Assessment of spatial orientation in Alzheimer's disease:

Theoretical and clinical implications

Lili Liu, M.Sc.

School of Physical and Occupational Therapy

McGill University, Montreal

March, 1993

A Thesis Submitted to

the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for

the degree of Ph.D. in Rehabilitation Science

ISBN 0-315-87954-8

Abstract

The purpose of this project was to develop a reliable and valid battery for the assessment of spatial orientation skills (SOS) in persons with Alzheimer's disease (AD). The battery, comprised of 13 subtests, was administered to 97 normal control subjects, 25 subjects with early AD and 10 with late AD. The test-retest reliability of the battery was based on the test results of 33 normal control subjects and 25 early AD subjects. Inter-rater reliability was determined using four trained raters who evaluated 27 normal control subjects and the same 25 early AD subjects. Content validity was established using a panel of six experts and construct validity was determined by comparing the performance of the normal control and early AD groups. To establish criterion validity, the Global Deterioration Scale (GDS) was used as the criterion. For the AD group, eight subtests demonstrated acceptable test-retest and inter-rater reliability coefficients (ICC $\geq$ .70). For the control group, three subtests had acceptable test-retest coefficients and four had acceptable inter-rater coefficients. The internal consistency of the battery was acceptable as shown by overall Cronbach's alpha of .86 for AD subjects and .72 for control subjects, and was further analyzed using factor analysis which yielded five factors. Logistic regression provided evidence for good construct validity. Scores on the SOS subtests were able to differentiate the three groups of subjects established on the basis of the GDS scores (GDS 1 and 2, GDS 3 and 4, and GDS 5). A preliminary shortened version of the battery was developed using six subtests which demonstrated high test-retest and inter-rater reliability. The performance of subjects with AD on the battery is discussed with respect to its implications for the theoretical basis and clinical assessment of spatial orientation in AD.

Résumé

Le but de ce projet était de développer une batterie d'évaluations fiable et valide des habiletés d'orientation spatiale (SOS) pour les personnes ayant la maladie d'Alzheimer (MA). La batterie comprenant 13 sous-tests fut administrée à 97 sujets contrôles normaux, 25 sujets ayant une MA précoce et 10 sujets ayant une MA avancée. La fiabilité test-retest de la batterie fut basée sur les résultats de 33 sujets contrôles et 25 sujets avec MA précoce. La fiabilité inter-juges a été déterminée par 4 examinateurs entraînés qui ont évalué 27 sujets contrôles et les mêmes 25 sujets avec MA précoce. La validité de contenu fut établie à l'aide d'un panel de six experts et la validité de construit fut déterminée en comparant la performance des groupes contrôle et MA précoce. Afin d'établir la validité de critère, le "Global Deterioration Scale" (GDS) fut utilisé comme critère. Pour le groupe MA, 8 sous-tests ont démontré des coefficients fiabilité test-retest et inter-juges acceptables (ICC $\geq$ .70). Pour le groupe contrôle, trois sous-tests avaient des coefficients test-retest acceptables et quatre étaient acceptables pour les coefficients inter-juges. La consistance interne de la batterie fut acceptable telle que démontrée par l'alpha de Cronbach (.86 pour le groupe MA et .72 pour le groupe contrôle) et fut analysée en plus à l'aide d'une analyse factorielle qui a produit cinq facteurs. Une régression logistique a mis en évidence une bonne validité de construit. Les résultats aux sous-tests de la SOS ont différencié les trois groupes de sujets établis selon les performances au GDS (GDS 1 et 2, GDS 3 et 4, et GDS 5). Une version préliminaire raccourcie de la batterie fut développée utilisant six sous-tests qui démontraient des fiabilités test-retest et inter-juges élevées. La performance des sujets MA avec la batterie est discutée tout en tenant compte des bases théoriques et de l'évaluation clinique de l'orientation spatiale dans la MA.

