significant [F(Contrast) (1,47) = 6.9, p = .01]. Since Hypothesis H1 was about the differential effect of valence on

Figure 13
Recall and Recognition
by Valence and Group
means with SE



recall within each group, post-hoc tests were performed, first using the Spjotvoll-Stoline test. There were no significant effects of valence on recall within any group. Because this left the significant contrast difficult to interpret, the less conservative Least Significant Difference t-test was then performed on neutral and negative valence words recalled within each group; the results were significant only in the MemDep group [p = .03]. Thus, the significant contrast was due to the fact that valence had no significant effect on recall in the Mem or Control groups, while negative valence

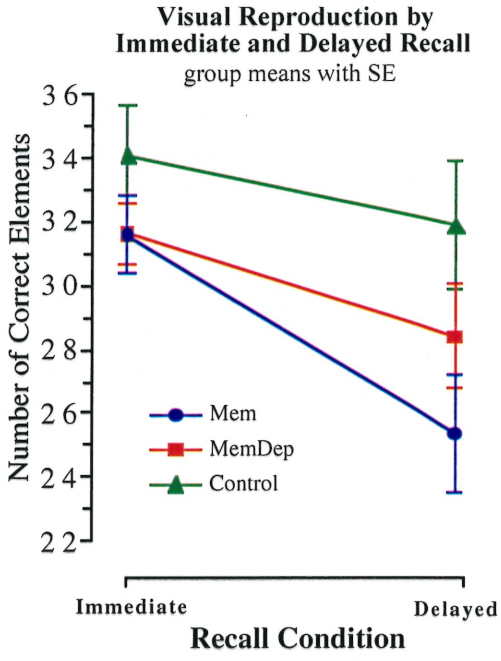
decreased recall in the MemDep group.

Recognition was also analyzed with a two-way ANOVA with valence as the repeated-measures factor and Group as the independent variable. There was a significant effect of Valence [F(Valence) (1,46) = 4.42, p = .04], no effect of group [F(Group) (2,46) = .83, p = .44], and no Group by Valence interaction [F(Group-by-Valence) (2,46) = 2.9, p = .06] (**Figure 13, Recognition**).

Visual Reproduction Test

A two-way ANOVA with Trial (Immediate and Delayed Recall) as the repeated-measures factor and Group as the independent variable showed a significant interaction of Trial by Group [F(Interaction) (2,49) = 3.36, p = .043]. A Spjotvoll-Stoline post-hoc test showed that the Mem group recalled fewer visual elements at Delayed than Immediate Recall (p < .001); there were no significant differences between Immediate and Delayed recall for the MemDep (p = .06) or Control (p =.72) groups (**Figure 14**).

Figure 14



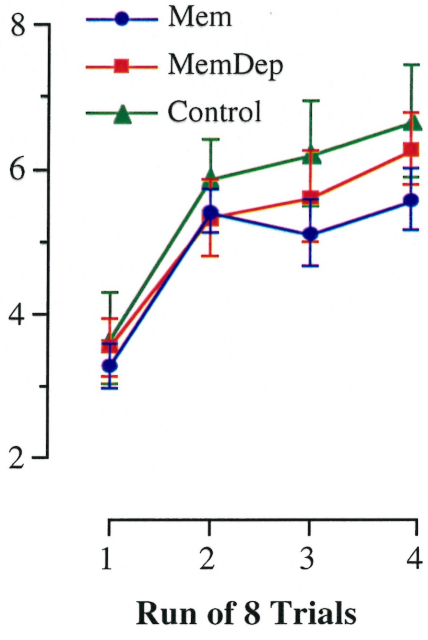
8.3.2 Attention, Psychomotor and Verbal Fluency

Conditioned Associative Learning Test (CAL)

For the CAL the variables of interest were: (1) the number of correct first response in the first, structured part of the test, separated by runs of 8 trials; (2) the number of prior response repetitions and (3) the number of trials necessary to reach criterion (up to a maximum of 64). Since the first variable were not significantly correlated with the other two, it (i.e., the number of correct first responses in the first part, by runs of 8 trials) was analyzed separately. A two-way ANOVA with Trial as the repeated measures factor and Group as the independent variable was performed. There were no significant group effects [F(Trial) (2,48) = .82, p = .45] (**Figure 15**). Most but not all subjects improved over successive trials, making more correct first responses in trials 3 and 4 than in trial 1 and 2. However, in 7 of 24 Mem subjects (29 %), as well as in 17 MemDep subjects (18 %) and 1 of 9 Control subjects, the performance was actually worse over the second half of the test.

Figure 15

Correct First Responses by Run of 8 Trials
Means with SE



Because this test's other components (Prior Response Errors [Errors], and Trials to Criterion [Trials]) were correlated (Pearson's r^2 = .295, p < .001), a Multivariate Analysis of Variance was performed, with Errors and Trials as the dependent variables and Group as the Independent factor. There were no significant group differences (Wilks's Lambda = .94, p = .57). **Figure 16** shows the individual prior response errors. There were also no group differences in the proportion of subjects in each group who reached the criterion (i.e., 8 consecutive correct first responses within the first 32 unstructured trials) [χ^2 Criterion (df = 2, N = 52) = 3.18, p = .21]. Criterion was achieved by 7 of 24 Mem subjects (29 %), 10 of 18 MemDep subjects (55 %) and 5 of 9 Control subjects (55 %) (**Figure 17**).

Figure 16

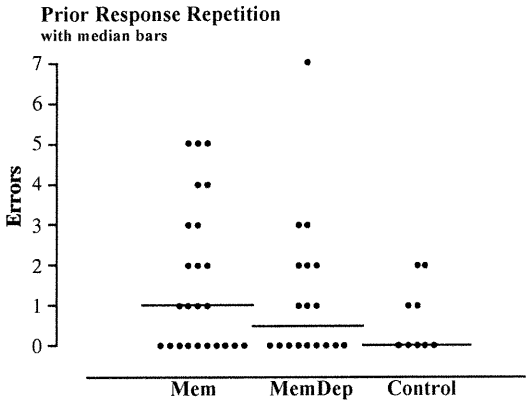
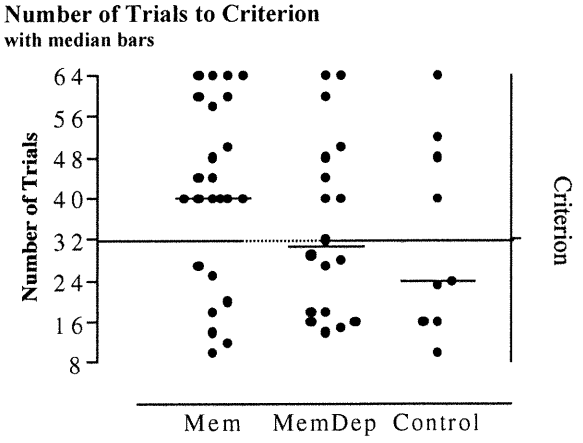


Figure 17



Selective Attention

In this test subjects scanned a computer screen and had to signal as quickly as possible whether or not a target (a black square) was present by pressing a Yes or No key. Each trial featured 2, 4 or 7 distractor objects (black or white circles and white

squares). Forty randomly distributed trials were in the Target Present condition and 40 were in the Target Absent condition. Reaction time performances were evaluated with a three-way ANOVA with *Condition* (target present or absent) and *Number* (of on-screen distractor objects: 2, 4, 7 or 10) as the repeated-measures factors and Group as the independent variable. Although there were significant effects of Condition [F(Condition) (1, 48) = 98, $p < .001$] and Number [F(Number) (6,144) = 25.6, $p < .001$], there were no significant Group effects on any measure [F(Group) (2,48) = .76, $p = .47$; F(Group by Condition) (2,48) = .17 $p = .85$; F(Group by Number) (6, 144) = .35, $p = .91$; F(Group by Condition by Number) (6, 144) = .55, $p = .77$] **(Figure 18)**.

Figure 18

Selective Attention: Reaction Times by presence or absence of target and number of distractors

Group means with SE

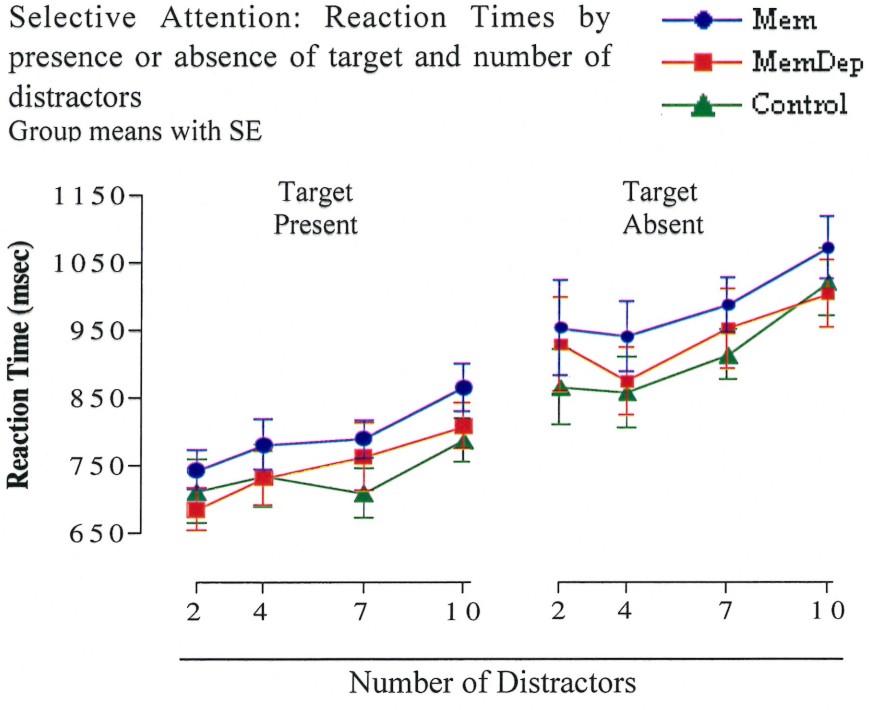


Table 2

False Negatives and Positives in the Selective Attention Test

False Negatives	Mean Errors	SE	Range
Mem	.76	.18	0-3
MemDep	1.17	.35	0-4
Control	.22	.15	0-1
False Positives			
Mem	1.04	.63	0-16
MemDep	1.17	.31	0-4
Control	.22	.22	0-2

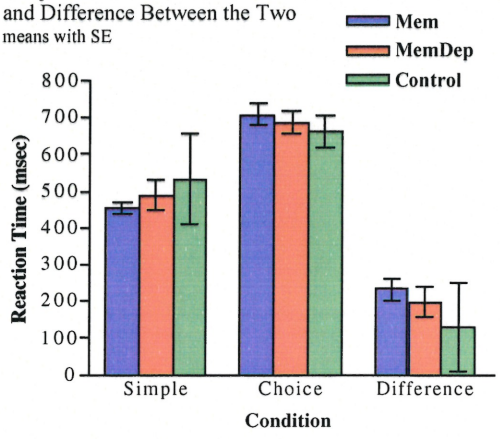
Table 2 shows the range of mean number of false positive and false negative errors (i.e., signaling that the target was present when it was not, and vice-versa). In order to make the results more relevant, the false negative (from the Target Present trials) and false positive (from the Target Absent trials) data were recoded as "accurate" for each subject who had a perfect performance or only a single error; and as "inaccurate" for each subject who had two or more errors. This is based on the following rationale: Subjects were instructed to respond by signaling whether the target was present, within two very explicit conditions: they had to answer as accurately as possible *and* as quickly as possible. A single error might have been an accident; more than one suggests a pattern of behaviour. Thus if one gives subjects who had a single error the

benefit of the doubt and classify them with the "perfect responders," one is left with two contrasting groups: those who followed directions carefully, and those who only followed one of the directions: speed. A Chi Square test was performed on the two types of errors (false negative or false positive) separately. There were no differences in the frequency of false positive errors [χ^2 (False Positives)= .84, df = 2, p = .66], or false negatives [χ^2 (False Negative) = 4.6, df = 2, p = .09]. There were no significant differences in mean reaction time between accurate and inaccurate responders either in the Target Absent [Mann-Whitney U test, z = -.61, two-tailed p corrected for ties = .55], or Target Present trials [Mann-Whitney U, z = -.69, two-tailed p corrected for ties = .49].

Reaction Time

Simple and Choice Reaction Times were not significantly correlated [df = 50, Pearson's r = .22, p = .127], and were evaluated with one-way ANOVAs. There were no significant group differences [F(Simple Reaction Time) (2,49) = .54, p = .59; and F (Choice Reaction Time) (2,49) = .40, p = .67]; neither were there any significant differences between the median Choice and median Simple reactions times [F (Difference) (2,49) = .82, p = .45] (**Figure 19**).

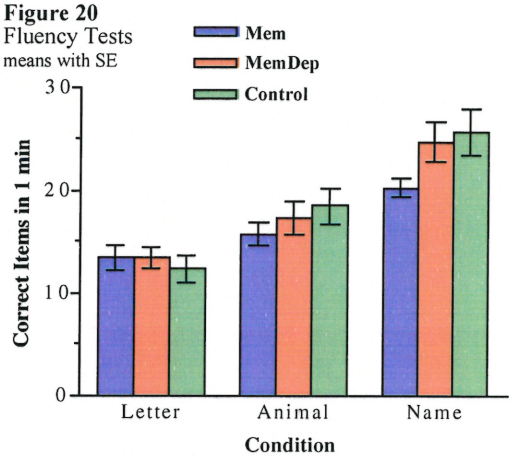
Figure 19
Simple and Choice and Reaction Time and Difference Between the Two means with SE



Verbal Fluency

One-way ANOVAs were performed with Group as the independent variable and Animal, Letter or Name Fluency as the dependent variable. There were no Group effects for Animal [F(Animal) (2, 29) = .82, p = .45] or Letter Fluency [F(Letter) (2,49) = .18, p = .84], but there was a significant group difference in Name Fluency [F(name) (2, 48) = 3.53, p = .04] (**Figure 20**). There were no significant pairwise differences.

Figure 20
Fluency Tests
means with SE



8.3.3 Neuroendocrine

Cortisol concentrations were determined via salivary cortisol assays from 5-times daily samples taken for three consecutive days. The frequency distribution of the grand mean for all subjects is shown in **Figure 21**, and the scatterplot of individual means within each group is

Figure 21
Three-day mean cortisol concentrations

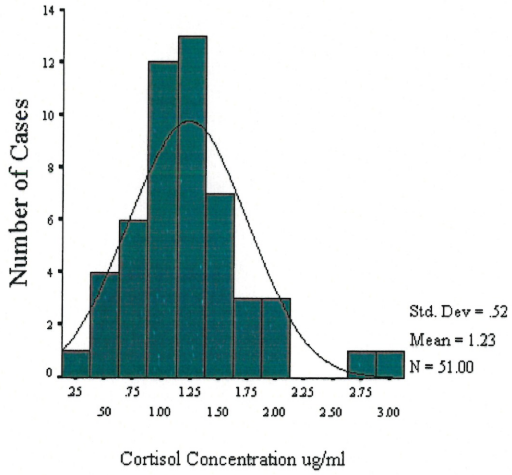


Figure 22
Individual 3-Day Mean Cortisol Concentrations
by Group
with median bars

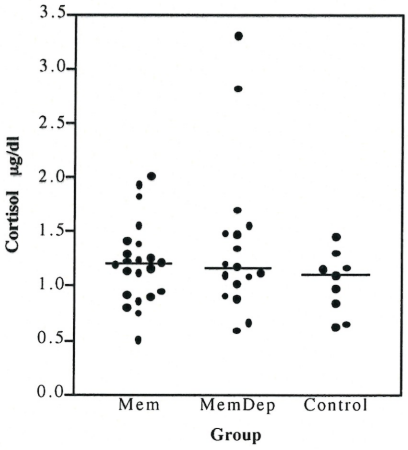
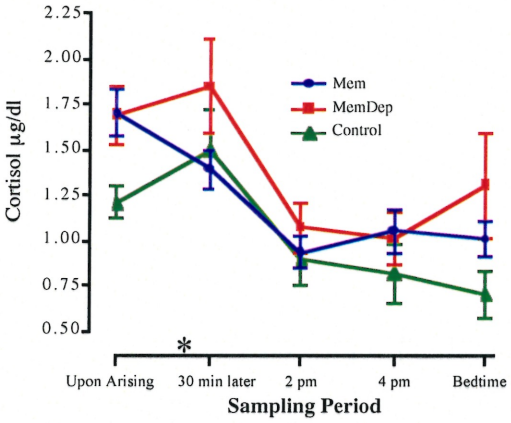
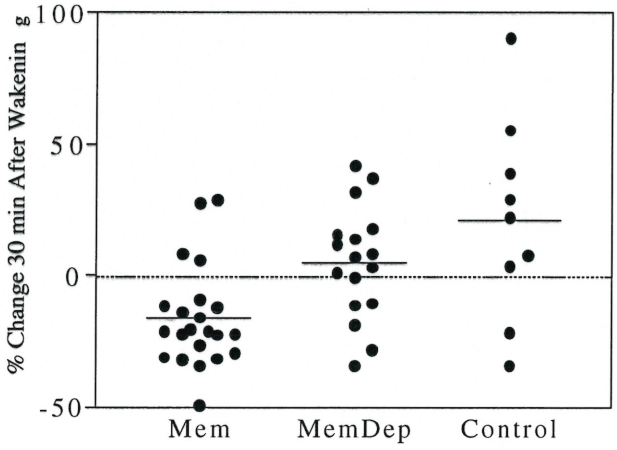


Figure 23
Cortisol Concentrations by Sampling Period
group means with SE



depicted in **Figure 22.** There were no group differences in 3-day mean cortisol concentrations [F (2,45) = p .4325]. Intergroup comparisons were also made with a two-way repeated-measures ANOVA on the infradian data, using the five sampling times averaged over three days as the repeated-measures factor and Group as the independent va was a significant interaction of time by group [F (8,180) = 2.39, p = .018]. Spjotvol Stoline post-hoc tests showed a difference between Control and MemDep, and Mem and MemDep, 30 min after awakening (p = .03) (**Figure 23**). The MemDep group appears to show a rise at Bedtime compared to the Control group (Figure 23), but it was not significant (Spjotvoll-Stoline test p = .054).

Figure 24
Scatterplot of Morning Cortisol Rise
by Group with means bars



The morning cortisol change following awakening was then analysed on its own by

submitting the percent changes in salivary cortisol 30 min after awakening (i.e., the percent difference from the first to the second saliva sample of the day) to a one-way ANOVA. The results were significant [F(Morning Cortisol) (2, 45) = 8.5, $p < .0001$] (**Figure 24**). A Spjotvoll-Stoline test showed that that the Mem Group had the most extreme mean value—a mean decrease in cortisol instead of the expected increase 30 min after awakening—and that this mean was significantly different from both the MemDep group mean ($p = .04$) and the Control group mean ($p = .01$); these latter two means did not differ significantly ($p = .35$). Thus, the MemDep group had the highest absolute cortisol concentrations 30 min after awakening, but the Mem group had a cortisol decrease in that same 30 min period that was significantly different from the increase found in the other two groups.

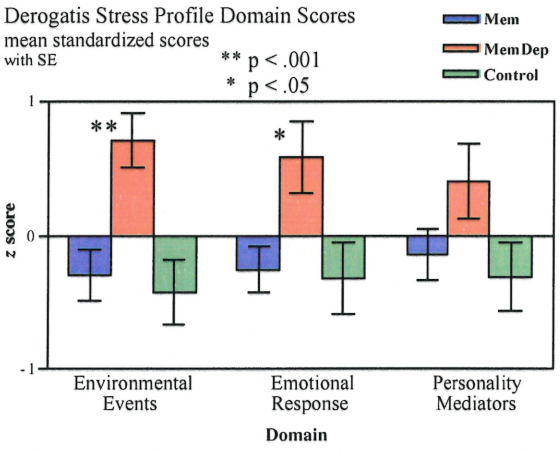
8.3.4 Psychosocial

Derogatis Stress Profile

In order to permit comparisons across stress measures, standardized scores were used for analyses and graphs. The untransformed scores for the Derogatis Stress Profile and the Symptom Checklist-90 (which is treated in the next section) are given in **Tables 3 and 4**. The Derogatis Stress Profile Domains—Environmental Events, Emotional Response, and Personality Mediators—were first analyzed in a Multivariate ANOVA. There was a significant group effect [Wilks's Lambda (df: 6, 90) = .71, $F = 2.79$, $p = .016$]. Subsequent univariate F tests revealed significant group differences in the

Environmental Events [F (2,47) = 7.65, p = .001] and Emotional Response [F (2,47) = 4.55, p = .02] Domains, but not in the Personality Mediator Domain [F (2,47) = 2.07, p = .14] (**Figure 25**).

Figure 25



Next, the eleven Derogatis Stress Profile Dimensions—Time Pressure, Driven Behaviour, Attitude Posture, Relaxation Potential, Role Definition, Vocational Satisfaction, Domestic Satisfaction, Health Posture, Hostility, Anxiety and Depression—along with the Derogatis Subjective Stress rating, were also analyzed inthe form of standardized scores; a Multivariate ANOVA revealed a significant group effect [Wilks's Lambda = .38, F (24, 72) = 1.84, p = .025]. Since higher stress in the MemDep group was hypothesized *a priori* (Hypothesis H2), contrast analysis was then carried out on the eleven Derogatis Stress Profile Dimensions plus the subjective stress rating; the results are depicted in **Figure 26** in increasing order of the associated F statistic. Thus, the MemDep group had higher stress scores on 10 of 12 subscales, seven of which were statistically significant in contrast tests.

Table 3
Derogatis Stress Profile

Derogatis Stress Profile

Subscale	Group (means ± standard deviation)		
	Mem n = 25	Mem Dep n=16	Control n = 8
Emotional Response Domain	30.32 ± 11.11	40.82 ± 13.37	29.44 ± 10.31
Environmental Events Domain	30.84 ± 8.84	40.00 ± 7.32	29.67 ± 6.60
Personality Mediators Domain	56.40 ± 9.12	61.69 ± 10.86	54.78 ± 7.45
Depression	8.36 ± 4.27	14.13 ± 6.02	6.67 ± 4.85
Domestic Satisfaction	10.44 ± 3.78	14.25 ± 4.64	11.22 ± 5.19
Driven Behavior	10.20 ± 4.06	11.06 ± 3.87	8.78 ± 2.91
Attitude Posture	8.48 ± 2.04	9.82 ± 3.51	8.67 ± 1.87
Anxiety	12.20 ± 4.19	15.63 ± 4.88	11.33 ± 2.87
Health Posture	8.28 ± 4.31	10.63 ± 3.24	7.00 ± 2.29
Hostility	9.76 ± 5.13	11.06 ± 5.64	11.44 ± 4.42
Time Pressure	11.68 ± 3.84	12.13 ± 3.05	12.89 ± 4.29
Relaxation Potential	11.36 ± 1.91	12.5 ± 2.73	12.22 ± 2.17
Role Definition	14.68 ± 3.58	16.19 ± 3.51	12.22 ± 2.77
Subjective Stress Rating	45.40 ± 2.73	66.25 ± 17.46	53.33 ± 18.71
Vocational Satisfaction	12.12 ± 3.06	15.13 ± 3.65	11.44 ± 3.40

Table 4

Psychiatric Symptom Checklist-90

Symptom Checklist-90			
Subscale	Group (means ± standard deviation)		
	Mem n = 24	Mem Dep n=16	Control n = 8
Anger	.27 ± .65	.59 ± .65	.40 ± .44
Anxiety	.40 ± .49	.93 ± .52	.35 ± .42
Depression	.65 ± .55	1.62 ± .92	.31 ± .21
Obsessive Compulsive	.82 ± .65	1.54 ± .87	.65 ± .19
Paranoid Ideation	.57 ± .75	.90 ± .60	.38 ± .60
Phobic Anxiety	.24 ± .39	.34 ± .33	.27 ± .40
Psychoticism	.37 ± .47	.76 ± .62	.15 ± 15
Interpersonal Sensitivity	.62 ± .52	1.23 ± .71	.51 ± .53
Somatization	.45 ± .56	1.12 ± .73	.32 ± .16

Figure 26
Derogatis Stress Profile Dimension Scores
mean standardized scores
with SE

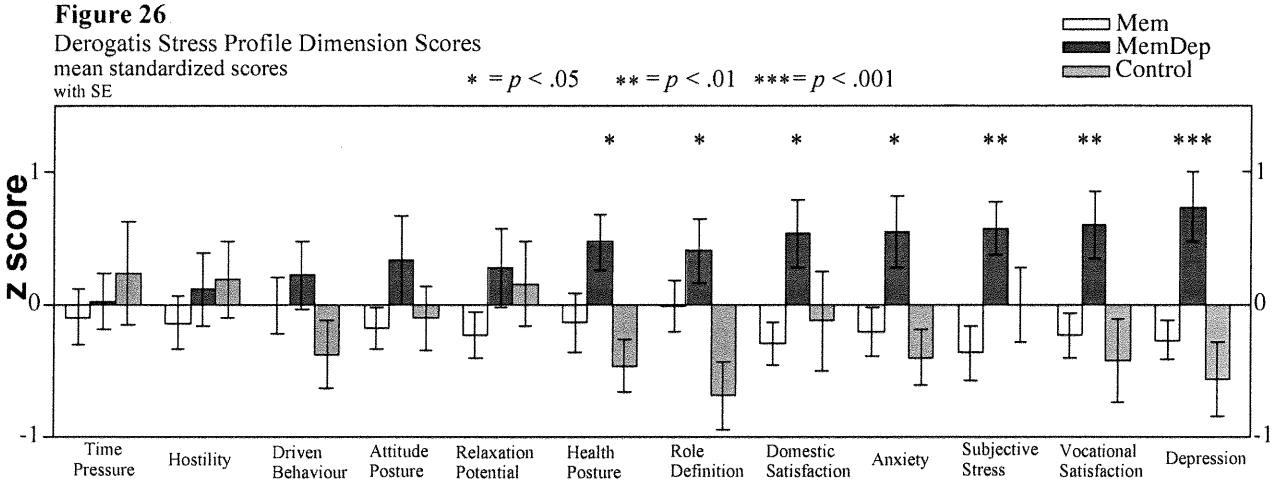
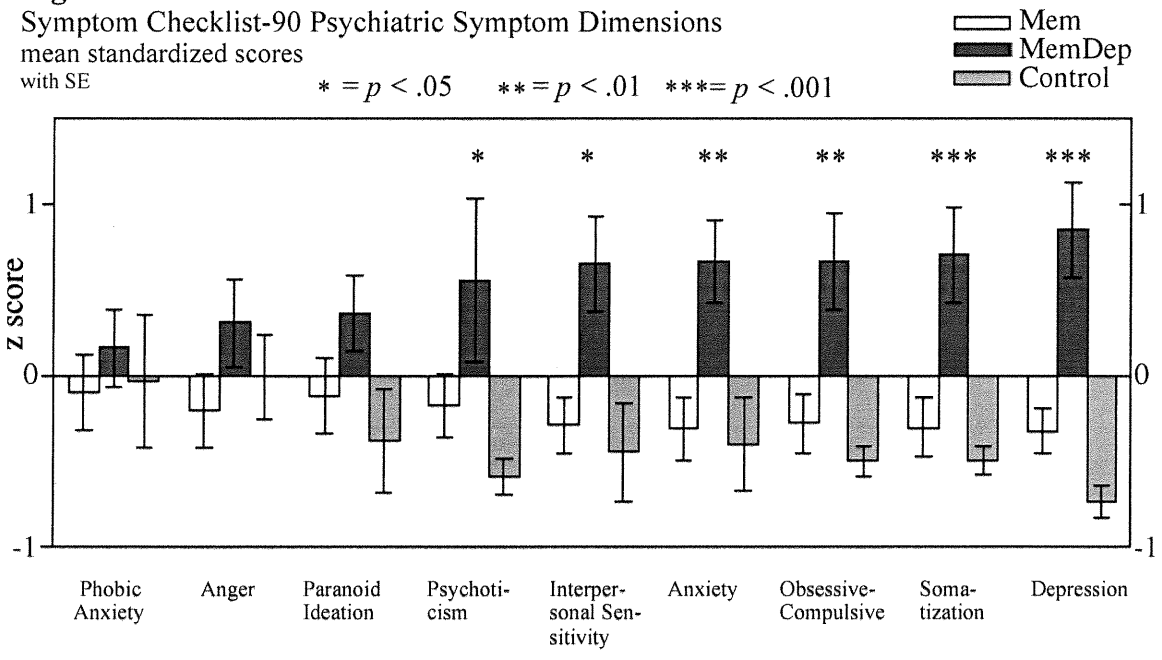


Figure 27
Symptom Checklist-90 Psychiatric Symptom Dimensions
mean standardized scores
with SE



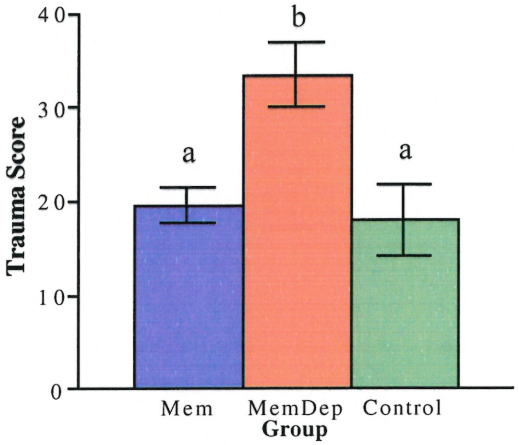
Symptom Checklist-90

An identical analysis as the above was carried out for the nine subtests of the Symptom Checklist-90. A Multivariate ANOVA was carried out on the subtests. There was a significant effect of group [Wilks's Lambda = .45, F (18,74) = 1.98, p = .022]. Contrast analysis The MemDep group had significantly higher scores than either of the other two groups on all nine subscales (6 of which were significant at $p <$ .05 in univariate F tests). The mean standardized scores of the three groups on the nine subscales of the SCL-90, along with the significance levels of the associated univariate tests, are depicted in **Figure 27**. (The untransformed scores are given in **Table 4**.)

Trauma Symptom Checklist (TSC)

The TSC data were analyzed with a one-way ANOVA; there was a significant group difference [F(Trauma) (2,46) = 8.7, p = .0006], with a post-hoc Spjotvoll-Stoline test showing that the MemDep group has significantly higher Trauma scores than the Mem Group (p = .003) and the Control group (p = .02). There was no difference between the Mem and Control

Figure 28
Trauma Symptom Checklist-33
Means with SE



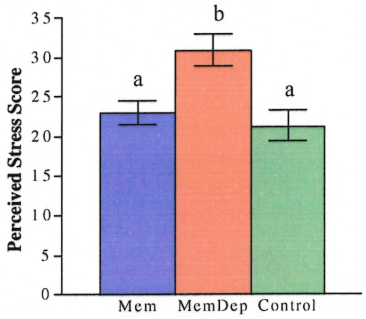
Group means not sharing a superscript letter are significantly different $p < .05$

groups (p = .95) (**Figure 28**).

Perceived Stress Scale (PSS)

The PSS data were also analyzed with a one-way ANOVA; there was also a significant group difference [F(Perceived Stress) (2,47) = 7.5, p = .0015]; a contrast analysis showed that MemDep group has higher perceived stress than the two other groups [F(Contrast) (1,46) = 14.8, $p < .001$] (**Figure 29**).

Figure 29
Perceived Stress Scale
Group means with SE



Group means not sharing a supescript letter are significantly different $p < .05$

Stress Scales and Trauma Symptom Checklist

The Trauma Symptom Checklist is designed to measure present day symptoms related to psychologically disrupting events that happened in the past, whereas the Derogatis Stress Profile, the Symptom Checklist-90, and the Perceived Stress Scales have no such theoretical and empirical linkage to the past.

During the evaluation of subjects with the Alzheimer Disease Assessment Scale—cognitive subscale (ADAS-cog), subjects were induced to produce 3 minutes of oral language; in order to evaluate spoken language ability, the ADAS-cog *Manual* instructs the tester to prompt the subject with 3 standard questions, which are used as needed only in order to have the subject speak for 3 minutes. The first prompt was "Tell me about your family" along with "Tell me more about your family" in the event

that a 15-second lull ensued before a minute had passed. The second prompt was "Tell me about where you born and raised." (The third prompt, "Tell me about the jobs you had or the work you did," was rarely needed.) I was moved and surprised at the tragic stories that came out of many of the volunteers at these prompts. Though anecdotal in the purest sense, the stories that were heard from these volunteers could not but bring to mind the powerful influence that long ago painful experiences can have in the present on one who has not yet come to terms with them. Thus, when one considers the strong group effect on the Trauma Symptom Scale (TSS), with the MemDep group scoring significantly higher than the other two groups, it seems appropriate to check whether sequellae of traumatic events, as indexed by the TSS, could account for group differences found in stress levels as indexed by the Derogatis Stress Profile scores, as well as in psychiatric symptoms as indexed by the Symptom Checklist-90. In order to test this hypothesis—that much or all of the variance in stress and psychiatric symptomatology values can be explained by differences in traumatic symptoms—three Multivariate Analyses of Variance were carried out using the TSC as covariate. First, the MANOVA on the Derogatis Stress Profile subscales was repeated; with the TSC as covariate, the resulting MANCOVA was no longer significant [Wilks's Lambda = .49, F (24, 68) = 1.20, p = .272]. The significance levels of the TSC covariate by Derogatis Stress Profile (DSP) Dimension are listed in **Table 5.**

Table 5

Significance levels by Derogatis Stress Profile Dimension of the Trauma Symptom Checklist Covariate

DSP Dimension	t-value	*p*
Anxiety	5.82	<.001
Attitude Posture	3.86	<.001
Depression	5.09	<.001
Dominance Posture	2.9	.01
Driven Behaviour	1.64	.11
Health Posture	0.49	.63
Hostility	2.75	.01
Relaxation Potential	0.77	.44
Role Definition	4.97	<.001
Subjective Stress Rating	3.34	.01
Time Pressure	1.96	.06
Vocational Satisfaction	2.54	.02

Second, the same MANCOVA was carried out for the higher-level DSP Domains: Environmental Events, Emotional Response, and Personality Mediators. Just as with the DSP Dimensions, the addition of the Trauma Symptom Checklist (TSC) as covariate made the group means no longer significantly different from each other [Wilks's Lambda = .88, F (6,86) = .97, p = .45]. The significance levels of the TSC covariate for the three Derogatis Stress Profile Domains are listed in **Table 6**.

Table 6
Significance Level of the Trauma Symptom Covariate by Derogatis Stress Profile Domain

DSP Domain	t-value	*p*
Environmental Events	2.94	.005
Emotional Response	5.92	<.001
Personality Mediators	4.92	<.001

The same procedure was followed for the group differences previously found in psychiatric symptoms as indexed by the Symptom Checklist-90. Again, the previously significant group effect was rendered non-significant [Wilks's Lambda = .67, F (18,72) = .89, p =.59]. Results of the regression analysis of the TSC covariate for the nine Symptom Checklist-90 subscales are given in **Table 7.** Finally, an ANOVA on the Perceived Stress Scale was repeated, using the TSC as covariate. The effect of group was no longer significant [F (2, 45) = 1.35, p = .27; Trauma covariate: F (1,2) = 33.7, p <.001]. Thus, when the Trauma Symptom Checklist is entered as a covariate there are no significant group differences in psychiatric symptoms or self-rated stressful events and feelings.

Table 7
Significance level of the Trauma Symptom Covariate by Symptom Cheklist-90 Subscale

SCL-90 Subscales	t-value	p
Somatization	4.19	<.001
Obsessive-Compulsive	6.53	<.001
Interpersonal Sensitivity	5.22	<.001
Depression	7.14	<.001
Anxiety	3.98	<.001
Anger	4.23	<.001
Phobic Anxiety	1.57	.122
Paranoid Ideation	4.64	<.001
Psychoticism	4.72	<.001

Social Support, Self-Efficacy, Daily Activities

The Social Support data were analyzed with a one-way ANOVA. There were no significant group differences [F(Social Support) (2, 41) = 1.53, p = .23]. The Mem group reported (mean ± standard error) 14.6 ± 2.99 supportive persons in their social circles, the MemDep group 18.4 ± 7.69 and the Control group 33.3 ± 13.3.

There was no significant group difference in Self-Efficacy scores [F (2, 48) = 2.66, p = .08] [means and SD for Mem, MemDep and Control groups respectively: 33.2 ± 3.8; 29.0 ± 7.1; and 31.1 ± 6.6].

The 22 subscales of the Daily Activities Data were analyzed with a Multivariate ANOVA. There were no significant group differences [Wilks's Lambda = .21, F =

1.33, $p = .17$].

8.4 Cognition and Cortisol

Possible relationships between cortisol concentrations and cognition were tested by means of bivariate regression analyses between several cortisol and cognition measures. For cortisol, the measures analyzed were: the means of the 5 daily sampling times (upon arising, 30 min later, 2 PM, 4 PM, bedtime), the grand mean (of all time points and days), the maximum and minimum mean cortisol values, and the percent change in cortisol concentration from Upon Rising to 30 min later. For cognition, two composite measures were devised with the intention of loading them both with demanding cognitive tasks, in order to provide more of an opportunity for any cognitive effects of cortisol to manifest themselves (a small disturbance in cognitive functioning is more liable to be detected in demanding than in easy cognitive tasks). The composite measures were (1) a Paragraph Recall measure and (2) an Attention measure. The Paragraph measure was comprised of the total items recalled in the paragraph recall task (immediate and delayed recall combined) (maximum: 52). For the Attention measure the values for Choice Reaction Time, Absent Target reaction time in the Selective Attention test, and number of trials to Criterion in the Conditioned Associated Learning test, were transformed to standardized scores and then summed. These particular measures were chosen because (1) paragraph recall as indexing hippocampal-mediated, declarative memory function is a well-recognized

outcome measure in studies of symptomatic change in AD; and (2) measures of attention or concentration are also known to be important, although somewhat neglected, in evaluating the impact of AD on cognition (Mohs, Knopman, Petersen, Ferris, Ernesto, Grundman, Sano, Bieliauskas, Geldmacher, Clark & Thal, 1997). For both the Paragraph Recall and Attention measures, the summing of related variables is based on the following rationale: each variable within a given, related domain contains some "true" value of the variable in question, plus some random measurement error. Since the measurement error is random, summing over variables will tend to diminish the effect of error, and to increase the reliability of the sum score (Statsoft, 1999). For the Attention measure, the variables were chosen based on their difficulty, i.e. Choice Reaction times are normally slower than Simple Reaction times, and signaling the Absence of a target also takes longer than to signal its Presence; finally, the Criterion measure (number of trials to the criterion of 8 correct responses in a row) was considered difficult because no subjects attained a perfect performance. For the Paragraph Recall measure the range of scores was 2-44, with a mean of 22.4 and a S.D. of 10.2; higher scores indicate better performance. For the Attention measure the range of scores was –3.6 to 8.2 with a mean of -.05 and a S.D. of 2.3; lower Attention scores indicate better performance (e.g. shorter Choice Reaction time).Age, Education, Geriatric Depression Scale score and Memory Complaints scores were also included in the analysis. The results of all bivariate regression analyses are given in **Table 8**.

Table 8. Cortisol-Cognition Correlation Coefficients

	age	Education	GDS	MAC-Q	Attention	Paragraph Recall
Upon Rising	.07	.02	<.01	.21	.11	-.14
30 min later	.20	-.03	.14	.08	.07	-.05
2:00 PM	.22	-.05	.07	.10	.12	-.003
4:00 PM	-.01	.18	-.05	.15	.04	-.14
Bedtime	.23	-.06	.16	.27	.23	-.14
Grand Mean	.19	<.01	.12	.20	.13	-.12
Max	.17	.01	.01	.09	.14	-.14
Min	.12	.05	.13	.24	.07	-.07
MornCort Change	.14	-.10	.17	-.23	-.12	.17
Age		-.12	.16	.20	-.32*	-.43**
Education			-.11	.20	-.21	.15
GDS				.41**	-.01	-.10
MAC-Q					.02	-.27
Attention						-.33*

Legend:

Upon Rising to *Bedtime*: circadian cortisol concentrations
Grand Mean: 3-day mean cortisol
Max and Min: 3-day mean maximum and minimum cortisol
MornCort Change: Percent cortisol change from *Upon Rising* to *30 min later*
GDS: Geriatric Depression Scale
MAC-Q: Subjective Memory Complaints scale
Attention: Choice + Absent Target Reactions Times + Trials to Criterion in Conditioned Associative Learning test
Paragraph Recall: Immediate + Delayed Paragraph Recall

* = $p < .05$ ** = $p < .01$ (2 tailed)

As can be seen in the table, there were no significant correlation between any of the

chosen cortisol and cognitive measures. The two composite cognitive measures were significantly correlated, with about 10% of the variance in one of the measures accounted for by variance in the other measure.

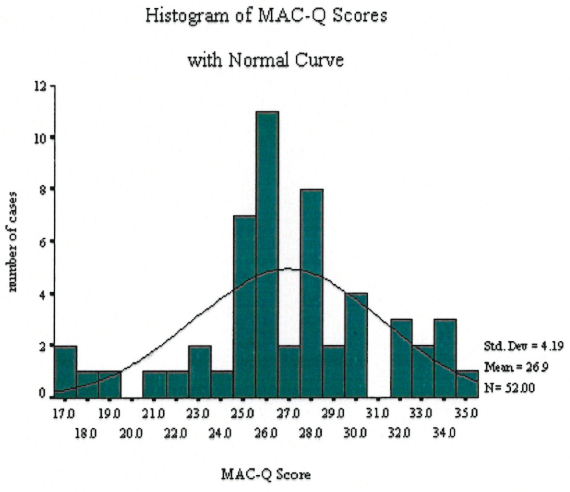
8.4.1 Memory Complaints, Cortisol and Cognition

Split median tests have been criticized on the grounds that they provide less power than regression analyses, and oversimplify a possible real-world quantitative relationship by reducing it to a qualitative binary state (Cohen, 1983). But the Memory Assessment Clinics Questionnaire (MAC-Q) was designed to do just that, to dichotomize memory complaints as either significant (at or above the cut-point of 25 out of a possible 35) or not (7-24 points), rather than to measure their severity. At the most basic level, one could ask: With every other factor confounded, do subjective memory complaints have any systematic relationship with cognitive performance or cortisol function? Since there were few differences between the Mem and MemDep groups as presently defined, it was decided to perform analyses of cortisol concentrations and cognitive performance by groups defined as above or below the MAC-Q median. This procedure tests the following hypothesis:

> Regardless of possible mediating factors (e.g. personality characteristics, mood, psychiatric status), subjects with high subjective memory complaints will, as a group, have higher cortisol concentrations and lower levels of declarative memory performance than subjects with no subjective memory complaints.