Table of Contents

Abstract ii

Résumé iii

List of Tables xi

List of Figures xiv

Preface xv

1.0 INTRODUCTION 1

1.1 Diagnosis and Pathology of Alzheimer's Disease 2

1.2 Cognitive Functions in Persons with Alzheimer's Disease 6

1.2.1 Primary and working memory 7

1.2.2 Attention 8

1.2.3 Secondary memory: Declarative memory 10

1.2.4 Secondary memory: Procedural memory 12

1.2.5 Language 13

1.3 Spatial Orientation in Alzheimer's Disease 15

1.3.1 Definition 15

1.3.2 Spatial skills in Alzheimer's disease 16

1.3.3 Figure-ground discrimination 16

1.3.4 Personal orientation 18

1.3.5 Extrapersonal orientation 20

1.3.6 Visuospatial constructional ability 22

1.3.7 Mental representation 24
1.3.8 Functional spatial skills 28
1.3.9 Functional spatial orientation: Way-finding and wandering behaviour 29
1.3.10 Summary 34
1.4 Test Development 35
1.4.1 Reliability 37
1.4.2 Validity 39
1.5 Rationale for the Study 42
1.6 Objectives of the Study 44
2.0 METHODS 46
2.1 Subjects 46
2.1.1 Normal Control Subjects 46
2.1.2 Alzheimer Subjects 49
2.2 Materials 54
2.2.1 Mental Status 54
2.2.2 Spatial Orientation Skills (SOS) Battery: Perceptual Spatial Skills Subtests 55
2.2.2.1 Left-Right Discrimination 56
2.2.2.2 Figure-Ground Discrimination 57
2.2.2.3 Position in Space 59
2.2.2.4 Spatial Relations 60

2.2.3 Spatial Orientation Skills (SOS) Battery: Cognitive Spatial Skills Subtests 63
2.2.3.1 Porteus Maze Test 64
2.2.3.2 Map-Reading Test 64
2.2.3.3 Road-Map Test 68
2.2.3.4 Corsi Block Tapping Test 69
2.2.3.5 Modified Stylus Maze Test 72
2.2.3.6 Mental Representation of the Familiar Environment 75
2.2.3.7 Sketch Map of the New Environment 77
2.2.4 Spatial Orientation Skills (SOS) Battery: Functional Spatial Abilities 77
2.2.5 Other Tests of Cognitive Function 80
2.2.5.1 Attention 81
2.2.5.2 Language 81
2.2.5.3 Verbal memory 82
2.2.5.4 Clock drawing 82
2.3 Procedure 82
2.3.1 Content validity 83
2.3.2 Test-retest and inter-rater reliability 84
2.3.3 Construct validity 85
2.3.4 Criterion validity 87
2.4 Data Analysis 88

3.0 RESULTS 95
3.1 Content validity 95
3.2 Descriptive and Comparative Analysis of the SOS Battery 96
3.2.1 Comparative analysis of subtests of SOS Battery 101
3.2.2 Functional spatial abilities 101
3.3 Correlation Between SOS Subtests and Demographic/Mental Status Variables 105
3.4 Test-Retest and Inter-Rater Reliabilities of SOS Subtests 105
3.5 Internal Consistency 109
3.5.1 Subtest-to-subtest correlations 109
3.5.2 Subtest-to-total score correlations 117
3.5.3 Internal consistency: Cronbach's coefficient alpha 117
3.5.4 Factor analysis using control group data 120
3.5.5 Selection of subtests for a shortened SOS Battery 122
3.5.6 Factor analysis of shortened SOS Battery using both AD and control groups 124
3.5.7 Subtest-to-total correlations for the shortened battery 128
3.5.8 Cronbach's alphas for the shortened battery 130
3.5.9 Internal consistency: Summary 130
3.6 Validity 132
3.6.1 Construct validity 132
3.6.2 Criterion validity 134