The median MAC-Q score in the present total sample is 26, a value which occurs eleven times. Thus, splitting the group into subjects with MAC-Q scores <26 and ≥ 26 produces groups of n = 16 and n = 36 respectively. Although MAC-Q scores were distributed normally according to the Kolmogorov-Smirnov statistic [KS (52) = .11, Lilliefors p = .16], a visual examination of the histogram suggests abnormal distribution more strongly than the KS statistic suggests normality (**Figure 30**). At any rate, raising the criterion for subjective memory complaints from 25 to 26 on the MAC-Q is akin to increasing the specificity of the test at the expense of its sensitivity, a move which also increases its positive predictive value at the expense of some negative predictive value. The sample tested for this study had on the whole[1] never previously undergone cognitive evaluation, making for an unselected (or rather, self-selected) population, in which, as discussed above, test cut-off points creating high specificity rather than sensitivity are (arguably) preferable.

Figure 30
Histogram of MAC-Q Scores

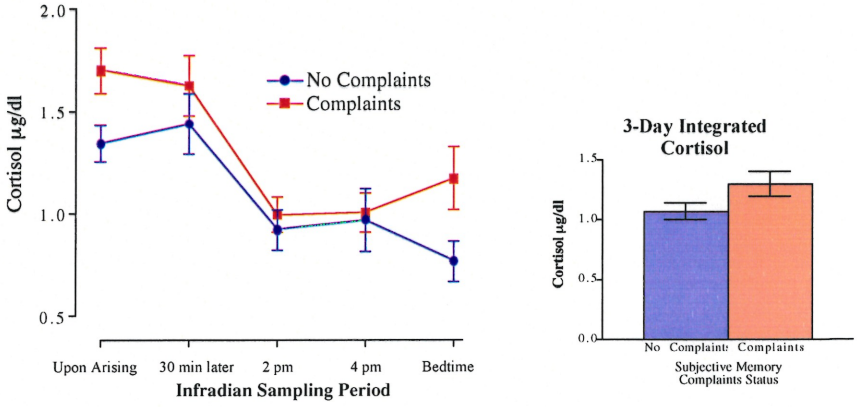


[1] Except for one subject, whose file was not consulted during this study.

A repeated-measures ANOVA on the infradian cortisol (i.e., the five daily mean cortisol concentrations) by MAC-Q (i.e., above or below the median memory complaint score) revealed no group differences [*Group* effect: F (1, 46) = 1.49, p = .23; *time* effect: F (4, 184) = 25.4, p < .001; *group by time* effect F (4, 184) = 1.9, p = .11]. There was no significant difference between the Complaints and No Complaints group for the grand mean cortisol [t (46) = 1.42, one-tailed p = .081] (**Figure 31**). There was also no difference in the morning Cortisol change [*Complaints* mean change (± standard error): -5.7 % ± 3.9; *No Complaints* mean change:+ 8.4 % ± 9.7; Welch's t (17.38) p = .19, two-tailed].

In order to test the total recall and rate of forgetting in these two groups, a two-way ANOVA was performed on the two components of the composite Paragraph recall measure: i.e. Immediate and Delayed recall of story elements. *Time* (of recall) was the repeated-measures factor, and Group (Complaints and No Complaints) the independent variable. There was a significant effect of time [F (1,48)= 12.94, p = .001) and

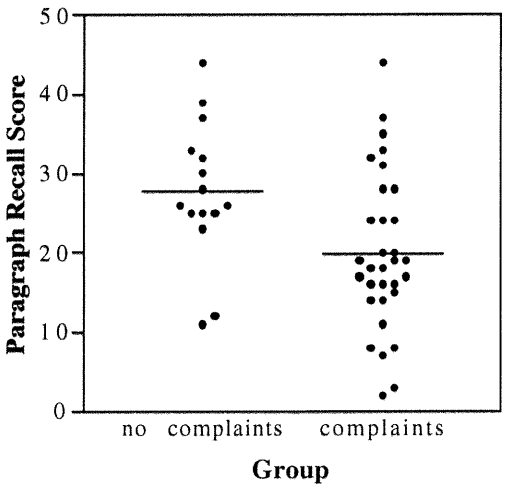
Figure 31
Infradian and 3-Day Mean Cortisol Concentrations by Memory Complaints Split-Median means with SE



group [F (1,48) = 8.33, p = .006], but no interaction [F (2,46) = .23, p = .80]. Since age was significantly correlated with the Paragraph Recall measures (Table 9), it along with education were entered as covariates in an ANOVA; this did not substantially alter the group effect [F (1,45) = 5.29, p = .007; *age* covariate: Beta = -.35, t = -2.75, p = .008; *education* covariate: Beta = .16, t = 1.31, p = .20]. Thus, both groups forgot story elements at an equivalent rate from Immediate to Delayed Recall, but in total (Immediate and Delayed recall combined), the Memory Complaints group remembered fewer items(**Figure 32**).

Figure 32

Scattergram of Paragraph Recall by Memory Complaints Split-Median
with means bars



8.4.2 Modelling Subjective Memory Complaints

A limited number of variables were tested in stepwise linear regression for their predictive value as regards subjective memory complaints scores. The goal of this procedure was to test a parsimonious model that would explain the maximum amount of variability in Subjective Memory Complaints with the minimum number of theoretically or empirically sound factors. The approach chosen was to select variables from dissimilar domains, based on the expectation that the predictors of memory

complaints would be multifactorial; i.e., not only cognitive (since in this study memory complaints were generally not strongly related to cognitive performances). In addition, variables were selected on the basis of there being significant Group differences on them, based on the rationale that since one of the groups had no memory complaints (Control group), a variable that was signicantly different in an ANOVA (if not significantly different in all three pairwise comparisons) was at least more likely to possess predictive utility with regards to memory complaints than one on which all three groups performed alike.

The variables thus selected were *Trauma* and *Name fluency* as well as morning change and grand mean *Cortisol. Trauma* was included as a predictor variable because of its power to account for reported stress and psychiatric symptom score differences between the Mem, MemDep and Control groups. *Name* fluency was chosen because it was one of the few cognitive performance measures on which group differences were significant (when analyzed as a univariate ANOVA). *Grand mean cortisol* was chosen because in this study we found significant correlations between the MAC-Q memory complaints questionaire and depressive symptoms as indexed by the Geriatric Depression Scale (Table 9); and it is a robust finding in psychiatry that depression is associated with high cortisol (Board, Wadeson and Perskey, 1957; Lupien et al, 1999). Depression itself was entered in the final block (see below), with *morning change cortisol*, which was chosen because of the significant Group effect in the

morning cortisol change, with the Mem group (representing "pure" memory complaints, i.e. unaccompanied by depressive symptoms) showing an anomalous cortisol decrease instead of the expected increase in the first 30 min after awakening (Pruessner, Wolfe et al, 1997).

In a stepwise multiple linear regression model with the criterion for inclusion of variables set at $p < .05$ and the criterion for removal at $p > .15$, and pairwise deletion of cases (the correlation matrix was examined beforehand to ensure that the missing data—about 4% of cases—were distributed randomly between variables), *age* and *education* were entered in the first block, in order to ensure that the effect of succeeding variables entered were unconfounded with basic demographic variables. In the second block *trauma* and *name fluency* were entered, and in the third *grand mean cortisol* and *morning cortisol change* were entered. Scores from the Geriatric Depression Scale (GDS) were also entered in this third block in order to see if depressive symptoms as indexed by GDS scores (which correlate with MAC-Q scores with an $r = .4$, $p = .003$), could account for any more variance in MAC-Q scores once the *Trauma* variable was accounted for.

Age and education were not significant variables and did not enter the equation. From the second block, Trauma was entered first and was significant [F(Trauma) (1,43) = 6.1, $p = .017$, adjusted R square = .1]. Name was entered next and was also significant

[F(Name) (2,42) = 5.3, p = .009, adjusted R square = .16]. Of the two variables in the third block, the GDS was not significant and was not entered, but grand mean cortisol and morning cortisol change were entered as significant [F (Cortisol) (4, 40) = 5.6, p = .001, adjusted R square = .29]. The final model contained as predictor variables Trauma (Beta = .46, p = .001), Name (Beta = -.26, p = .05), grand mean cortisol (Beta = .35, p = .01) and morning cortisol change (Beta -.30, p = .03), with an adjusted R square of .29. Variables not entered [with their Beta and p values) were Age (Beta = .17, p = .20), , education (Beta = .17, p = .16) and GDS scores (Beta = .28, p = .10)].

9. Discussion

9.1 Abstract of Significant Results

- In this study, the study groups were comparable in age, education and MMSE scores.
- There was no absolute differences in paragraph recall, but the MemDep group recalled fewer negative- than positive-valence items.
- The Mem group had lower performances at delayed than at immediate recall in the Visual Reproduction task.
- There was a significant group effect in Name fluency (with the Control, MemDep and Mem groups in rank order from best to worst), but no significant pairwise differences.
- the MemDep group had the highest absolute cortisol concentrations 30 min after awakening, but the Mem group had a cortisol decrease in that same 30 min period that was significantly different from the increase found in the other two groups.
- The MemDep group had significantly higher stress, psychiatric, and trauma symptom scores than the two other groups.
- Subjects with memory complaints above the median performed worse in a paragraph recall task.

9.2 Overview of Study Population

It is remarkable that one third of our total recruited sample (MemDep and Dep groups comprised 33% of the sample) were seniors whose scores on the Geriatric Depression

Scale were above the cut-off normally indicative of depression. These seniors with possible undiagnosed depression were three times as numerous as the treated depressed group (the treated "Anti-Dep" Group, who underwent evaluation but were excluded from the present study, comprised 10% of the total sample). Thus, the association of stress and memory complaints with depression appears to be warranted: a presenting complaint of memory problems, or excessive stress, should at least trigger an examination for signs of depression. However, another third of our total recruited sample was comprised of seniors with memory complaints only, and no significant, or only a mild degree of, depressive symptomatology. On the whole there were few cognitive differences between subjects with memory complaints only and those with both complaints and depressive symptoms.

9.3 Methodological considerations

The external validity of an observational design such as the present study could be questioned on the grounds that the categorization of the subjects was artificial and simply has no counterpart in real life: subjective memory complaints alone, or complaints with depressive symptoms, do not constitute disease entities. However, the relevance of this work to the world outside the laboratory can be affirmed with reference to Oakes (1972), who advanced what might be called the "external validity in spite of" argument; and to Mooke (1983), who advanced what might be called the "external *in*validity in spite of" argument. Both of these arguments are "in spite of"

because both Oakes (1972) and Mooke (1983) advocate the value of a given research project *in spite of* potential claims of invalidity that could be leveled against it. Mooke (1983) describes four possible goals of "externally invalid" research: (1) to investigate "what can" happen (in the laboratory, for example) versus "what does" happen (in "real life"); (2) to test "what should" happen for theoretical reasons; (3) to demonstrate the power of a well-described and theoretically grounded phenomenon; and (4) to produce conditions which have no counterpart in life in order to obtain information otherwise unattainable. The design and execution of this study could be framed in terms of the second and fourth propositions above. "What should" happen for theoretical reasons was described in the research hypotheses. And the artificial groupings of subject into the Mem, MemDep and Control groups—which would not be encountered in real life or in normal practice—was effected because only by creating such arbitrary distinctions was it possible to systematically observe the effects of depressive symptoms on the relation of memory complaints to objective memory performance.

With Oakes (1972), the argument as it applies to this project it that any reliable findings produced by it should be accepted as valid, with the limitation that they perhaps could not be generalized to other samples, which might have other characteristics. It does not seem unreasonable to assert, however, that the present sample would not be too dissimilar to other samples that could be formed from

persons presenting to a Memory Clinic. But the possibility of generalizing the present findings to other samples would decrease the further removed in provenance and composition those other samples would be from the original studied sample. In other words, the present findings would most apply to another sample of volunteers recruited by newspaper advertisements to participate in a research project on stress and aging at the *Institut universitaire de gériatrie de Montréal*. Next most similar in composition and provenance would be a sample of self-referred subjects presenting at a Memory Clinic with subjective memory complaints; next after that, perhaps, would be a sample referred to the clinic by a primary care physician. These findings should be generalized with caution to any sample which departs substantially from the basic demographic characteristics of the present sample. But the differences which resulted in this study were presumably robust enough to become apparent in spite of the heterogeneity of entry characteristics. These results could thus perhaps best be generalized to other samples of equal heterogeneity (i.e., without strict inclusion and exclusion criteria) while keeping close to our sample's basic characteristics (e.g. 74% women, mean age around 67 with at least a High School education). Of note, there were not enough subjects of both sexes within the three groups to include sex as a factor in analyses. Independent *t*-tests performed on a representative result from the total sample from each domain tested did not reveal significant sex differences. But it remains unknown whether these results apply equally to both sexes, or only indeed to

samples containing an proportion of men and women equivalent with that obtained in the present study.

It could be questioned whether by excluding 24% of the tested sample (i.e., subjects who were taking anxiolytics or anti-depressants or who were only depressed, with no subjective memory complaints), we invalidated our results. However, one can argue that on the contrary it becomes more possible to generalize to other populations (specifically, memory clinic populations—restricted to those not taking anxiolytics more than twice or week and also not taking anti-depressant medicines) than it would have been had we included all tested subjects. The methodology employed in defining the study group was an attempt to strike a middle ground between exhaustive inclusion and exclusion criteria which, as discussed in the introduction, can lead to experimentally well-controlled but unrepresentative groups, and purely observational studies where the studied subjects are representative of the population, but an endless number of possible confounding factors remain uncontrolled. In certain cases, for example in pathophysiological or natural history studies of AD, a maximum degree of homogeneity of research samples can be logically required (Berg et al, 1982). But in other cases, we are arguing, a lesser degree of homogeneity can be accepted without any untoward effect on the validity of the findings.

Had we *not* excluded subjects on anxiolytics or anti-depressant medication, it would

have been all but impossible to account for the possible cognitive effects of the different medicines and different doses in different subjects. In addition, the forming of a "Control" group after the fact (it had not been planned) could be questioned. It is important to note however that we decided to include the Control group after the testing had been completed, but before any data analysis. It was unknown how the Control group would perform. They were included in order to provide a reference point in drawing any contrast we might draw between the Mem and MemDep groups. The Control group were not added post-hoc in the sense of after seeing the initial results; the only analysis that had been done at that point was the tabulation of the MAC-Q (for memory complaints) and the GDS (for depression symptoms), as well as the concomitant medication questionaire.

Another potential critique of the present work relates to the large number of outcome variables, and the statistical corrections employed, or not employed, for multiple comparisons. In most studies on cognitive aging an extensive test battery is used. We used not only an extensive neuropsychological test battery, but we tested multiple domains (e.g. social support, cortisol function), which is in line with a recent recommendation that research on MCI study not only cognitive function, but also subjective memory complaints, and other non-cognitive variables (Petersen et al, 2001). Because this is an exploratory study; the few significant results that have been found will need to be replicated in a confirmatory study (as part of an longitudinal

project in progress), where they would form the basis of hypotheses defined a priori (and thus needing no multiplicity adjusments).

The power of this study to demonstrate all possible group differences could also be questioned. Indeed, some "nearly significant" (p values between .05 and .1) results were possibly due to a too small number of subjects, particularly in the Control group (n = 9). Although there was a fairly consistent rank ordering of outcome measures, from Control to MemDep to Mem, few of them approached statistical significance and performances of the Control group in particular must be interpreted with caution.

The use of Bonferroni adjustments for multiple comparisons is generally regarded as too conservative, because it inflates the Type II error rate (Pernerger, 1998). However, the use of other statistical techniques to counter inflated Type I errors due to multiple testing is controversial (Aicken, 1998; Pernerger, 1999). Some advocate that no multiplicity adjustments be made for exploratory studies (Bender and Lange, 1999). Accordingly, for example, no corrections were made for multiple univariate F-tests following a significant MANOVA on the Derogatis Stress Profile scores. In this study the Spjotvoll-Stoline adjustements were actually used infrequently within individual outcome measures. On the other hand, it is not generally considered advisable to adjust alpha levels across testing domains (i.e. for independent outcome measures). Granted that not all of our outcome measures were completely independent, but just

because there were so many outcome measures (because of the exploratory nature of the study), had we adjusted the alpha level across all outcome measures the alpha would have been extremely conservative and that would have all but guaranteed no positive results, and a high risk of Type II error.

However, just as the definitions and cut-points used to study other categories of aging-associated cognitive impairment have been criticized (Smith, Ivnik, Petersen, Malec, Kokmen and Tangalos, 1991), so too in this study were group definitions and cut-points problematic. This was shown by the significant group difference in a split-median test of paragraph recall by MAC-Q memory complaint scores, where there were no significant differences between the groups as originally defined. The present findings would have doubtless been different had the cut-points on the MAC-Q and the GDS depression score been moved just one point up or down; likewise had subjects taking anti-depressants or anxiolytics more than occasionally not been excluded from the analyses. Thus the critique that was made above, and in the literature, on the arbitrariness of the operational definitions of the various mild cognitive impairment categories extant, could justifiably apply to the definition of the categories used in the present study. On the other hand, although there was some degree of arbitrariness in the cut-points and the instruments used to define the study groups, it was minimal. That problems with cut-points were encountered in spite of the groups being defined as simply as could be while still addressing the research

question, serves to highlight the problem of "premature homogenization" of study groups in cognitive aging research. The state of knowledge about the earliest signs of AD, and the borderline between normal and pathological brain aging—or "usual" and "successful" aging, to use the terms introduced by Rowe and Kahn (1987) to address some of the heterogeneity of seniors' cognitive performance—is still insufficient to allow one to "carve nature at its joints" (Kendler and Gardner, 1998), i.e. categorize people as they are categorized by nature.

9.4 Depression and Cognition

The MemDep group recalled fewer negative-valence items in a paragraph recall task at both Immediate and Delayed free recall. We presented negative- and neutral-valence material. This is in contrast to Danion et al. (1995), who presented neutral-, negative- and positive-valence material to depressed subjects, and found greater recall of both negative and positive material. Our findings also diverged from those of Bradley et al (1994), who studied as we did a non-clinical sample; they also found greater recall of negative material in high negative affect subjects (once anxiety was partialled out), but found greater implicit memory for negative items. We did not test for a valence effect in implicit recall. The effect of valence on Recognition was non significant, but in the same direction as found with Bradley et al (1994), i.e. higher for negative- than for neutral-valence material. A possible explanation relates to the depth of encoding. Silberman and colleagues (Silberman, Weingartner, Laraia, Byrnes and

Post, 1983) found that normal and depressed subjects rated the emotionality of verbal material identically, but depressed subjects were more dependent on high emotional valence for recognition, while they were not helped by high valence during free recall. Because depressed patients rated the emotionality of the material the same as the normal group, Silberman et al (1983) speculated that shallow processing in depressed subjects (i.e. when not specifically directed towards more effortful processing) made them more dependent on high emotionality for memory processes.

Our findings of impaired recall in the MemDep group of negative-valence words extends the work of Stip and Lecours (1992) and suggests the presence of interference, or tying up of attentional resources to negative-valence material that is congruent with their mood. As regards the finding of increased recognition of negative-valence material, it would be interesting to repeat this experiment while varying the degree of self-relevance. The story used in paragraph recall could be considered highly self-relevant to our subjects, featuring as it does a protagonist who forgets things and finally consults a doctor because of worries about memory. In future it could also be useful to verify whether subjects were conscious of the emotional charge of the stories, which as Calev (1996) suggests could affect the mood-congruence effect, and may be a cause of these inconsistent results.

These findings support Hypothesis 1: "Since depressed persons have a cognitive bias

against stimuli of emotional of negative valence, the MemDep group will recall fewer neutral than emotional elements at both immediate and delayed recall, whereas there will be no difference in recall by valence for the Mem group." However, whether such a cognitive bias against negative-valence self-relevant material would hold in other experimental conditions which differed in degree of self-relevance, and the subject's awareness of the story's emotional charge, is an empirical question. It also does not give any information as to the magnitude of the effect such a bias would have in daily life. Most importantly, it does not differentiate between "negative valence" and general "emotionality"—perhaps the same effect would have been found with positive vs. neutral valence items (rather than negative vs. neutral). In addition, the effect found in this study could be due to some factor other than valence or emotionality of the story elements; concreteness, for example (Silberman et al, 1983).

9.5 Psychomotor and Frontal Function

Psychomotor Tests

That we failed to find a group effect on psychomotor tests such as Simple or Choice Reaction Time, or the Selective Attention test, is reassuring. Nebes and Brady (1992) found a robust (across 61 experimental conditions) increase in the effect of increasing task complexity on reaction time in AD patients. Both normal controls and depressed subjects had slower psychomotor speed as the complexity of the tasks (and thereby information processing demands) was increased; but the AD patients could be

differentiated on the basis of their proportionately greater slowing. This merits careful follow up with tasks where the range of complexity is wider. Our Selective Attention task for example had 2, 4, 7 or 10 distractors which gradually cluttered the visual field subjects had to search for the target. It is possible that a differential slope effect would only have become apparent if we had tried up to 20 or 30 distractors, or varied task complexity by other means that varying the number of irrelevant features in a visual search task. It is possible that the increasing information processing demands the subjects faced as the Selective Attention decision task progressed from a two- to a ten-distractor condition, was not burdensome enough to produce differential mean group increases in response times by task complexity.

Frontal Function Tests

There were no group differences on the Conditioned Associative Learning test, a probe of frontal lobe function. In general, although the MemDep group showed a cognitive bias against recall of negative material, they were not significantly impaired on frontal tests. This suggests that emotional bias in depressive symptomatology might be an earlier marker of cognitive deficits related to depression, rather than typical "frontal" deficits per se. However, we did not test for other memory functions that are believed to be very specific to the frontal lobes, such as memory for temporal order (Milner, Petrides and Smith 1985), therefore frontal deficits in either of the studied groups can not be ruled out.

Although verbal fluency tasks are used as index of frontal functioning (Binetti, Magni, Padovani, Cappa, Bianchetti and Trabucchi, 1996), other investigators have found that fluency tasks are sensitive instruments for detecting AD but that they are not related to either frontal or frontal/parietal perfusion as indexed by single photon emission computed tomography (Pasquier, Lebert, Grymonprez and Petit, 1995). Monsch at al (1992) found that category fluency distinguished between AD and normal subjects better than name or letter fluency; they ascribe this to the deterioration of access to semantic knowledge in AD patients. This interpretation is also advanced by Geffen and collaborators (Geffen et al, 1993), who studied normal, AD and depressed subjects and found that this latter group showed a deficit compared to normal controls only in effortful declarative memory tasks; in fluency tasks, both depressed and normal subjects, but not AD subjects, showed greater performance in category rather than in letter fluency (Geffen et al, 1993). That the Mem group showed an impairment in Name fluency relative to the MemDep and Control groups is difficult to interpret because Name fluency is neither clearly a purely "verbal fluency" task such as Letter fluency, nor a purely "semantic fluency" task such as Animal fluency. Fluency for first names has characteristics of both declarative and semantic memory, in that it is possible that the names that come to mind are from the subject's social circle. This could be explored systematically, for example by first having subjects perform a Name Fluency task and then in having them enumerate all their

friends by first name. The conservative conclusion to draw is that as regards risk factors for AD in this context, it is probably not relevant that the Mem group had lower Name fluency than the other two groups, given the lack of difference on the more sensitive (possibly because more clearly semantic) Animal category fluency.

9.6 Stress and Psychiatric Symptoms

Hypothesis 2, "Since stress may cause depression, the MemDep group will have higher self-reported stress than the Mem group" seems to be supported by the data. On virtually every subscale of the Derogatis Stress Profile (DSP), the MemDep group reported higher stress. The DSP measures a person's current stress status in terms of eleven stress dimensions, which load onto three stress domains: environmental, emotional response and personality mediators. Relative to the two other groups, the MemDep group had higher environmental stress, and higher stress due to emotional response to stress, but not higher stress due to personality factors. This suggests that the origin of stress in these individuals is external more than internal. It is also interesting in the light of the effect that including Trauma symptomatology had on the stress measures: it made the group differences no longer statistically significant. The use of ANCOVA in observational studies such as this one where the covariate is also an outcome measure calls for extreme caution and should only be done with clear theoretical justification (Porter and Raudenbush, 1987). Such justification provides from the known deleterious effects of early traumatic experiences on coping skills

(Heim and Nemeroff, 1999). There was no direct interview about traumatic experiences in this study, and it is not certain that the high Trauma Symptom scores actually relate to past traumatic experiences. But the lack of personality mediator effects on the DSP, together with the power of trauma symptomatology to account for different reported stress levels, suggests that for some subjects with both memory complaints and depressive symptoms, impaired coping skills due to early traumatic experiences may be a problem.

9.7 Cortisol

9.7.1 Cortisol and Depression

Hypothesis 3, "Since depressed persons have higher cortisol levels than normal controls, the MemDep group will have higher cortisol than the Mem group," was not supported. Although the MemDep group had higher cortisol means at 3 of the 5 daily cortisol sampling times (30 min after rising, 2 p.m., and bedtime), these differences were not statistically significant. The Memory Complaints group (as defined by a split-median test) was the one that actually presented neuroendocrinal abnormalities: a morning cortisol drop instead of a normal and expected rise, combined with an overall trend towards a higher 3-day mean basal cortisol.

9.7.2 Cortisol and Cognition

Hypothesis 4, "Since high basal cortisol levels can have deleterious effects on declarative memory, declarative memory scores will be inversely correlated with basal

cortisol levels in both groups," was not supported. Lupien and colleagues' (1994) findings on the negative relationship between basal cortisol and declarative memory had used a different sampling methodology, i.e. twenty-four-hour sampling with an indwelling forearm catheter. Their significant findings as regards declarative memory only were found in a "positive slope elevated" group that had both high basal cortisol and increasing basal cortisol over a span of years. Other investigators have also used special procedures to increase precision. Pruessner, Wolf et al (1997) used multiple salivary sampling every 10-15 min for 30 minutes to an hour after awakening; our subjects were instructed simply to take a saliva sample upon awakening and again half an hour later. It is possible that the trend towards a morning cortisol drop in the Memory Complaints group was due to their taking either the first or the second sample later than instructed, thereby missing the brief morning cortisol peak. Home-based sample collection procedures have that pitfall of increased feasibility at the cost of decreased precision of measurement. Another technique we could have used is to vastly increase the redundancy of testing (beyond the three consecutive days we already used), which can lead to significant correlations where potential results from fewer sampling points are lost within data variability (Pruessner, Gaab, Hellhammer, Lintz, Schommer and Kirschbaum, 1997).

9.7.3 Cortisol and Cognition in Depression

Hypothesis 5, "Since depression is related to high cortisol in seniors, and high cortisol

is related to declarative memory deficits, the MemDep group should have impaired declarative memory function relative to the Mem and Control groups" was not supported apart from the mood-congruence effect already covered by Hypothesis 1. The only cognitive deficit that obtained with the MemDep group was reduced recall and increased recognition of self-relevant negative-valence material in a paragraph recall task.

9.7.4 Links and Divergences With Other Studies

In our study we did not find clear evidence supporting the hypothesized relationships between cortisol and cognition, nor between cortisol and depressive symptoms. However, we did find that subjects with memory complaints (whether defined as the Mem group or as all subjects with complaints above the median) had a tendency towards having both higher integrated (3-day) mean cortisol concentrations, and a morning cortisol drop instead of the normal and expected cortisol rise. Subjects with memory complaints above median also performed worse on a paragraph recall task (immediate and delayed recall combined). That we even obtained such a result is somewhat surprising considering that in the literature one finds evidence of the lack of relationship between cortisol and cognition in cross-sectional studies, such a relationship is salient only in longitudinal studies (Lupien, Lecours, Schwartz, Sharma, Hauger, Meaney and Nair, 1995). Thus the cortisol-cognition and cortisol-depressive symptom data gathered in this study, although failing to provide

convincing evidence that any of the groups studied is truly at risk for the development of AD, also form the beginning of a longitudinal study being carried out by Dr. Sonia Lupien. As such these data provide promising leads to investigate. For example, previously a significant correlation between the obssessive-compulsive subscale (OCS) of the psychiatric Symptom Checklist-90 and cortisol slope (falling or rising integrated cortisol levels over years) had been noted (Lupien, Lecours et al, 1995). In this study the MemDep subjects had higher OCS scores than the other two groups (as shown in Figure 27, Section 7.6). There was also a significant correlation between MAC-Q memory complaint scores and that same OCS [$r = .54$, $p < .001$]. The focus of the OCS scale is on behaviours identified with the clinical syndrome of Obsessive-Compulsive Disorder (OCD); however, behaviors indicative of more general cognitive problems (e.g. "mind going blank," "trouble remembering") also load on this dimension. A relation between subjective memory complaints and OCD has not previously been noted in the literature, which suggests that the significant correlation is due to the cognitive items within the obssessive compulsive subscale, rather than to a developing psychiatric disorder. This lends some support to the usefulness of the MAC-Q as an indicator of at-risk status (the MAC-Q can also be completed in two or three minutes, compared to the hour or more the SCL-90 can require). Whether subjects with high memory complaints do in fact show an increasing cortisol slope with years remains to be determined.

In addition, the morning cortisol change measure was intriguing in that even the Control group in this study showed a very modest rise in cortisol in the first half hour of the day that was well below the 50-75% range reported by Pruessner for healthy subjects (Pruessner, Wolfe et al, 1997). Whether this discrepancy relates to inadequate adherence to study directions during the saliva sampling, to laboratory error, to the well-known and troublesome individual variability in cortisol secretion (Meaney, O'Donnell, Rowe, Tannenbaum, Steverman, Walker, Nair and Lupien, 1995), or to real HPA-axis abnormality even in the "Control" group (who were not operationally defined to be normal) is unknown.

9.8 Predictors of Memory Complaints

The multiple linear regression of selected factors on MAC-Q does not signify causality, but it is interesting that in this model, age and education did not enter into the equation, and neither did the GDS depression scores, once Trauma had been entered. Thus one psychosocial factor (Trauma), along with one neuropsychological (name fluency) and two neuroendocrinal factors (grand mean and morning change cortisol), together could account for about one third of the variance in MAC-Q scores. This begins to draw a picture that is more complex than simply a relationship, or a lack of relationship, between memory complaints and objective memory performance.

9.9 Psychoneuroendocrine Frailty

The basic controversy in the literature is whether depressive symptoms make

subjective memory complaints somehow less valid as indicators of objective memory problems because of their stronger association with a mediating variable, namely depressive symptomatology. But the whole problematic enterprise of determining the precise, predictable relationship between subjective memory complaint and objective cognitive performance may be obviated by conceptualizing memory complaints differently, as symptoms of frailty. Frailty refers not only to a person's physical susceptibility to disease and injury but also more broadly to the spread between the demands and opportunities of a person's situation and her or his capacities to meet and make use of them (Raphael, Cava, Brown, Renwick, Heathcote, Weird, Wright and Kirwan, 1995; Lebel et al, 1999). Again, frailty is not just a synonym of physical weakness or proneness to injury and disease. McGougall and Balyer, for example, described decreased confidence in memory and depression as symptoms of mental frailty (McDougall and Balyer, 1998). Depression and social isolation have also been found to be significant predictors of frailty (Strawbridge, Shema, Balfour, Higby and Kaplan, 1998). Bedi's (1999) conceptual treatment of depression as an imbalance between individually evaluated stressors and neurotransmitter resources to produce appropriate behavioural responses is squarely in the spirit of this multi-disciplinary study because it does justice to both sides of human experience—the subjective and objective—without reducing one to the other (Velmans 1993). Clinically, therefore, subjective memory complaints are important for much more than for what they can tell

us (or fail to tell us) about a person's objective cognitive capacities. Seen as an attempt to redirect resources in order to help the person's ability to cope with a perhaps unrealized adverse situation (the sequellae of a traumatic experience, perhaps?), memory complaints become an invitation to explore the patient's inner life.

9.10 Screening for Cognitive Impairment

It should be noted here that the proliferation of terms used to refer to subtypes of MCI is rather confusing. Two recent papers contained a new coinage and a new proposed definition of an existing term, which could help rationalize matters. Unverzagt and colleagues reporting on the Indianapolis study of Health and Aging coined the term "medically unexplained memory loss" (MUML) (Unverzagt, Gao, Baiyewu, Ogunniyi, Gureje, Perkins, Emsley, Dickens, Evans, Musick, Hall, Hui and Hendrie, 2001). Both Circumscribed Memory Impairment (CMI) and MUML are contained within the Cognitive Impairment, No Dementia (CIND) category (CIND includes cases where the cognitive impairment is attributable to non-disease factors, e.g. mental retardation or cultural differences); however, the term "MUML" does not convey the impression that the only cognitive impairment is with memory. By conrast, the term "circumscribed" in the "Circumscribed Memory Impairment" does contain the suggestion that the deficit is limited to memory. "MUML," however, is a label given to an anomalous situation regarding memory, without ruling out other possible deficits. In addition, Ritchie and Touchon (2001) have proposed that the term MCI be

thought of as a significant deficit in any cognitive domain, with the assumption that it is not part of the normal aging process. The use of MCI and MUML in this way would create a simpler nosological picture: CIND would refer to cognitive impairment due to any cause except diagnosed dementia; MCI would refer to a cognitive deficit that is assumed to be abnormal; and MUML would refer specifically to an abnormal memory deficit.

The lack of redundancy between cognitive screening instruments in this study—with seventeen of 52 subjects scoring beyond the cut-point for suspected cognitive impairment on one of the three screening instruments (MMSE, its extended form the 3MS, and the ADAS-cog), but only four on any two of those instruments, and a single subject on all three, is perhaps surprising. But those cut-points are a matter of still-evolving consensus. The MMSE is mostly used to screen for possible AD; the more sensitive 3MS has only theoretically been studied for its ability to detect mild cognitive impairment (Kristjansson et al, 1996), and the ADAS-cog is mostly used in dementia drug research. The classification of seventeen, four, or one subject as possibly cognitively impaired on one, two or three of those instruments is mostly a product of the experimental cut-points used in this study to define theoretical "mild cognitive impairment"cases. Although cases of reversible dementia are a distinct minority in the clinic, one chart-review study found that 23% of cases presenting at a memory clinic had potentially reversible dementia; but only 3.6% of cases were

actually reversible (Freter, Bergman, Gold, Chertkow and Clarfield, 1998)—nevertheless, those cases that did reverse all had high cognitive screening scores, which highlights the importance of research in this liminal zone between normal and pathological cognitive aging.

Other important and yet unresolved questions surrounding AD screening include: (1) the effects on those who are actually healthy but screen positive of receiving the dire prognosis of AD; (2) the cost to society and patients of intensive evaluations of these suspected cases who turn out to be false positives; (3) the false sense of security imparted to the others who screen negative but who really have incipient AD; and (4) the uncertain benefit to those patients who do screen positive and do have AD. Screening for cognitive impairment can be compared to screening for diabetes. Just as the population of patients with diabetes or pre-diabetic Syndrome X is expanding (partly because of modified screening standards [American Association of Clinical Endocrinologists, 2001]), it seems inevitable that the population of possibly pre-AD MCI patients will also expand in coming years. The effort to improve early detection of AD through conceptual and experimental work on its prodromal stages is similar to the effort to manage and prevent diabetes through such conceptual and nosological advances as the description of "Syndrome X" (Reaven, 1994). There are similar sorts of problems, too, with the ethical issues that arise from the costs involved in screening entire populations and then in offering, witholding or even recommending such

treatments as are or become available. As the state of knowledge advances it may be possible in the near future to determine with some fair degree of acuracy who is showing very early symptoms of AD. But when one considers that at present there is no consensus as to drug treatment of AD (Pryse-Phillips, 1999; Gauthier, 1999), it seems unlikely that government Health Services or private insurers would reimburse acetylcholinesterase inhibitors for a large minority of adults aged 65 and above to possibly prevent AD (assuming for the sake of argument that such a treatment would be beneficial). However, as knowledge accumulates on the neurobiology of AD, attitudes and practices will likely change. In a discussion on the ethics of limiting access to services, Sommerville (2000) argued that an ethical case could be made that physicians have the duty to inform their patients of all potential therapeutic options where the benefits would be reasonably expected to outweigh the risks. In the case of diabetes, faced with rising incidence rates, and the looming possibility that insurers would not reimburse chronic drug treatment of large segments of the diabetic population, physicians are being motivated to pay greater attention to the possibilities of early intervention, particularly dietary intervention (Eschwege, 2000). Thus, in the case of MCI, physicians may have a reason to discuss with their MCI or even with their worried-well elderly patients the option of taking long-term Vitamin E (Hogan and Black, 2001), or of modifying their diets to include sufficient quantities of spinach or strawberries (Joseph, Shukitt-Hal, Denisova, Prior, Cao, Martin, Taglialatela,

Bickford, 1998) in order to obtain a likely preventive or disease-slowing effect. Just as with prenatal folic acid recommendations for the prevention of neural tube defects, it is conceivable that in the near future, prevention-oriented recommendations would even extend to the prenatal period—at least that is one implication of the work of Cermak and colleagues on the long-term memory-enhancing effects of prenatal choline supplementation in rats (Cermak, Holler, Jackson and Blusztajn, 1998).

SADAS-cog: Time for a New Gold Standard of Quick Cognitive Screening Tests?

The SADAS-cog proved easy and quick to administer and permitted finer differentiation between subjects. Its criterion validity was supported by preliminary MRI data (Appendix 5). The Mini-Mental State Examination is too entrenched in clinical practice to be quickly replaced even by a superior test, but in centers where the extra time the ADAS-cog takes to administer (about 10 min vs. about 3 for the MMSE) is available, it should be considered. A training video for the standardized form of the test (SADAS-cog) is available from the test's developers (Standish et al, 1996)2, which answers concerns raised about standardization across centres (Pena-Casanova, 1997). The American Academy of Neurology recently recommended the use of general cognitive screening instruments in all elderly patients (Petersen et al,

---

[2] Tim Standish, c/o: McMaster University - Department of Medicine. Email: wgrg@fhs.mcmaster.ca

2001), because of a higher prevalence of memory impairment with increasing age. As this becomes put into practice, it would be useful to be able to compare across centers not only the well-known MMSE but the more informative ADAS-cog.

9.11 Clinical Significance of Memory Complaints

The Mem group showed several anomalies: no cortisol rise in the morning; increased rate of forgetting in Visual Reproduction, and decreased name fluency. The MemDep group high trauma and stress, and a emotional bias in cognition. Doubtless some in the Control group were denying their cognitive difficulties with the MAC-Q, the answers to which did not always match their oral self-description as persons with serious memory problems.

The slight visual memory impairment in the Mem group is possibly of clinical importance. The entorhinal cortex and adjacent area is both important for visual memory (Gaffan and Parker, 1996; Suzuki, 1996) and one of the first brain areas to show signs of AD (Braak and Braak 1991, 1992, 1999; Gómez-Isla et al, 1996), even before the hippocampus (de Leon, Convit, Tarshish, DeSanti & Bobinski, 1999). Impaired visual recall has also been found to be a marker of Age Associated Memory Impairment (Soininen, Partanen, Pitkänen, Vainio, Hänninen, Hallikainen, Koivisto and Riekkinen PJ Sr, 1994). Thus, in spite of the negative finding on psychomotor tests, the slight impairment in visual memory in the Mem group relative to the two

other groups merits careful follow-up.

A split median comparison of memory complaint groups showed a clearer separation of cognitive performance. Increasing the cut-point on the MAC-Q may be indicated in unselected populations, in order to increase specificity of "true" memory complaint status. Some volunteers with memory complaints may be interested in research, or in company, or in stimulating exchanges with study personnel; a higher MAC-Q score could help detect those who really are motivated by memory problems. This goal however only holds for the purposes of research; for the purposes of clinical care of seniors, any degree of memory complaint should be seriously as a possible sign of some kind of disturbance, though not necessarily cognitive per se.

Even if memory complaints have zero relationship to objective memory performance, they can still be an important call for help. Murrel and Norris (1984) found an interaction between resources, life events and depression such that depression was the result of a combination of weak resources and high adverse events; subjects with weak resources and low or moderate adverse events actually showed improvement in mood. Access to resources is extremely important to the well-being of seniors and lack of resources is an important cause of frailty (Lebel, Leduc, Kergoat, Latour, Leclerc, Béland & Contandriopoulos, 1999). Resources include coping capacity and sense of self-mastery, as well as social and community resources and, not least, health care

resources (Lebel et al, 1999). If a senior experiences high or increasing adverse events with an unchanging access to resources, an imbalance will be created along with much distress. In this light, subjective memory complaints—complaining to somebody—can serve the much-needed purpose of reorienting external resources, e.g. by motivating memory clinic workers and researchers to examine them and thereby, somehow, help them cope with a yet ill-defined problem. (The attempt, one imagines, is not always successful: indeed, involvement with a memory assessment clinic can add to the burden of patients' definition of their problem [Cremin , 1992].) As we have seen with the MemDep group's high level of traumatic symptoms, the story does not necessarily begin and end with memory.

Finally, the interest of subjective memory complaints is that they are one of only two ways to detect cognitive decline transversally (at one given time). The other is through informant interview, which does not apply for isolated patients. Of course both are subjective, and not always accurate (McGlone, Gupta, Humphrey, Oppenheimer, Mirsen and Evans, 1990). But they have the advantage of obviating reliance on transversal neuropsychological data in order to detect decline. If, as Morris et al (2001), one is to assume that in the absence of disease, memory performances should remain relatively stable during aging, then some kind of longitudinal reference data is needed in order to detect decline in individuals (rather than statistically, in groups). The difficulty with neuropsychological data to detect decline at a single point in time

is that on most measures, the performance of AD patients, questionable cases and normal controls overlap (Storandt and Hill, 1989). Morris and colleagues hypothesized that the success of their group in identifying incipient-AD MCI cases up to 9.5 years before AD diagnosis was partly due to the diagnostic process taking place without information on neuropsychological test scores (Morris et al, 2001). Individual differences in test scores can obscure the distinction between normal aging and incipient AD (Morris et al, 2001). These differences could be "averaged out" when making group comparisons such as in this present study, but any statistically significant difference, such as performance in paragraph recall by subjects with memory complaints above or below the median, does not apply to any given individual. Such individual information is a matter of clinical judgment (O'Connor et al, 1991).

An Apgar Scale for Mild Cognitive Impairment

The focus on subjective memory complaints in this thesis was inspired by a feeling that in research on cognitive aging, the point of view of the person concerned—the patient or research participant—had been neglected in the emphasis on standardized psychometric testing. These concerns can perhaps be brought into focus when one considers two recent papers, Feinstein's (1999) paper contrasting psychometrics and "clinimetrics" (the art and science of making a clinical judgment), along with Morris et al's (2001) paper "Mild cognitive impairment represents early-stage Alzheimer

Disease." The research success of Morris and colleagues (2001) is due to just that approach that Feinstein (1987, 1999) said was being neglected: clinical diagnosis. Feinstein (1999) used the example of Virginia Apgar's (1953) development of the Apgar rating scale for the health of newborn babies to discuss two complementary approaches, psychometrics and clinimetrics (Feinstein 1987). Psychometrics consists in giving standardized tests and then in following rules to produce a categorization. Clinimetrics consists in asking such simple questions as "How are you?" and then in following up with additional questions as needed to come to a diagnosis. Clinimetrics relies on clinical wisdom rather than on algorithms for making judgments—and it always includes the patient's own point of view.

The Morris et al (2001) team used the Clinical Dementia Rating (CDR) (Berg, 1988), which is basically a semi-structured interview that relies on clinical judgment, not psychometric cut-off scores, in the process of formulating a simple and easy to understand summary score. Virginia Apgar's scale was also eminently simple (far simpler than the CDR) and informed by clinical judgment of the important factors to include. Heeding Feinstein's (1999) call for greater simplicity and clinical cogency of rating scales in the neuropsychology of aging would mean developing an Apgar rating system for the aging mind. The expectation at the outset of this study was that the MAC-Q memory complaints questionaire could perhaps function in such a way; but the MAC-Q is not suitable because it samples only a single domain, the subjective,

and of course the objective is also important. This study showed however that the MAC-Q is useful as a summary score of memory worries. Since those worries do not necessarily reflect cognitive functioning, they remain an important part of the total clinical picture (important because of the possibilities they open up for further enquiry by the clinician), but not the complete picture.

At present the only widely used and understood summary score is the MMSE, which is actually little informative with regards to mild deficits. Feinstein (1999) described six principles embodied in Apgar's (1953) scale, which could be used to develop an Agpar for MCI: they were selection of variables based on clinical experience and judgment; evenly weighted variables (which provides the virtue of simplicity of use and scoring); heterogeneity of variables; ease of usage; face validity; and source of observations that includes the essential one of the subject or patient concerned (this is one Feinstein [1999] ascribes to Agpar's spirit of doing things, since she of course did not query the babies on how they felt).

The ADAS-cog takes about ten minutes to complete and it was criterion-validated within this small study for subjects with much lower scores than are generally studied (Appendix 5). The ADAS-cog is perhaps the closest "Agpar for MCI" presently being used, being deficient only in heterogeneity of variables (e.g. no psychomotor or attention measures) and in not incorporating the patient's subjective judgment in the

clinical picture. Thus development of an adaptation or extension of the ADAS-cog (incorporating some measure of the subject's or patient's own point of view) is thus a possible further development of the present work.

The CDR scale (Berg, 1988) as applied by Morris and collaborators (2001) is of proven usefulness in detecting incipient AD, but as mentioned above the process of determining a CDR score may be too laborious for primary care practice, where simple and quick tests are needed (Petersen et al, 2001). Conceivably, subjective memory complaints alone could trigger not only a search by the primary care physician for a possible underlying depressive disorder or excessive stress in the patient's life, but also a referal to neuropsychology or to a memory clinic for the more in-depth clinical interview needed for the CDR.

Clinimetrics and psychometrics are complementary, not opposites (Feinstein, 1999). Both clinical and experimental neuropsychology or medicine make use of both approaches. Only the controlled, rule-bound procedures used in this study could have yielded such specialized information as the deficit in recall of negative information by the MemDep group. Such findings can help define and guide the choice of cogent variables to include in forming a clinical judgment. Clinical judgment in turn can help guide and inspire experimental research. In this sense the experimental and clinical domains are like mutual tributaries bringing the fresh water of new ideas to each other.

10. Conclusion

Tuokko and Frerichs (2000) have pointed out that although most clinical neuropsychologists use both individual and normative comparison standards in cognitive assessment, such a dual approach has yet to find its way into research. Using the MAC-Q or some other means of assessing subjective memory complaints could be a way of integrating both approaches: it is an objective record of a person's own comparison standard.

Lamberty and Bieliauskas (1993) defined five broad categories of cognitive and affective presentations in seniors: normal; depressed with no cognitive deficits; depressed with motor-related cognitive deficits; depressed with broad cognitive deficits; and not depressed with broad cognitive deficits. It would seem that given the ubiquity of subjective memory complaints in seniors, and the present findings relating complaints to subtle cognitive impairments, depression, high traumatic symptomatology or some combination thereof, a sixth category should be considered: normal with memory complaints. The normal should perhaps be put in quotation marks to signify the unknown demarcation between the earliest signs of AD and normal cognitive aging.

A recent report on predictors of progression to AD is of some relevance. Subjects with mild cognitive impairment were tested, followed up and retested two years later in

four cognitive areas: language, attention, motor visuospatial function, and verbal fluency (Bozoki, Giordani, Heidebrink, Berent and Foster, 2001). Nearly half of those who at initial evaluation had at least one cognitive impairment along with memory impairment progressed to AD by the two-year follow up evaluation, compared to only 6% of those who at baseline had memory impairments only. Motor visuospatial function (Block Design) was the most frequent abnormality apart from memory in those who progressed to AD. Thus, in view of the impaired declarative memory of the Memory Complaints (in split-median tests) group compared to the No Complaints group, together with the impaired Visual Reproduction performances of the Mem group compared to the MemDep and Control groups, for future and continued evaluation of these same subjects, the inclusion of a Block Design appears warranted, along with more frequent (every 10 minute) in-home salivary cortisol sampling, or in-laboratory venipuncture sampling. In addition, because of known sex differences in the relation of cortisol and cognition (Seeman, McEwen, Singer, Albert and Rowe, 1997) it seems indicated to include sex as a factor in follow-up studies: not studying either men or women, but comparing equivalent numbers of each.

The clinical significance of subjective memory complaints as a sign of cognitive at-risk status (AD or other dementing process) is doubtful. What is less doubtful is that memory complaints are often associated with depression and high levels of stress and possibly post-traumatic stress disorder. A careful history taking may reveal recent or

distant trauma which could be helped with psychotherapy, social work, and appropriate anti-depressant therapy. Subjective memory complaints may perhaps best be conceptualized as a marker of "psychoneuroendocrine frailty," showing slightly elevated cortisol secretion, and slightly lowered cognitive performance, and—when combined with depression—greatly increased stress levels and traumatic symptomatology.

For the purposes of research, memory complaints provide important information about the subject's cognitive self-evaluation, and can in some cases sound an early warning of declining cognitive capacities, increasing basal cortisol, and eventually, possibly, Alzheimer Disease. For the purposes of clinical care, memory complaints are useful as indicators of mental frailty: a perceived, and disturbing gap between the challenges of everyday life and available cognitive resources. This frailty may be due to depression, post-traumatic stress syndrome, or incipient AD. In the short term however memory complaints do not necessarily indicate incipient AD. Their presence in an elderly patient should motivate the clinician to search for an underlying psychological or cognitive cause, and to follow-up regularly in the long term. Memory complaints do not point in one direction only but in many possible directions. In taking heed of them the dignity and sovereignty of the subjective point of view is honored.

Acknowledgements

Thanks to Dr. Serge Gauthier for the French translation of the ADAS-cog.

Note

The dataset used in this project can be viewed by the members of the thesis committee on the Internet at http://www.yahoo.com/r/bc. The user name is: springafternoon. The password is: stendhal. The file is named: PhD Data Set. The key to the codes is named: Key. The files will be maintained at this site until 1 February 2002.

11. Appendix 1: Neutral and negative valence story

STORY

Protagonist: Male Emotional Valence Material is in Bold

Bob ( ) was a math ( ) teacher ( ) in a highschool ( ). He has lbeen living with his brother ( ) since **his wife's death ( )**. Six months ago ( ), he decided to retire ( ) because he found it more and more difficult **to get used to new methods of teaching ( )**. He laughs when he says that he is not from **the computer generation ( )**. For example he says, it's his **grandchildren ( )** who taught him how to use **the bank machine (ATM) ( )** and his VCR ( ).

Last week ( ), he went to his doctor ( ) at the General Hospital(because he thought he **might have Alzheimer's disease ( )**. He was worried because lately, he forgot ( ) the birthday ( ) of his **best friend( )** Tom( ) and often forgot to lock up his front door( ). He also noticed that he sometimes forgets to take his **blood pressure medication ( )**. His doctor told him that it was probably just **stress ( ) and fatigue ( )**. He was very relieved( ).

RESULTS SUMMARY

	Immediate Recall	Delayed Recall	Recognition
Total	/26	/26	/10
Emotional	/13	/13	/ 5
Neutral	/13	/13	/ 5

**(emotional questions)

I. What is the man's name ?
A) Marc
B) Bruce
C) Bob
D) Ronald

2. Where

does he live?* *
A) in a home for elderly people
B) at his brother's
C)in his own house
D) with his daughter

What did he do for a living?
A) he was an insurance broker
B) he was a doctor
C) he was an english teacher
D) he was a math teacher

4. Why has he retired?* *
A) because he did not have enought discipline anymore
B) because he did not have a computer
C) because he could not get use to the new methods of teaching
D) because he was tired

5. Who showed him how to use the ATM?
A) his grandchildren
B) the cashier
C) his sister
D) his students

6. Where did he go last week?

1. his daughter's
2. his doctor
3. the garage
4. the bank

7. Where was that?
A) in Quebec city
B) on Sherbrooke street
C) at the Jewish General Hospital
D) at the General Hospital

8. Why?-*
A) he was worried he might have Alzheimer
disease
B) because he had an appointment
C) because his wife had just passed away
D) because he broke his hip

9. What has he forgotten?**
A) his medicare card
B) where her doctor's office was
Q the date of his wedding
D) his best-friend's b-day

10. What did his doctor tell him?**
A) it was probably just stress and fatigue
B) it might be a tumor
C) that he had to make more tests
D) that he should better watch his diet.

12. Appendix 2: Consent forms

Centre de recherche du Centre hospitalier Côte-des-Neiges

4565, chemin do la Reine-Marie. Montr6al, Qc, H3W 1W5. T61. : (5 14) 340-3540. FAX : (5 14) 340-3548

CONSENT TO PARTICIPATE IN A RESEARCH STUDY

NEUROPSYCHOLOGICAL EVALUATION

I,_____ consent to participate in the following research study, in the conditions described below.

TITLE OF PROJECT: Psychoneuroendocrine Mechanisms of Late-Onset Depression, Dementia and Normal Aging.

RESPONSABLE(S): Sonia Lupien, Ph.D.

GOAL OF THE STUDY:

Dr. Sonia Lupien is conducting a research study to find out more about the effects of stress (daily hassles, loss of friends and/or spouse etc.) on the occurrence of depressive symptoms and/or mild memory loss during human aging. Recent reports in the scientific literature indicate that the hormonal response to stress (which is related to the secretion of a specific hormone called 'cortisol' can have negative effects (if it is secreted for a long period of time, due to stress or other factors). on memory and emotion in people. These negative effects on memory and emotion are related to an impairment in the functioning of a structure in the brain that plays a significant role in memory and emotion. My participation in this study may help scientists to understand how and to what extent this hormone is related to mild memory loss and/or depressive symptoms during aging.

NATURE OF MY PARTICIPATION:

If I agree to participate, I will be asked to participate in a neuropsychological assessment lasting approximately 60 minutes, which will measure my abilities for memory and attention. This testing will I I be performed using a Macintosh computer and I will receive clear instructions as to the different tasks I will be asked to perform. There are no adverse effects related to taking these tests, although I may feel tired after the 60 minute evaluation. However, a 10 min. pause will be taken in the middle of the evaluation and I can ask for other pauses if I wish.

BENEFITS OF ALL THESE PROCEDURES :There is no advantage that can result from my participation in this study except that of contributing to a better understanding of the effects of stress on the occurrence of depressive symptoms and/or mild memory loss during human aging. I will be reimbursed for my expenses (meals, travel, etc.) during my participation in the study.

DISADVANTAGES OF PARTICIPATING IN THIS STUDY: There are no direct

disadvantages that can result from my participation in this study. However, I will have to retrain from smoking or consuming caffeine containing foods and beverages (coffee, tea, cola drinks, chocolate candy, and cocoa) one hour before taking saliva samples once a month at my home, which could create some frustration. Finally, due to the frequency of saliva sampling, I could be tired.

RISK :It is clear that my participation to this research project does not imply any medical risk. It is also clear that my participation to this research study will have no effect on any treatment that I am receiving or may receive in the future at the Montreal Geriatric Institute.

INFORMATION ON THE PROJECT : Researchers or their assistants will have to answer, at my satisfaction, any questions I might have concerning this research project.WITHDRAWAL OF PARTICIPATION : It is clear that my participation to this research study is entirely voluntary; it is also clear that I can, at any moment, stop my participation in. the research study. If this happens, I could ask that data concerning my participation be destroyed.

ACCESS TO FILE : I accept that the researchers responsible for this project have access to my medical file only for the purpose of establishing any concordance between the results of this research project and my health status.

AUTHORIZATION TO TRANSMIT RESULTS : I authorize the researchers responsible for the research study to transmit the results of my participation to my treating physician if this is pertinent

YES [ ] NO [ ]

Name and address of treating physician :

CONFIDENTIALITY

All research records will be kept in a locked filing cabinet accessible only to the principal investigators or their designated assistants will have the key. Moreover, data will be coded in order to prevent any assistant from making a link between a patient's name and test results. If a doctor other than my treating physician wishes to have access to these data, this person will obtain them only after I have given a signed authorization. I will receive, if I desire, some feedback on the results of the study, when the study is finished. Data on this study will be kept in a file for a period of 20 years and only the researchers of this study will have access to this file.

QUESTIONS OR RESEARCH RELATED PROBLEMS

Should I have any other questions or research related problems, I may reach Dr. Lupien at (5[redacted]or [redacted]

If you have any questions about your rights as a patient, or as a research subject, you can phone the person responsible for research related complaints at the Montreal Geriatric Institute, Dr. Céline Crowe at 3 4 0 - 3513.

I declare having read and understood the terms of the present consent form

Signature of Participant

Signed in on _____ 19__

Signature of Witness

I _____certifiy (a) having explained to the signatory the terms of the present consent form; (b)
having answered to the questions that he/she have asked regarding the research project and (c) having
clearly indicated to him/her that he/she is free to stop his/her participation in the study at any time.

Signature of Responsible of project director or representative

Signed in on _____ 19__

The persons responsible for this project can be reached at the Research Center of the Montreal Geriatric Hospital, 4565, Queen Mary, Montreal, Quebec, Canada, H3W-lW5; tel : (514) 340-3540; fax (51 4) 340-3548.

Centre de recherche du
Centre hospitalier Côte-des-Neiges
4585 chemin de la Reine-Marie Montreal Qc, H3W I 1W5

CONSENT TO PARTICIPATE IN A RESEARCH STUDY
SALIVA SAMPLING

I, consent to participate in the following resea.rch study, in the conditions described below.

TITLE OF PROJECT: Psychoneuroendocrine Mechanisms of Late-Onset Depression, Dementia and Normai Aging.

RESPONSABLE(S) Sonia Lupien, Ph.D.

GOAL OF THE STUDY:

Dr. Sonia Lupien is conducting a research study to find out more about the effects of stress (daily hassles, loss of friends and/or spouse etc.) on the occurrence of depressive symptoms and/or mild memory loss during human aging. Recent reports in the scientific literature indicate that the hormonal response to stress (which is related to the secretion of a specific hormone called "cortisol" can have negative effects (if it is secreted for a long period of time, due to stress or other factors) on memory and emotion in aged people. These negative effects on memory and emotion are related to an impairment in the functioning of a structure in the brain that plays a significant role In memory and emotion. My participation in this study may help scientists to understand how and to what extent this hormone is related to mild memory loss and/or depressive symptoms during aging.

NATURE OF MY PARTICIPATION:

If I agree to Participate, Dr. Lupien or her associate will show me how to take a sample of my saliva using a little filter paper. Dr. Lupien will use my saliva in order to Measure the level of my stress hormones every month of the incoming year. This will help the researchers assess whether my stress hormones change throughout the year. In order to take a sample of my saliva, I will use a little filter paper of 3 inches. This little filter paper has 2 parts : a handling part (1 inch) that I will be able to handle and on which I will write my initials, time and date of the sampling, and a saliva sampling part (2 inches) that I will place in my mouth until the filter paper is totally wet with my saliva. Once the filter paper is totally wet with my saliva, I will gently take it out

of my mouth and place it in a plastic bag that Dr. Lupien or her associate will provide me, and I will let the paper air dry before closing the plastic bag.

I will be asked to provide 5 saliva samples per day for three consecutive days. I understand that I can decide to stop taking saliva samples at any time during the study without any jeopardy to the medical care I am receiving or might receive at the Hospital.

During the day that I take the saliva samples, I will be asked to take a saliva sample at the time of awakening, and 30 minutes later (for example, If I wake up at 8h00 am, I will be asked to take a saliva sample at 8h00 am and at 8H30 am I will then take 2 additional saliva samples at 2h00 and 4h00 pm the same day. Although I will be asked to refrain from smoking or consuming caffeine containing foods and beverages (common caffeine containing foods and beverages are coffee, tea, cola drinks, chocolate candy, and cocoa) after awakening and 45 minutes later, I will be free to smoke or consume caffeine containing foods and beverages for the rest of the day (i.e. I will not have to retrain from smoking or consuming caffeine containing beverages before taking the 2h00 and 4h00 pm saliva samples).

During the same day, I will be asked to answer short questionnaires containing questions about the stress that I have experienced in the last month, and the daily hassles that may have disturbed me within the last month.

BENEFITS OF ALL THESE PROCEDURES

There is no advantage that can result from my participation in this study except that of contributing to a better understanding of the effects of stress on the occurrence of depressive symptoms and/or mild Memory loss during human aging. I will be reimbursed for my expenses (meals, travel, etc.) during my participation in the study.

DISADVANTAGES OF PARTICIPATING IN THIS STUDY

There are no direct disadvantages that can result from my participation in this study. However, I will have to refrain from smoking or consuming caffeine containing foods and beverages (coffee, tea, cola drinks, chocolate candy, and cocoa) one hour before taking saliva samples once a month at my home, which could create some frustration. Finally, due to the frequency of saliva sampling, I could be bred.

RISK

It is clear that my participation to this research project does not imply any medical risk. it is also clear that my participation to this research study will have no effect on any treatment that I am receiving or may receive in the future at the Montreal Geriatric Institute.

INFORMATION ON THE PROJECT : Researchers or their assistants will have to answer, at my satisfaction, any questions I might have concerning this research project.

WITHDRAWAL OF PARTICIPATION : It is clear that my participation to this research study is entirely voluntary; it is also clear that I can, at any moment, stop my participation in the research study, If this happens, I could ask that data concerning my participation be destroyed.

ACCESS TO FILE : I accept that the researchers responsible for this project have access to my medical file only for the purpose of establishing any concordance between the results of this research project and my health status.

AUTHORIZATION TO TRANSMIT RESULTS : I authorize the researchers responsible for the research study to transmit the results of my participation to my treating physician if this is pertinent

YES [ ] NO [ ]

Name and address of treating physician :

CONFIDENTIALITY

All research records will be kept in a locked filing cabinet accessible only to the principal investigators and their designated assistants will possess the key. Moreover, data will be coded in order to prevent any assistant from making a link between a patients name and test results. If a doctor other than my treating physician wishes to have access to these data, this person will obtain them only after I have given a signed authorization. I will receive, if I desire, some feedback on the results of the study, when the study is finished. Data on this study will be kept in a file for a period of 20 years and only the researchers of t th is study will have access to this file.

QUESTIONS OR RESEARCH RELATED PROBLEMS

Should I have any other questions or research related problems, I may reach Dr. Lupien at (514) 762 - 3048 or (514) 591-9600.

It you have any questions about your rights as a patient, or as a research subject, you can phone the person responsible for research related complaints at the Montreal Geriatric Institute, Dr. Céline Crowe at 3 4 0 - 3513.

I declare having read and understood the terms of the present consent form

Signature of Participant _____

Signed in_____, on 19___

Signature of Witness_____

Signed in_____, on 19___

I _____ certify (a) having explained to the signatory the terms of the present consent form; (b) having answered to the questions that he/she have asked regarding the research' project and (c) having clearly indicated to him/her that he/she is free to stop his/her participation in the study at any time.

Signature of Responsible of project or representative

_____

The persons responsible for this project can be reached at the Research Center of the Montreal Geriatric Hospital, 4565 Queen Mary, Montreal, Quebec, Canada, H3W-1W5; tel : (514) 340-3540; fax (51 4) 340-3548.

(TO BE COMPLETED IN 3 COPIES)

Version of December 4, 1997

13. Appendix 3: Statement on work done

It was my idea to study a group of volunteers with subjective memory complaints with the goal of trying to determine whether the validity of those memory complaints was altered by the presence of concomittant depressive symptoms. It was my idea to situate this work in the context of the evolving definitions of mild cognitive impairment and the search for reliable early markers of incipient Alzheimer Disease. Dr. Lupien directed my reading and research within the framework of the literature on

stress and stress hormones.. Dr. Lupien also advised me about various methodological aspects, including testing instruments and outcome measures. Most instruments used were the ones she and her collaborators have been using in various longitudinal studies of memory, stress and aging. I chose to include the 3MS and the ADAS-cog along with the commonly used MMSE.

14. Appendix 4: Publication plan

The publication plan has not been finalized, but future submissions to peer-reviewed journals will include papers on the following subjects:

- Review of the use of the term "cognitive impairment, no dementia" (CIND). In recent proposals, CIND would contain MCI, which would constitute a significant deficit in any cognitive domain, as well as MUML, which would specifically refer to a memory deficit. CIND includes benign, non-progressive cases, other non-aging-related causes of cognitive impairment (e.g. mental retardation, alcoholism), and also incipient AD cases. All unsuspected AD cases within the CIND category could eventually be recategorized as either MUML (for isolated memory impairments) or MCI (for other cognitive deficits). This would better reflect the many ways that AD could progress. The challenge is developing an operational definition of MCI or MUML that would produce an "incipient AD – enriched" sample while excluding a greater proportion of non-progressive cases.

- Report on subjective memory complaints. Complaints were poorly related to cognitive performance in our cross-sectional study; however this could be improved by selecting a higher cut-off on the MAC-Q, a commonly used memory complaint measure. Memory complaints may perhaps best be conceptualized as markers of "psychoneuroendocrine frailty," signaling the presence of high stress, inadequate coping mechanisms, and in the presence of concomitant depressive symptoms, signs of post-traumatic stress disorder.

- Report on infradian salivary cortisol concentrations in seniors with memory complaints with or without depressive symptoms. Subjects with "pure" memory complaints (without depressive symptoms) failed to show the expected cortisol rise in the first 30 min after awakening. A model comprising cortisol concentrations averaged over three days, cortisol morning rise, traumatic symptoms and verbal fluency accounted for one third of the variance in memory complaints.

- Report on the the impact of traumatic symptoms on subjective memory complaints or depression in elderly humans.

15. Appendix 5: Hippocampal Volumetry

As part of a post-doctoral project by Dr. Jens Pruessner at the Montreal Neurological Institute, all subjects that I evaluated as part of this project were slated to undergo

Magnetic Resonance Imaging of their brains. At the time of this writing, 12 subjects had undergone the procedure. In order to verify the validity of the SADAS-cog as compared to the MMSE and the 3MS, a correlational analysis between scores on those cognitive screening measures and hippocampal volumes was undertaken. The results are shown in **Table 9** by Pearson's *r*, number of subjects, and *p* value.

Table 9. Hippocampal Volumetry

Correlations between Cognitive Screening Instruments and right and left hippocampal (HIP) volumes.

	MMSE	3MS	RIGHT HIP	LEFT HIP
ADAS	-.6195	-.6406	-.7495	-.7489
	(12)	(12)	(12)	(12)
	P= .032	P= .025	P= .005	P= .005
MMSE		.9519	.3351	.3812
		(12)	(12)	(12)
		P= .000	P= .287	P= .221
3MS			.4290	.4898
			(12)	(12)
			P= .164	P= .106

16. Appendix 6: Contribution to the science

A survey of the literature revealed that in the search for a rational definition of an intermediate stage between normal aging and AD, various groups have advanced and tested multiple definitions, which have been proven to be mostly mutually exclusive, and (in the opinion of some investigators) unduly restrictive in their inclusion and

exclusion criteria. My contribution to the advancement of knowledge in neuropsychology in this thesis consists in demonstrating that a simple technique—asking about memory complaints—with light exclusion criteria, could be used to detect a possible psychoneuroendocrine abnormality, consisting of depressive symptoms, post-traumatic stress symptoms, mild hypercortisolism, or mild cognitive deficits compared to people with no memory complaints. Although this thesis provides no information as to whether memory complaints or the slight deficits in various domains found in some subjects presage AD, it does demonstrate that memory complaints are important clinically as signs of an abnormality, and human suffering, in the present.

17. References

Aicken M. (1998). Other methods for adjustment of multiple testing exist. [Letter]. *British Med J* 318: 127.

Almkvist O & Winblad B. (1999). Early diagnosis of Alzheimer dementia based on clinical and biological factors. *Eur Arch Psychiatry Clin Neurosci* 249 Suppl 3: 3-9.

American Association of Clinical Endocrinologists. (2001). ACE consensus conference on guidelines for glycemic control. Internet: http://www.aace.com/pub/press/releases/diabetesconsensuswhitepaper.php

American Psychiatric Association. (1987). Diagnostic and Statistical Manual for Mental Disorders—third edition, revised. Washington, D.C.: American Psychiatric Association.

American Psychiatric Association. (1994). Diagnostic and Statistical Manual for Mental Disorders—fourth edition. Washington, D.C.: American Psychiatric Association.

Anthony JC, LeResche L, Niaz U, von Korff MR & Folstein MF. (1982). Limits of the 'Mini-Mental State' as a screening test for dementia and delirium among hospital patients. *Psychol Med* 12:397-408.

Apgar V. (1953). A proposal for a new method of evaluation of the newborn infant. *Anesth Analg* 32: 260-7.

Arbuckle TY, Gold D & Andres D. (1986). Cognitive functioning of older people in relation to social and personality variables. *Psychology and Aging* 1: 55-62.

Arbuckle TY, Gold DP, Chaikelson JS & Lapidus S. (1994). Measurement of Activity in the Elderly: The Activities Checklist. *Canadian J of Aging* 13: 550-65.

Auer SR, Sclan SGD, Yaffee RA & Reisberg B. (1994). The neglected half of Alzheimer Disease: cognitive and functional concomitants of severe dementia. *J Am Geriatr Soc* 42: 1266-72.

Baddeley A. (1986). Working memory. Oxford: Oxford University Press.

Bazin N. (1991). [Memory, emotional staus and affective valence. Critical review.] [Article in French.] *Encephale* 17: 503-10.

Bedi RP. (1999). Depression: an inability to adapt to one's perceived life distress? *J Affect Disord* 54: 225-34.

Belleville S, Rouleau N & Caza N. (1998). Effect of normal aging on the manipulation of information in working memory. *Memory Cog* 26: 572-83.

Bench CJ, Frackowiak RSJ & Dolan RJ. (1995). Changes in regional cerebral blood flow on recovery from depression. *Psychol Med* 25: 247-51.

Bender R & Lange S. (1999). Multiple test procedures other than Bonferroni's deserve wider use. *British Med J* [Letter]. 318: 600.

Berg L (1988). Clinical Dementia Rating (CDR). *Psycho Bull* 24: 637-39.

Berg L, Hughes CP, Coben LA, Danziger WL, Martin RL & Knesevich J. (1982). Mild senile dementia of Alzheimer Type: research diagnostic criteria, recruitment, and description of a study population. *J Neurol Neurosurg Psychiatry* 45: 962-68.

Binetti G, Magni E, Padovani A, Cappa SF, Bianchetti A & Trabucchi M. (1996). Executive dysfunction in early Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 60: 91-3.

Blackford RC & La Rue A (1989). Criteria for diagnosing age-associated memory impairment: proposed improvement from the field *Dev Neuropsychol* 5: 295-306.

Blass JP. (1996). Age-Associated Memory Impairment and Alzheimer's Disease. (Editorial). *J American Geriatri Soc* 44:209-11.

Board F, Wadeson R & Persky H. (1957). Depressive affect and endorine function: blood levels of adrena cortex and thyroid hormone in patients suffering from depressive reaction. *Arch Neurol Psychiatry* 78: 612-20.

Born J, Hitzler V, Pietrowsky R, Pauschinger P & Fehm HL. (1989) Influences of cortisol on auditory evoked potentials (AEPs) and mood in humans. *Neuropsychobiol* 20: 145-51.

Bowen J, Teri L, Kukull W, McCormick W, McCurry SM & Larson EB. (1997). Progression to dementia in patients with isolated memory loss. *The Lancet* 349: 763-65.

Bozoki A, Giordani B, Heidebrink JL, Berent S & Foster NL. (2001). Mild cognitive impairments predict dementia in nondemented elderly patients with memory loss. *Arch Neurol* 58: 411-6

Braak H & Braak E. (1991). Neuropathological staging of Alzheimer's disease. *Acta Neuropathol (Berl)* 82: 239-59.

Braak H & Braak E. (1992). The human entorhinal cortex: normal morphology and lamina-specific pathology in various diseases. *Neurosci Res* 15: 6-31.

Braak H & Braak E. (1999). Neuropathological stages of Alzheimer Disease. *In*: De Leon MJ. (Ed.) An Atlas of Alzheimer's Disease. New York: The Parthenon Publishing Group. pp. 57-74.

Bremmer JD. (1999). Alterations in brain structure and function associated with post-traumatic stress disorder. *Semin Clin Neuropsychiatry* 4: 249-55.

Briere J & Runtz M. (1989). The Trauma Symptom Checklist (TSC-33): Early Data on a New Scale. *J Interpersonal Violence* 4: 151-63.

Broca P. (1878). Anatomie comparée de circonvolutions cérébrales. Le grand lobe limbique et la scissure limbique dans le série des mammifères. *Rev Anthropol* 1: 385-498.

Buckner RL, Kelley EM & Peterson SE. (1999). Frontal cortex contributes to human memory formation. *Nature neurosci* 2: 311-3.

Burke WJ, Roccaforte WH, Wengel SP, Conley DM & Potter JF. (1995). The reliability and validity of the Geriatric Depression Rating Scale administered by telephone. *J Am Geriatr Soc* 43: 674-9.

Cahil L & McGaugh JL. (1996). Modulation of memory storage. *Curr Opin Neurobiol* 6: 237-42.

Calev A. (1996). Affect and memory in depression: evidence of better delayed recall of positive than negative affect words. *Psychopathology* 29: 71-6.

Canadian Study of Health and Aging Work Group (1994). Canadian Study of Health and Aging: study methods and prevalence of dementia. *Can Med Assoc J* 150: 899-913.

Caroll BJ, Curtis GC & Mendel J. (1976). Neuroendocrine regulation in depression. I. Limbic system-adrenocortical dysfunciton. *Arch Gen Psychiat* 33:1039.

Carroll BJ, Curtis GC & Mendels J. (1976). Cerebrospinal fluid and plasma free cortisol concentrations in depression. *Psychol Med* 6:235-44

Cermak JM, Holler T, Jackson DA & Blusztajn JK. (1998). Prenatal availability of choline modifies development of the hippocampal cholinergic system. *FASEB J* 12:349-57

Churchland PS. (1988). Neurophilosophy: Toward a unified science of mind and brain. Cambridge, Mass: MIT Press.

Cohen J & Cohen P. (1983). Applied Multiple Regression / Correlation Analysis for the

Behavioral Sciences. 2nd Ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; pp. 414-17. Cited in: Robinson-Whelen & Storandt, 1992.

Cohen J. (1983). The cost of dichotomization. *Applied Psychological Measurement* 7: 249–253.

Cohen J. (1988). Statistical Power Analysis for the Behavioral Sciences. Hillsdale, New Jersey: Laurence Erlbaum Associates, Publishers.

Cohen S, Karmarck T & Mermelstein R. (1983). A Global Measure of Perceived Stress. *J Health and Social Behav* 24: 385-96.

Cohen S, Tyrell DA & Smith AP. (1993). Negative life events, perceived stress, negative affect, and susceptibility to the common cold. *J Personality and Social Psychol* 64: 131-40.

Craik FIM. (1994). Memory Changes in Normal Aging. *Current Directions in Psychol Sci* 3:155-63.

Cremin MC. (1992). Feeling old versus being old: views of troubled aging. *Soc Sci Med* 34: 1305-15.

Crook T, Bartus RT, Ferris SH, Whitehouse P, Cohen GD & Gershon SG. (1986). Age-Associated Memory Impairment: Proposed Diagnostic Criteria and Measures of Clinical Change. Report of a National Institutes of Health Working Group. *Dev Neuropsychol* 2: 221-76.

Crook TH, Feher E & Larrabee GJ. (1992). Assessment of Memory Complaints in Age-Associated Memory Impairment: The MAC-Q. *International Psychogeriatrics* 4: 165-76.

Cummings J & Khachaturian Z. (1996). Definitions and diagnostic criteria. In: Gauthier S. (Ed.) Clinical Diagnosis and Management of Alzheimer's Disease. Toronto: Butterworth-Heinemann. pp. 3-13.

Danion JM, Kauffmann-Muller F, Grange D, Zimmermann MA & Greth P.J. (1995). Affective valence of words, explicit and implicit memory in clinical depression. Affect Disord 34: 227-34

Danion JM, Willard-Schroeder DW, Zimmermann MA, Grangé DM, Schlienger J-L & Singer L. (1991). Explicit memory and repetition priming in depression. *Arch Gen Psychiatry* 48: 707-11.

Davis KL; Davis BM; Greenwald BS; Mohs RC; Mathe AA; Johns CA & Horvath TB (1986). Cortisol and Alzheimer's disease, I: Basal studies. *Am J Psychiatr* 143: 300-5.

De Marco RR. (2000). The epidemiology of major depression: implications of occurrence, recurrence,and stress in a Canadian community sample. *Can J Psychiatry* 45:67-74.

DeLeon M, McRae T, Tsai J, Heorge A, Marcus D, Freedman M, Wolf A & McEwen B. (1988). Abnormal cortisol response in Alzheimer's Disease linked to hippocampal atrophy. *Lancet* 2: 391-2.

DeLeon MJ, Convit A, Tarshish C, DeSanti S & Bobinski M. (1999). MRI studies of the

hippocampal formation: contributions to the early diagnosis of Alzheimer Disease. *In*: De Leon MJ. (Ed.) An Atlas of Alzheimer's Disease. New York: The Parthenon Publishing Group. pp. 31-55.

Derogatis LR & Fleming MP. (1997). The Derogatis Stress Profile: A Theory Driven Approach to Stress Management. In: Zalaquett CP & Wood RJ (Eds.). Evaluating Stress: A Book of Resources. Lanham, MD, USA: Scarecrow Press, Inc. pp. 113-140.

Derogatis LR, Lipman RS & Covi L. (1973). SCL-90: An Outpatient Psychiatric Rating Scale—Preliminary Report. *Psychopharm Bull* 9:13-28.

Derouesne C. (2000). [Memory and affect]. [Article in French]. *Rev Neurol (Paris)* 156: 732-7

Deshmuck VD & Deshmuck SV. (1990). Stress-adaptation failure hypothesis of Alzheimer's Disease. *Med Hypotheses* 32: 293-5.

Detoledo-Morrell L, Sullivan MP, Morrell F, Wilson RS, Bennett DA & Spencer S. (1997). Alzheimer's Disease: In Vivo Detection of Differential Vulnerability of Brain Regions. *Neurobiol Aging* 18: 463-68.

Dobkin PL, Pihl RO & Breault C. (1991). Validation of the Derogatis Stress Profile using laboratory and real world data. *Psychotherapy and Psychosomatics* 56: 185-96.

Dolan RJ; Bench CJ; Brown RG; Scott LC & Frackowiak RS. (1994) Neuropsychological dysfunction in depression: the relationship to regional cerebral blood flow. *Psychol Med* 4: 849-57

Dresser R. (2000). Weighing the benefits of new Alzheimer's treatments. [Letter.] *Science* 289: 86.

Ebly EM, Hogan DB & Parhad IM. (1995). Cognitive impairment in the nondemented elderly. Results from the Canadian Study of Health and Aging. *Arch Neurol* 52: 612-619.

Elias MF, Beiser A, Wolf PA, Au R, White RF & D'Agostino RB. (2000). The preclinical phase of alzheimer disease: a 22-years prospective sutdy of the Framingham cohort. *Arch Neurol* 57: 808-13.

Elias MF, Wolf PA, D'Agostino RB, Cobb J & White LR. Untreated blood pressure level is inversely related to cognitive functioning: the Framingham Study. *Am J of Epidemiol* 138: 353-64.

Ellis HC, Thomas RL & Rodriguez IA. (1984). Emotional mood states and memory: elaborative encoding, semantic processing, and cognitive effort. *J Exp Psychol Learn Mem Cogn* 10: 470-82.

Ellis HC. (1985). Depressed mood states and memory: a ressource allocation model. *Soc Sci Newsletter* Fall Issue: 1-3

Erkinjuntti T, Østbye T, Steenhuis R & Hachinski V. (1997). The effect of different diagnostic criteria on the prevalence of dementia. *New England J Med* 337(23): 1667-74.

Eschwege E (2000). [Epidemiology of type II diabetes, diagnosis, prevalence, risk factors,

complications.] (Article in French). Arch Mal Coeur Vaiss 93 Spec No 4:13-7

Esteban-Santillan C, Praditsuwan R, Ueda H & Geldmacher DS. (1998). Clock drawing test in very mild Alzheimer's Disease. *J Am Geriatr Soc* 46: 1266-9.

Estes WK. (1978). The information-processing approach to cognition: a confluence of metaphors and methods. In: Estes WK. (Ed.). Handbook of Learning and Cognitive Processes, Vol 5, Human Information Processing. New York: Wiley; pp. 1-18.

Falgatter AJ, Roesler M, Sitzmann L, Heidrich A, Mueller TJ & Strik W. (1997). Loss of functional hemispheric asymetry in Alzheimer's dementia assessed with near-infrared spectroscopy. *Brain Res Cogn Brain Res* 6: 67-72.

Feinstein AR. (1987). Clinimetrics. New Haven, CT: Yale University Press.

Feinstein AR. (1990). The inadequacy of binary models for the clincial reality of three-zone diagnostic decisions. *J Clin Epidemiol* 43: 109-13.

Feinstein AR. (1999). Multi-item "instruments" vs Virginia Apgar's principles of clinimetrics. *Arch Intern Med* 159: 125-8.

Ferris SH, Crook T, Sathananthan G & Gershon S. (1976). Reaction time as a diagnostic measure of cognitive impairment in senility. *J Am Geriat Soci* 24: 529-33.

Flicker C, Ferris SH & Reisberg B. (1991). Mild cognitive impairment in the elderly: Predictors of Dementia. *Neurology* 41: 1006-9.

Flicker C. (1988). Neuropsychological evaluation of treatment effects in the elderly: a critique of tests in current use. *Psychopharmcol Bul* 24: 535-55.

Folstein MF, Fostein SE & McHugh PR. (1975). "Mini-Mental State." A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12: 189-98.

Foster JK, Behrmann M & Stuss DT. (1995). Aging and visual search: Generalized cognitive slowing or selective deficit in attention? *Aging Cog* 2: 279-99.

Freter S, Bergman H, Gold S, Chertkow H & Clarfield AM. (1998). Prevalence of potentially reversible dementias and actual reversibility in a memory clinic cohort. *Can Med Assoc J* 159: 657-62.

Gaffan D & Parker A. (1996) Interaction of perirhinal cortex with the fornix-fimbria: memory for objects and "object-in-place"memory. *J Neurosci* 16: 5864-9.

Gauthier S, Thal LJ & Rossor M. (1996). The future diagnosis and treatment of Alzheimer's disease. In: Gauthier S. (Ed.) *Clinical Diagnosis and Management of Alzheimer's Disease.* London: Martin Dunitz. pp.359-65.

Gauthier S. (1999). Do we have a treatment for Alzheimer Disease? Yes. *Arch Neurol* 56: 738-9.

Gazzaley AH, Thakker MM, Hof PR & Morrison JH. (1997). Preserved Number of Entorhinal Cortex Layer II Neurons in Aged Macaque Monkeys. *Neurobiol Aging* 18: 549-53.

Geda YE & Petersen RC. (2001). Clinical trials in mild cognitive impairment. In: Gauthier S. (Ed.). Alzheimer's disease and related disorders annual 2001. London, UK: Martin Dunitz Ltd. Pp

Geertz C. (1973/1977). The Interpretation of Cultures: Selected Essays. New York: Basic Books.

Geffen G, Bate A, Wright M, Rozenbilds U, Geffen L. (1993). A comparison of cognitive impairments in dementia of the Alzheimer type and depression in the elderly. *Dementia* 4 : 294-300.

George MS, Ketter TA & Post RM. (1994). Prefrontal cortex dysfunction in clinical depression. *Depression* 2: 59-72.

Geula C. (1998). Abnormalities of neural circuitry in Alzheimer's disease: hippocampus and cortical cholinergic innervation. *Neurology* 51(Suppl.1): S18-S29.

Gilewski MJ, Zelinski EM & Schaie KW. (1990). The memory functioning questionnaire for assessment of memory complaints in adulthood and old age. *Psychol Aging* 5: 482-90.

Goldman J & Côté L. (1991). Aging of the Brain: Dementia of the Alzheimer's Type. In: Kandel ER, Schwartz JH & Jessel TM (Eds.) Principles of Neural Science. (Third Edition). New York: Elsevier.

Gómez-Isla T, Price JL, McKeel Jr DW, Morris JC, Growdon JH & Hyman BT. (1996). Profound Loss of Layer II Entorhinal Cortex Neurons Occurs in Very Mild Alzheimer's Disease. *J Neurosci* 16: 4491-4500.

Goya L, Rivero F & Pascual-Leone AM. (1995). Glucocorticoids, stress and aging. *In:* Timiras PS, Quay WB, Vernadakis A. (Eds.). Hormones and aging. Boca Raton, Florida: CRC Press, Inc.

Graf P & Schacter DL. (1985). Implicit and explicit memory for new associations in normal and amnesic patients. *J Exp Psychol: Learning Mem Cog* 11: 501-18.

Graham JE, Minitski AB, Mogilner AJ, Gauvreau D, Rockwood K. (1996). Symptoms and signs in dementia: synergy and antagonism. *Dementia* 7:331-5.

Graham JE, Rockwood K, Beattie BL, Eastwood R, Gauthier S, Tuokko H & McDowell I. (1997). Prevalence and severity of cognitive impairment with and without dementia in an elderly population. *The Lancet* 349: 1793-96.

Gray JA (1995). The contents of Consciousness: a Neuropsychological conjecture. *Behavioral and Brain Sciences* 18: 659-722.

Greenwald BS, Mathé AA, Mohs RC, Levy MI, Johns CA & Davis KL. (1986). Cortisol and Alzheimer's Disease, II: Dexametahson Suppression, Dementia Severity, and Affective Symptoms. *Am J Psychiatry* 143: 442-446.

Grut M, Jorm AG, Fratiglioni L, Forsell Y, Viitaen M & Winblad B. (1993). Memory complaints of elderly people in a population survey: variation according to dementia stage and depression. *J Am Geriat Soc* 41(12): 1295-1300.

Gussekloo J, Westendorp RGJ, Remarque EJ, Lagaay AM, Heeren TJ & Knook DL. (1997). Impact of mild cognitive impairment on survival in very elderly people: cohort study. *Brit Med J* 315:1053-4.

Gust DA, Gordon TP, Hambright MK & Wilson ME. (1993). Relationship between social factors and pituitary-adrenocortical activity in female rhesus monkeys (Macaca mulatta). *Horm Behav* 27: 318-31.

Habib M & Poncet M. (1991). Conduite diagnostique face à un sundrome démentiel. In: Habib M, Joanette Y & Puel M. (Eds.) Démences et syndrome démentiels. Paris: Edisem, 1991.

Haist F, Gore JB & Mao H. (2001). Consolidation of human memory over decades revealed by functional magnetic resonance imaging. *Nature neurosci* 4: 1139-45.

Halbreich U, Asnis GM, Zumoff B, Nathan NS & Shindledecker R (1984). Effect of age on cortisol secretion in depressives and normals. *Psychiat Res* 13: 221-9.

Hammen C. (1992). Life events and depression: the plot thickens. *Am J Communitry Psychol* 20: 179-93.

Hanley J.A. (1989). Receiver operating characteristic (ROC) methodology: the state of the art. *Crit Rev Diagn Imaging* 29:307-35

Hänninen T, Hallikainen M, Koivisto K, Helkala E-L, Reinikainen KJ, Soininen H, Mykkänen L, Laakso M, Pyörälä K & Riekkinen P. (1995). A Follow-Up Study of Age-Associated Memory Impairment: Neuropsychological Predictors of Dementia.. *J Am Geriatr Soc* 43: 1007-1015.

Hänninen T, Hallikainen M, Koivisto K, Partanen P, Laakso M, Reinikainen KJ & Soininen H. (1997). Decline of frontal lobe functions in subjects with age-associated memory impairment. *Neurology* 48: 148-53.

Hart RP, Kwentus JA, Talylor JR & Hamer RM (1988). Productive naming and memory in depression and Alzheimer's type dementia. *Arch Clin Neuropsychol* 3: 313-22.

Heim C & Nemeroff CB. (1999). The impact of early adverse experiences on brain systems involved in the pathophysiology of anxiety and affective disorders. *Biol Psychiatry* 46: 1509-22.

Henkin RI, McClone RE, Daly R & Bartter FC. (1967). Studies on auditory thresholds in normal man and in patients with adrenal cortical insufficiency: the role of adrenal cortical steroids. *J Clin Investi* 46: 429-45.

Hodges JR & Patterson K. (1995). Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomic and diagnostic implications. *Neuropsychologia* 33: 441-59.

Hogan DB & Black SE. (2001). Dissenting opinion on Vitamin E. In: Patterson C, Gauthier S, Bergman H, Cohen C, Feightner JW, Feldman H, Grek A & Hogan DB. (2001). The recognition, assessment and management of dementing disorders: conclusions fro the

Canadian consensus conference on dementia. *Can J Neurol Sci* 28: Suppl. 1: S3-S16.

Huberty CJ & Morris JD. (1989). Multivaraite analysis versus multiple univariate analyses. *Psychol Bull* 105: 302-8.

Ihl R, Frolich L, Dierks T, Martin E-M & Maurer K. (1993). Differential validity of psychometric tests in dementia of the alzheimer's type. *Psychiatr Res* 44:93-106.

Issa AM, Rowe W, Gauthier S & Meaney MJ. (1990). Hypothalamic-pituitary-adrenal activity in aged, cognitively impaired and cognitively unimpaired rats. *J Neurosci* 10: 3247-54.

Izaks GJ, Gussekloo J, Dermout KMT, Heeren TJ & Ligthart GJ. (1995). Three-year follow-up of Mini-mental state examination (MMSE) in octo- and nonagenarians. *Psychol Med* 25: 841-8.

James O & Davies DM. (1987). Assessing social supporty and satisfaction in the elderly: development of a bried assessment instrument, the index of social support. *Int J Geriatric Psychiatry*: 2: 227-33.

Jolles J, Verhey FRJ, Riedel WJ & Houx PJ (1995). Cognitive impairment in elderly people: predisposing factors and implications for experimental drug studies. *Drugs & Aging* 7: 459-79.

Jonker C, Launer LJ, Hooijer C & Lindeboom J. (1996). Memory complaints and memory impairment in older individuals. *J Am Geriatr Soc* 44: 44-9.

Joseph JA, Shukitt-Hale B, Denisova NA, Prior RL, Cao G, Martin A, Taglialatela G, Bickford PC. (1998). Long-term dietary strawberry, spinach, or vitamin E supplementation retards the onset of age-related neuronal signal-transduction and cognitive behavioral deficits. *J Neurosci* 18:8047-55

Jost BC & Grossberg GT. (1996. The evolution of psychiatric symptoms in Alzheimer's disease: a naturalistic study. *J Am Geriatr Soc* (Letter) 44: 1137-8.

Kandel ER, Schwartz JH & Jessel TM. (Eds). (1991). Principles of Neural Science. (Third edition.) New York-Amsterdam-London-Tokyo: Elsevier.

Kandel ER. (1991). Cellular Mechanisms of Learning and the Biological Basis of Individuality. In: Kandel ER, Schwartz JH & Jessel TM. (Eds). (1991). Principles of Neural Science. (Third Edition). New York-Amsterdam-London-Tokyo: Elsevier.

Kaszniak AW. (1986). The neuropsychology of dementia. In: Grant I & Adams KM. (Eds.) Neuropsychological Assessment of Neuropsychiatric Disorders. New York: Oxford University Press.

Kaye JA. (1998). Diagnostic challenges in dementia. *Neurology* 51, Suppl 1: S45-S52.

Kendler SK & Gardner CO, Jr. (1998). Boundaries of major depression: an evaluation of DSM-IV criteria. *Am J Psychiatry* 155: 172-77.

Kertesz A & Mohs RC. (1996). Cognition. In: Gauthier S (Ed.). Clinical Diagnosis and Management of Alzheimer's Disease. London: Martin Dunitz. pp.155-74.

Knopman DS, DeKosky ST. Cummings JL, Chui H, Corey-Bloom J, Relkin N, Small GW, Miller B & Stevens JC. (2001). Practice parameter: diagnosis of dementia (an evidence-based review). Report fo the Quality Standards Subcommitee of the American Academy of Neurology. *Neurology* 56: 1143-53.

Koivisto K, Reinikainen KJ, Hänninen T, Vanhanen M, Helkala E-L, Mykkänen L, Laakso M, Pyörälä K & Riekkinen J. (1995). Prevalence of age-associated memory impairment in a randomly selected population in Finland. *Neurology* 45: 741-47.

Kopelman MD. (1985). Rates of forgetting in alzheimer-type dementia and Korsakoff's syndrome. *Neuropsychologia* 23: 623-38.

Koven MG. (1977). Health of the elderly and use of health services. *Public Health Rep* 92: 9-15.

Kraemer HC (1981). Coping Strategies in Psychiatric Clinical Research. *J Consulting Clin Psychol* 49: 309-19. Reprinted in: Kazdin AE (Ed.) (1992). Methodological Issues & Strategies in Clinical Research. Washington, D.C.: American Psychological Association.

Kraemer HC, Moritz DJ & Yesavage J. (1998). Adjusting mini-mental state examination scores for age and education level to screen for dementia: correcting bias or reducing validity? *Int Psychogeriatr* 10: 43-51.

Kral VA. (1958). Neuro-psychiatric observations in an old peoples homes: studies of memory dysfunction in senescence. J *Gerontol* 1334: 169-76.

Krieger DT, Allen W, Rizzo F, Kireger HP. (1970). Characterization of the normal temporal pattern of plasma corticosteroid levels. *J Clin Endocrinol Metab* 32: 266-84.

Kristjansson B, Hill G, McDowell I & Lindsay J. (1996). Optimal Cut-Points When Screening for More Than One Disease State: An Example from the Canadian Study of Health and Aging. *J Clin Epidemiol* 49:, 1423-28.

Kumar A & Cummings JL. (2001). Depression in neurodegenerative disorders and related conditions. In: Gauthier S & Cummings JL. (Eds.). Alzheimer's disease and related disorders annual 2001. London, UK: Martin Dunitz Ltd. pp 123-141.

Kupfermann I. (1991). Hypothalamus and Limbic System: Peptidergic Neurons, Homeostasis, and Emotional Behavior. In: Kandel ER, Schwartz JH & Jessel TM (Eds.) (1991). Principles of Neural Science. (Third Edition). New York: Elsevier.

Laakso M. (1996). MRI of Hippocampus In Incipient Alzheimer's Disease. University of Kuopio (Finland) Series of Reports Department of Neurology No. 37. Available on the Internet at: http://www.uku.fi/laitokset/neuro/37the.htm.

Lamberty GJ & Bieliauskas LA. (1993). Distinguishing between depression and dementia in the elderly: a review of neuropsychological findings. *Archives Clin Neuropsychol* 8: 149-70.

Landfield P, Waymire J & Lynch G. (1978). Hippocampal aging and adrenocortioids: quantitative correlations. *Science* 214: 581-4.

Larrabbee GJ & Crook T. (1988). Assessment of drug effects in age-related memory disorders: clinical, theoretical and psychometric considerations. *Psychopharm Bull* 24: 515-22.

Larrabee GJ & Crook T. (1994). Estimated Prevalence of Age-Associated Memory Impairment Derived From Standardized Tests of Memory Function. Int *Psychogeriatrics* 6: 95-104.

Larrabee GJ & McEntee WJ. (1995). Age-Associated memory impairment: Sorting out the controversies. (editorial). *Neurology* 45: 611-14.

Leach J & Levy R. (1994). Reflections on the NINCDS/ADRDA criteria for the diagnosis of Alzheimer's Disease. *Int J Geriatr Psychiatr* 9:173-9.

Lebel P, Leduc N, Kerogat MJ, Latour J, Leclerc C, Béland F, Contandriopoulos A-P. (1999). [A dynamic model of frailty]. [Article in French]. *L'Année Gerontologique* 1999: 84-94.

Levine B, Stuss DT & Milberg WP. (1997). Effects of aging on conditioned associative learning: process analyses and comparison with focal frontal lesions. *Neuropsychology*, 11:367-81.

Levy R (1994). Report of the Working Party of the International Psychogeriatric Association in collaboration with the World Health Organization: Aging-Associated Cognitive Decline. *International Psychogeriatrics* 6: 63-8.

Lewinson PM, Allen NB, Seeley JR & Gotlib IH. (1999). First onset versus recurrence of depression: differential processes of psychosocial risk. *J Abnorm Psychol* 108: 483-9.

Lezak MD. (1982). Neuropsychological Assessment. 2nd Edition. New York: Oxford University Press.

Ling M, Perry P & Tsuant M. (1981). Side effects of corticosteroid therapy. *Arch Gen Psychiatry* 38: 471-7.

Linn RT, Wolf PA, Bachman DL, Knoefel JE, Cobb JL, Belanger AJ, Kaplan EF & D'Agostino RB. (1995). The 'preclinical phase' of probable Alzheimer's disease. A 13-year prospective study of the Framingham cohort. *Arch Neurol* 52: 485-90.

Lupien SJ & Forget H. (1995). Glucocorticoïdes et cognition. *Médecine/Science* 11: 1524-30.

Lupien SJ & McEwen BS. (1997). The acute effects of corticosteroids on cognition: integration of animal and human model studies. *Brain Research Rev* 24: 1-27.

Lupien SJ, de Leon M, De Santi S, Convit A, Tarshish C, Nair NPV, Thakur M, McEwen BS, Hauger RL & Meaney MJ. (1998). Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature neurosci* 1(1): 69-73.

Lupien SJ, Gaudreau S, Tchiteya BM, Maheu F, Sharma S, Nair NPV, Hauger RL, McEwen BS & Meaney MJ. (1997). Stress-Induced Declarative Memory Impairment in Healthy Elderly Subjects: Relationship to Cortisol Reactivity. *J Clin Endocrinol Metab* 82(7): 2070-5.

Lupien SJ, Gillin C, Frakes D, Soefje S & Hauger RL. (1995). Delayed but not immediate effects of a 100 minutes hydrocortisone infusion on declarative memory performance in young normal adults. *Int Soc Psychoneuroendocrinol Abstr*: 25.

Lupien SJ, Lecours AR, Lussier I, Schwartz G, Nair NPV & Meany MJ. (1994). Basal Cortisol Levels and Cognitive Deficits in Human Aging. *J Neurosci* 14: 2893-2903.

Lupien SJ, Lecours AR, Schwartz G, Sharma S, Hauger RL, Meaney MJ & Nair NP. (1995) Longitudinal study of basal cortisol levels in healthy elderly subjects: evidence for subgroups. *Neurobiol Aging* 1996 Jan-Feb;17(1):95-105

Lupien SJ, Nair NPV, Brière S, Maheu F, Tu MT, Lemay M, McEwen BS & Meaney MJ. (1999). Increased cortisol levels and impaired cognition in human aging: implication for depression and dementia in later life. *Rev Neurosci* 10: 117-39.

Lupien SJ, Tannenbaum EM, Ohashi E, Meaney MJ. (1996). Daily salivary cortisol levels in healthy elderlies correlate with feelings of fatique, depression, and stress. *Neurosci Abstr* 22: 1889.

Lussier I, Peretz I, Belleville S & Fontaine F. (1989). Contribution of indirect measures of memory to clinical neuropsychology assessment. (abstract.) *J Clin Exp Neuropsychol* 11: 64.

Madden KM & Hogan DB. (1998). Drug Therapy for Alzheimer's Disease. *Can J CME* April 1998: 111-20.

Martignoni E, Costa A, Sinforiani E, Liuzzi A, Chiodini P, Mauri M, Bono G & Nappi G. (1992). The brain as a target for adrenocortical steroids: cognitive implications. *Psychoneuroendocrinol* 17: 343-54.

Mason JW. (1968). A review of psychoneuroendocrine research on the pituitary-adrenal cortical system. *Psychosom Med* 30: 576-607.

McDougall GJ & Balyer J. (1998). Decreasing mental frailty in at-risk elders. *Geriatr Nurs* 19: 220-4.

McDowell I, Kristjansson B, Hill GB & Hébert R. (1997). Community screening for dementia: the Mini-Mental State Exam and Modified Mini-Mental State Exam (3MS) Compared. *J Clin Epidemiol* 50: 377-83.

McEwen B. (1987). Glucocorticoids. In: Adelman G (Ed.). (1987). Encyclopedia of Neuroscience. Boston-Basel-Stuttgart: Birkhäuser. pp.467-8.

McEwen B. (1997). Protective and damaging effects of stress mediators. *New England J Medicine* 338: 171-9.

McGlone J, Gupta S, Humphrey D, Oppenheimer S, Mirsen T & Evans DR. (1990). Screening for early dementia using memory complaints from patients and relatives. *Arch Neurol* 47: 1189-93.

McKann G, Drachman D, Folstein M, Katzman R, Price D & Stadlan EM. (1984). Clinical

diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 34: 939-44.

Meaney MJ, O'Donnell D, Rowe W, Tannenbaum B, Steverman A, Walker M, Nair NP & Lupien S. (1995). Individual differences in hypothalamic-pituitary-adrenal activity in later life and hippocampal aging. *Exp Gerontol* 30:229-51.

Merriam EP, Thas ME, Haas GL, Keshavan MS & Sweeney JA. (1999). Prefrontal cortical dysfunction in depression determined by Wisconsin Card Sorting test performance. *Am J Psychiatry* 156: 780-2.

Milberg WP, Hebben N & Kaplan E. (1986). The Boston process approach to neuropsychological assessment. In: Grant I & Adams KM (Eds.). Neuropsychological assessment of neuropsychiatric disorders. New York: Oxford University Press. pp. 65-86.

Miller AH, Sastry G, Speranza AJ, Lawlor BA, Mohs RC, Ryan TM, Gabriel SM, Serby M, Schmeidler J & Davis KL. (1994). Lack of association between cortisol hypersecretion and nonsupression on the DST in patients with Alzheimer's Disease. *Am J Psychiatry* 151: 267-70.

Miller EK. (2000). The prefrontal cortex and cognitive control. *Nat Rev Neurosci* 1:59-65

Milner B, Petrides M & Smith L. (1985). Frontal lobes and the temporal organization of memory. *Human Neurobiol* 4: 137- 42.

Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, Sano M, Bieliauskas C, Geldmacher D, Clark C & Thal LJ. (1997) Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. *Alzheimer Dis Assoc Disord* 11 Suppl 2:S13-21.

Molloy DW, Alemayehu E & Roberts R. (1991). The Standardized mini-mental state examination. *Am J Psychiatry* 148: 102-5.

Monsch AU, Bondi MW, Butters N, Salmon DP, Katzman R & Thal LJ. (1992). Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Arch Neurol* 49: 1253-8.

Montorio I & Izal M (1996). The Geriatric Depression Scale: a review of its development and utility. *Int Psychogeriatr* 8:103-12.

Mook DG. (1983). In defense of external invalidity. *American Psychologist* 38: 379-87.

Moreau JL (1997). Validation of an animal model of anhedonia, a major symptom of depression *Encephale* 23:280-9

Morris JC, McKeel Jr DW, Storandt M, Rubin EH, Price JL, Grant EA, Ball MJ & Berg L (1991). Very mild Alzheimer's Disease: Informant-based clinical psychometric and pathologic distinction from normal aging. *Neurology* 41: 469-478.

Morris JC, Storandt M, Miller JP, McKeel DW, Price JL, Rubin EH & Berg L. (2001). Mild cognitive impairment represents early-stage Alzheimer disease. *Arch Neurol* 58: 397-405.

Morris JC, Storandt M, McKeel DW, Jr, Rubin EH, Price JL, Grant EA & Berg L (1996). Cerebral amyloid deposition and diffuse plaques in "normal" aging: Evidence for presymptomatic and very mild Alzheimer's disease. *Neurology* 46: 707-19.

Murrell SA & Norris FH. (1984). Resources, life events, and changes in positive affect and depression in older adults. *Am J Community Psychol* 12: 445-64.

Nebes RD & Brady CB (1992). Generalized cognitive slowing and severity of dementia in Alzheimer's disease: implications for the interpretation of response-time data. *J Clin Exp Neuropsychol* 14: 317-26.

Nilsberth C, Westlind-Danielsson A, Eckman CB, Condron MM, Axelman K, Forsell C, Stenh C, Luthman J, Teplow DB, Younkin SG, Näslund J & Lannfelt L. (2001). The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Aβ protofibril formation. *Nature Neurosci* 4: 887-93.

Oakes W. (1972). External validity and the use of real people as subjects. *American Psychologist* 27: 959-62.

O'Boyle M, Amadeo M & Self D. (1990). Cognitive complaints in elderly derpessed and pseudodemented patients. *Psychol Aging* 5: 467-8.

O'Brian JT & Levy R. (1992). Age associated memory impairment: too broad an entity to justify drug treatment yet. *British Med J* 304: 5-6.

O'Connor DW, Pollitt PA, Hyde JB, Miller ND & Fellowes JL. (1991). Clinical issues relating to the diagnosis of mild dementia in a British community survey. *Arch Neurol* 48:530-4

O'Hara MW, Hinrichs JV, Kohout FJ, Wallace RB & Lemke JH. (1986). Memory complaint and memory performance in the depressed elderly. *Psychol Aging* 1: 208-14.

Pablos-Mendez A & Cimino J. (1997). Screening. In: Miriam Rabkin (Ed.). General Medicine Ambulatory Care Syllabus. Columbian Presbytarian Medical Centre. On the Internet: cpmc.columbia.edu/whichis/private/aim/02SCREEN.html

Parnetti L, Lowenthal DT, Presciutti O, Pelliccioli G, Palumbo R, Gobbi G, Chiarini G, Palumbo B, Tarducci R & Senin U (1996). 1H-MRS MRI-Based Hippocampal Volumetry and 99mTc-HMPAO-SPECT in Normal Aging, Age-Associated Memory Impairment and Probable Alzheimer's Disease. *J Am Geriatr Soc* 44: 133-38.

Parry G & Brewin CR. (1988). Cognitive style and depression: symptom-related, event-related or independent provoking factor? *Br J Clin Psychol.* 27 (Pt 1) :23-35.

Pasquier F, Lebert F, Grymonprez L & Petit H. (1995). Verbal fluency in dementia of frontal lobe type and dementia of Alzheimer type. *J Neurol Neurosurg Psychiatry* 58: 81-4.

Patterson CJS & Gass DA. (2001). Screening for cognitive impairment and dementia in the elderly. *Can J Neurol Sci* 28 Suppl. 1: S42-S51.

Patterson C, Gauthier S, Bergman H, Cohen C, Feightner JW, Feldman H, Grek A & Hogan DB. (2001). The recognition, assessment and management of dementing disorders: conclusions fro the Canadian consensus conference on dementia. *Can J Neurol Sci* 28: Suppl 1: S3-S16.

Patterson MB, Mack JL, Geldmacher DS & Whitehouse PJ. (1996). Executive functions and Alzheimer's disease: problems and prospects. *Eur J Neurol* 3: 5-15.

Pena-Casanova J. (1997). Alzheimer's Disease Assessment Scale--cognitive in clinical practice. J Int Psychogeriatr 9 Suppl 1:105-14.

Pernerger TV. (1998). What's wrong with Bonferroni adjustments. *British Med J* 316: 1236-8.

Pernerger TV. (1999). Adjusting for multiple testing in studies is less important than other concerns. [Letter]. *British Med J* 318: 1288.

Persson G, Berg S & Nilsson L. (1991). Subclinical dementia. Relation to cognition, personality and psychopathology: a nine year prospective study. *Int J Geriatr Psychiat* 6: 239-47.

Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL & DeKosky ST. (2001). Practicie parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommitte of the American Academy of Neurology. *Neurology* 56: 1133-42.

Petersen RC, Smith GE, Kokmen E, Ivnik RJ & Tangalos EG. (1992). Memory function in normal aging. *Neurology* 42: 396-401.

Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG & Kokmen E. (1999). Mild cognitive impairment. Clinical characterization and outcome. *Arch Neurol* 56: 303-8.

Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG & Kokmen E. (1999). Mild cognitive impairment. Clinical characterization and outcome. *Arch Neurol* 56: 303-8.

Pianta RC & Egeland B (1994). Relation between depressive symptoms and stressful life events in a sample of disadvantaged mothers. *J Consult Clin Psychol* 62:1229-34

Pirozzolo FJ, Christensen KJ, Ogle KM, Hansch EC & Thompson WG. (1981). Simple and choice reation time in dementia: clinical implications. *Neurobiol Aging* 2: 113-7.

Porter AC & Raudenbush SW. (1987). Analysis of covariance: its model and use in psychological research. *J Counseling Psychol* 34: 383-92.

Pruessner JC, Gaab J, Hellhammer DH, Lintz D, Schommer N & Kirschbaum C. (1997). Increasing correlations between personality traits and cortisol stress responses obtained by data aggregation. *Psychoneuroendocrinology* 22: 615-25.

Pruessner JC, Wolf OT, Hellhammer DH, Buske-Kirschbaum A, von Auer K, Jobst S, Kaspers F & Kirschbaum C. (1997). Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. *Life Sci* 61:2539-49

Pryse-Phillips W. (1999). Do we have drugs for dementia? No. *Arch Neurol* 56: 735-7.

Rabbitt PMA. (1965). An age decrement in the ability to ignore irrelevant information. *J Gerontology* 20:233-8.

Raff H, Raff JL, Duthie EH, Wilson CR, Sasse EA, Rudman I & Mattson D. (1999). Elevated salivary cortisol in the evening in healthy elderly men and women: correlation with bone mineral density. *J Gerontol A Biol Sci Med Sci* 54:M479-83.

Raphael D, Cava M, Brown I, Renwick R, Heathcote K, Weird N, Wright K & Kirwan L. (1995). Frailty: a public health perspective. *Can J Public Health* 86: 224-7.

Reaven GM. (1994). Syndrome X:6 years later. *J Internal Medicine* 236 Suppl 736: 13-22.

Reed BR, Paller KA & Mungas D. (1998). Impaired acquisition and rapid forgetting of patterned visual stimuli in Alzheimer's. *J Clin Exp Neuropsychol* 20: 738-49.

Reisberg B, Ferris S, DeLeon M & Crook T. (1982). The global deterioration scale for assessment of primary degenerative dementia. *Am J Psychiatry* 139: 1136-9.

Rey A. (1964). L'examen clinique en psychologie. Paris: Presses Universitaires de France.

Ritchie K, Artero S & Touchon J. (2001). Classification criteria for mild cognitive impairment. A population-based validation study. *Neurology* 56: 37-42.

Robinson-Whelen S & Storandt M. (1992). Immediate and delayed prose recall among normal and demented adults. *Arch Neurol* 49: 32-4.

Rockwood K & Morris JC. (1996). Global staging measures in dementia. In: Gauthier S & Cummings JL (Eds.). Clinical Diagnosis and Management of Alzheimer's Disease. London: Martin Dunitz, 1996, pp 141-153.

Rohling ML & Scogin F. (1993). Automatic and effortful memory processes in depressed persons. *J Gerontol Psychol Sci* 48: p87-p95.

Rosen W. (1980). Verbal fluency in aging and dementia. *J Clin Neuropsychol* 2: 135-46.

Rosen WG, Mohs RC & Davis KL. (1984). A New Rating Scale for Alzheimer's Disease. *Am J Psychiatry* 141: 1356-64.

Rowe JW & Kahn RL (1987). Human aging: Usual and successful. *Science* 237: 143-149,

Rowe W, Steverman A, Walker M, Sharma S, Barden N, Seckl JR & Meaney MJ. (1997). Antidepressants restore hypothalamic-pituitary-adrenal feedback function in aged, cognitively-impaired rats. *Neurobiol Aging* 18: 527-33.

Royall DR. (1997). Some "depressive" symptoms may not imply depression. *J Am Geriat Soc* 45: 891-7.

Sachar EJ, Hellman L, Rofwarg HP, Halpern FS, Fukushima DK, Gallagher TF. (1973). Disrupted 24-hour patterns of cortisol secretion in psychotic depression. *Arch Gen Psychiatr* 28: 19-24.

Sandi C, Loscertales M & Guaza C. (1997). Experience-dependent facilitating effect of corticosteone on spatial memory formation in the water maze. *Eur J Neurosci* 9:637-42.

Sapolsky R. (1992). Stress, the Aging Brain, and the Mechanisms of Neuron Death. Cambridge, MASS: MIT Press.

Sapolsky RM, Krey LC & McEwen BS. (1986). The neuroendocrinology of stress and aging: the glucocorticoid cascade hypothesis. *Endocrinol Rev* 7: 284-301.

Schaie KW. (1983). The Seattle longitudinal study: A 21-year exploration of psychometric intelligence in adulthood. In: Schaie KW. (Ed.) Longitudinal studies of adult psychological development. New York: Guilford. pp.64-135.

Schmand B, Jonker C, Geerlings MI & Lindeboom J (1997). Subjective memory complaints in the elderly: depressive symptoms and future dementia. *Br J Psychiat* 171(10): 373-6.

Schmand B, Jonker C, Hooijer C & Lindeboom J. (1996). Subjective memory complaints may announce dementia. *Neurology* 46(1): 121-5.

Schulkin J, McEwen BS & Gold PW (1994). Allostasis, amygadala, and anticipatory angst. *Neurosci Biobehav Rev* 18: 385-96.

Scwarzer R. (Ed.). Self-Efficacy: Thought Control of Action. Washington, D.C.: Hemisphere.

Seeman TE, McEwen BS, Singer BH, Albert MS, Rowe JW. (1997). Increase in urinary cortisol excretion and memory declines: MacArthur studies of successful aging. *J Clin Endocrinol Metab* 82: 2458-65

Selkoe DJ. (1997). Alzheimer's Disease: Genotypes, Phenotype, and Treatments. *Science* 275: 630-1.

Silberman EK, Weingartner H, laraia M, Byrnes S & Post RM. (1983). Processing of emotional properties of stimuli by depressed and normal subjects. *J Nerv Ment Dis* 171: 10-14.

Small GW, La Rue A, Komo S & Kaplan A. (1997). Mnemonics Usage and Cognitive Decline in Age-Associated Memory Impairment. *Int Psychogeriat* 9: 47-56.

Small GW, Rabins PV, Barry PP, Buckholts NS, DeKosky ST, Ferris SH, Finkel SI, Gwyther LP, Khachaturian ZS, Lebowitz BD, McRae TD, Morris JC, Oakley F, Schneider LS, Streim JE, Sunderland T, Teri LA & Tune LE. (1997). Diagnosis and Treatment of Alzheimer Disease and Related Disorders: Consensus Statement of the American Association for Geriatric Psychiatry, the Alzheimer's Association, and the American Geriatrics Society. *J Am Geriatr Soc* 278: 1363-71.

Smith G, Ivnik RJ, Petersen RC, Malec JF, Kokmen E & Tangalos E. (1991). Age-associated memory impairment diagnoses: Problems of reliability and concerns for terminology. *Psychol Aging* 6: 551-558.

Smith GE, Peterson RC, Ivnik RJ, Malec JF & Tangalos EG. (1996). Subjective memory complaints, psychological distress and longitudinal change in objective memory

performance. *Psychol & Aging* 11: 272-9.

Snowdown J & Lane F. (1994). A Longitudinal Study of Age-Associated Memory Impairment. *Int J Geriatr Psychiatry* 9: 779-87.

Soininen HS, Partanen K, Pitkänen A, Vainio P, Hänninen T, Hallikainen M, Koivisto K & Riekkinen PJ Sr (1994). Volumetric MRI analysis of the amygdala and the hippocampus in subjects with age-associated memory impairment: Correlation to visual and verbal memory. *Neurology* 44: 1660-1668.

Somerville M. (2000). The Ethical Canary: Science, Society and the Human Spirit. Toronto, Canada: Penguin Books.

Spencer WD & Raz N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychol & Aging* 10: 527-39.

Spjotvoll E & Stoline MR (1973). An extension of the T-method of multiple comparison to include the cases with unequal sample sizes. *J Am Stat Association*, 68: 976-978.

Squire LR, Haist F & Shimamamura AP. (1989). The neurology of memory: Quantitative assessment of retrograde amnesia in two groups of amnesic patients. *J Neuroscience* 9: 828-39.

Squire LR. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys and humans. *Psychol Rev* 99: 195-23.

Standish TIM, Molloy DW, Bédard M, Layne EC, Murray EA & Strand D. (1996). Improved Reliability of the Standardized Alzheimer's Assessment Scale (SADAS) Compared with the Alzheimer's Disease Assessment Scale (ADAS). *J Am Geriatr Soc* 44: 712-16.

Stansbury K, Haley D & Koeneker A. (2000). Higher cortisol values facilitate spatial memory in toddlers. Brief report. *Ann N Y Acad Sci* 911:456-8.

Starkman MN, Gebarski SS, Berent S, Schteingart DE. (1992). Hippocampal formation volume, memory dysfunction, and cortisol levels in patients with Cushing's syndrome. Biol Psychiatr 32: 756-65.

StatSoft, Inc. (1999). Electronic Statistics Textbook. Tulsa, OK: StatSoft. Internet: http://www.statsoft.com/textbook/stathome.html

Steenhuis R, Østbye T. (1995). Neuropsychological test performance of specific diagnostic groups in the Canadian study of health and aging. J Clin Exp Neuropsychol 17: 773-85.

Sterling P & Eyer J. (1988). Allostasis: a new paradigm to explain arousal pathology. In: Fisher S, Reason J. (Eds.). Handobook of Life Stress, Cognition and Health. New York: John Wiley & Sons. pp. 629-49.

Stiles PG & McGarrahan JF (1998). The Geriatric Depression Scale: A comprehensive review. *J Clin Geropsychology*, 4: 89-110

Stip E, Lecours AR. (1992). Fonctionnement neuropsychologique du déprimé. *L'Encéphale* XVIII: 575-83.

Storandt M & Hill RD. (1989). Very mild senile dementia of the Alzheimer type. II. Psychometric test performance. *Arch Neurol* 46: 383-6.

Strawbridge WJ, Shema SJ, Balfour Jl, Higby HR & Kaplan GA. (1998). Antecedents of frailty over three decades of life in an older cohort. *J Gerontol B Psychol Sci* 53: S9-S16.

Suzuki WA. (1996) The anatomy, physiology and functions of the perirhinal cortex. *Curr Opin Neurobiol* 6: 179-86.

Teasdale JD & Russell ML. (1983). Differential effects of induced mood on the recall of positive, negative and neutral words. *Br J Clin Psychol* 22:163-71.

Teng EL & Chui HC. (1987). The Modified Mini-Mental State (3MS) Examination. *J Clin Psychiatry* 48: 314-18.

Teng EL, Chui HC & Gong A. (1989). Comparisons between the Mini-Mental State Exam (MMSE) and its modified version--the 3MS test. In: Hasegama K & Homma A. (Eds.) Psychogeriatrics: Biomedical and Social Advances. Selected Proceedings of the Fourth Congress of the International Psychogeriatrics Association. Tokyo: Excerpta Medica.

The Lundt and Manchester groups. (1994). Clinical and neuropathological criteria for frontotemproal dementia. *J Neurol Neurosurg Psychiatry* 57: 416-8.

Tierney MC, Szalai JP, Snow WG, Fisher RH, Nores A, Nadon G, Dunn E & St George-Hyslop PH. (1996). Prediction of probable Alzheimer's disease in memory-impaired patients: A prospective longitudinal study. *Neurology* 46: 661-65.

Tomb DA. (1994). The phenomenology of post-traumatic stress disorder. *Psychitr Clin North Am* 17: 237-50.

Tombaugh TN & McIntyre NJ. (1992). The Mini-Mental State Examination: A Comprehensive Review. *J Am Geriat Soc* 40: 922-35.

Tristmans L, Clincke G & Peelmans B. (1990). Does AAMI Constitute a Real Disease Entity? A Placebo-Controlled Double-Blind Study With Sabeluzole (R58 735) in a Patient Population with Real Memory Problems. *Drug Development Research* 20: 473-82.

Tuokko H & Frerichs R. (2000) Cognitive impairment with no dementia (CIND): longitudinal studies, the findings and the issues. *Clin Neuropsychol* 14: 504-25.

Tuokko H, Kristjansson E & Miller J. (1995). Neuropsychological detection of dementia: an overiew of the neuropsychological component of the Canadian Study of Health and Aging. *J Clin Exp Neuropsychology* 17: 352-73.

Unverzagt FW, Gao S, Baiyewu O, Ogunniyi AO, Gureje O, Perkins A, Emsley CL, Dickens J, Evans R, Musick B, Hall KS, Hui SL & Hendrie HC. (2001). Prevalence of cognitive impairment. Data from the Indianapolis study of health and aging. *Neurology* 57: 1655-62.

van Londen L, Goekoop JG, Zwinderman AH, Lanser JB, Wiegant VM & De Wied D. (1998). Neuropsychological performance and plasma cortisol, arginine vasopressin and oxytocin in patients with major depression. *Psychol Med* 28: 275-84.

van Reekum R, Simard M & Cohen T. (1999). The prediction and prevention of Alzheimer's Disease—towards a research agenda. *J Psychiatr Neurosci* 24: 413-430; p.417.

Veenema JC, Spruijt BM, Gispen WH, Van Hoof JARAM. (1997). Aging, dominance history, and social behavior in Java-Monkeys (Macaca fascicularis). *Neurobiol Aging* 18: 509-15.

Velmans M. (1991). Is human information processing consious? *Behavioral Brain Sciences* 14: 651-726.

Velmans, M. (1990) Consciousness, brain and the physical world. *Philosophical Psychology,* 3: 77-99.

Velmans, M. (1993) A reflexive science of consciousness. In: Experimental and Theoretical Studies of Consciousness. Ciba Foundation Symposium 174. Chichester: Wiley.

Velmans, M. (1996) What and where are conscious experiences? In Velmans M (ed.), The Science of Consciousness: Psychological, Neuropsychological and Clinical Reviews. London: Routledge.

Voytko ML, Olton DS, Richardson RT, Gorman LK, Tobin JR & Price DL. (1994). Basal forebrain lesions in monkeys disrupt attention but not learning and memory. *J Neurosci* 14: 167-86.

Waugh NC & Norman DA. (1965). Primary memory. *Psychol Rev* 72: 89-104.

Wechsler D & Stone CP. (1973). Wechsler Memory Scale. New York: Psychological Corp.

Wechsler D. (1945). A standardized memory scale for clinical use. *J Psychol* 19: 87-95.

Wecshler D. (1981). Wechsler Adult Intelligence Scale®—III. Available from: PsychCorp.Com

Weiner MF, Vobach S, Olsson K, Svetlik D & Risser RC. (1997). Cortisol secretion and Alzheimer's Disease progression. *Biol Psychiatry* 42:1030-8.

Weingartner H, Cohen RM, Murphy DL, Martello J & Gerdt C. (1981). Cognitive processes in depression. *Arch Gen Psychiatry* 38: 42-7.

Weitzman ED, Fukushima D, Nogeire C, Roffwarg H, Gallagher TF & Hellman L. (1971). Twenty-four hour pattern of the episodic secretion of cortisol in normal subjects. *J Clin Endocrinol Metab* 33:14-22.

Welford AT (1984). Between bodily changes and performance: some possible reasons for slowing with age. *Exp Aging Res* 10:73-8.

Welsh KA, Butters N, Hughes J, Mohs R & Heyman A. (1991). Detection of abnormal memory decline in mild cases of Alzheimer's Disease using CERAD neuropsychological measures. *Archives of Neurology* 48:278-81.

Welsh KA, Butters N, Mohs RC, Beekly BS, Edland S, Gillenbaum G & Heyman A. (1994). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A Normative study of the neuropsychological battery. *Neurology* 44: 609-14.

Wetherell JL, Gatz M, Johansson B & Pedersen NL. (1999). History of depression and other psychiatric illness as risk factors for Alzheimer disease in a twin sample. *Alzheimer Dis Assoc Disord* 13: 47-52.

Wilson RS & Evans DA. (1996). How clearly do we see our memories? *J Am Geriat Soc* 44(2): 93-4.

Wilson RS, Bacon LS, Kramer RL, Fox JH & Kaszniak AW. (1983). Word frequency effect and recognition memory in dementia of the Alzheimer type. *J Clin Neuropsychol* 5: 97-104.

Winer BJ, Brown DR & Michels KM. (1991). Statistical Principles in Experimental Design 3rd Edition. New York: McGraw-Hill.

Wolfson C, Moride Y, Perrault A & Vida S. (1997).A study of the efficacy, effectiveness and economic impact of Tacrine in alzheimer's disease. Ottawa, Ontario: Canadian Coordinating Office for Health Technology Assessment. Available from the internet at: http://www.ccohta.ca

Wolkowitz OM, Reus VI, Weingartner H, Thompson K, Breier A, Doran A, Rubinow D & Pickar D. (1990). Cognitive effects of corticosteroids. *Am J Psychiatr* 147: 1297-1303.

Wolkowitz OM, Weingartner H, Rubinow DR, Jimerson D, Kling M, Berretini W, Thompson K, Breier A, Doran A, Reus VI & Pickar D. (1993). Steroid modulation of human memory: biochemical correlates. *Biol Psychiatr* 33: 744-6.

World Health Organization. (1993). The International Classification of Diseases-10 Classification of Mental and Behavioural Disorders: diagnostic criteria for research. Geneva: World Health Organization.

Yastrubetskaya O, Chiu E & O'Connell S (1997). Is Good Clinical Research Practice For Clinical Trials Good Clinical Practice? *In J Geriatric Psychi* 12, 227-31.

Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M & Leirer VO. (1983). Development and Validation of a Geriatric Depression Screening Scale: A Preliminary Report. *J Psychia Res* 17: 37-49.

Youngjohn JR & Crook III TH. (1993). Stability of Everyday Memory in Age-Associated Memory Impairment: A Longitudinal Study. *Neuropsychology* 7(3): 406-16.

Zec RF, Landreth ES, Vicari SK, Feldman E, Belman J, Andrise A, Robbs R, Jumar V & Becker R. (1992). Alzheimer Disease Assessment Scale: useful for both early detection and staging of Dementia of the Alzheimer Type. *Alzheimer's Disease Assoc Disord* 6: 89-102.