3.6.3 Summary of validity studies 141

3.7 Results of Other Tests of Cognitive Function 142

3.7.1 Mental status 142

3.7.2 Attention 143

3.7.3 Language 143

3.7.4 Verbal memory span (Hebb verbal span) 144

3.7.5 Clock drawing test 144

3.7.6 Summary 145

3.8 Summary of Results 146

4.0 DISCUSSION 148

4.1 Psychometric properties of SOS Battery 148

4.1.1 Normative data 149

4.1.2 Test-retest and inter-rater reliability 151

4.1.3 Internal consistency of the full battery 151

4.1.4 Construct and criterion validity of the full battery 154

4.2 Characteristics of the Shortened SOS Battery 155

4.3 Variables that affect spatial skills 158

4.4 Theoretical Implications 162

4.4.1 Spatial skills in Alzheimer's disease 162

4.4.1.1 Figure-ground discrimination 162

4.4.1.2 Personal orientation 164

4.4.1.3 Extrapersonal orientation 166

4.4.1.4 Visuospatial constructional ability 167

4.4.1.5 Mental representation 168

4.4.1.6 Functional spatial orientation skills 171

4.4.1.7 Summary 173

4.4.2 Framework for studying spatial orientation in AD patients 176

4.5 Clinical Implications 178

4.5.1 Clinical assessment of spatial orientation 178

4.5.2 Clinical management of spatial disorientation in AD patient 182

4.6 Limitations of the study 184

4.7 Other contributions of the study 186

5.0 CONCLUSION 190

Bibliography 193

Appendices 218

Appendix A List of abbreviations 218

Appendix B General information form 219

Appendix C Consent forms 220

Appendix D Functional spatial abilities questionnaires (FSAQs) 222

Appendix E Questionnaires sent to caregivers 224

Appendix F Comparison of characteristics and mental status performance by the AD patients in stages 3 and 4 of the GDS 226

Appendix G Summary of results 227

Appendix H Content validation form for SOS Battery 228

Appendix I Comparison of performance on SOS Battery by the AD patients in stages 3 and 4 of the GDS 229

Appendix J Frequencies of self-rated responses on items of FSAQ 231

Appendix K Frequencies of proxy-rated responses on items of FSAQ 233

List of Tables

1. Characteristics of AD and control groups on the initial visit 52
2. Distribution of scores on subtests of the SOS Battery 97
3. Comparison of performance on perceptual spatial subtests 98
4. Comparison of performance on cognitive spatial subtests 99
5. Comparison of performance on the FSAQ 100
6. Correlation of SOS subtests with demographic/mental status variables for AD group 106
7. Correlation between SOS subtests and demographic/mental status variables for control group 107
8. Test-retest and inter-rater reliabilities for AD group 108
9. Test-retest and inter-rater reliabilities for control group 110
10. Correlation between perceptual subtests for AD and control groups 111
11. Correlation between cognitive subtests for AD group 112
12. Correlation between cognitive subtests for control group 113
13. Correlation between perceptual and cognitive subtests for AD and control groups 115
14. Correlation between perceptual and cognitive subtests for AD and control groups 116
15. Subtest-to-total Pearson product-moment correlation coefficients for AD and control groups 118

16. Cronbach's alpha coeffients for AD and control groups 119

17. Factor loadings of subtests for control group 121

18. Standardized subtest and total scores on shortened version of SOS battery: Control group compared to AD group 125

19. Factor loadings of subtests on shortened version of SOS battery for AD and control groups 127

20. Subtest-to-total correlation coefficients and Cronbach's alpha coefficients for AD group: Shortened version of SOS battery 129

21. Subtest-to-total correlation coeffcients and Cronbach's alpha coefficients for control group: Shortened version of SOS battery 131

22. Logistic regression and odds ratios for subtests of SOS battery 133

23. Comparison of the RNL scores obtained by subjects in the control and AD groups 135

24. Comparison of demographic characteristics and mental status of the control, early AD and late AD groups 137

25. Comparison of performance on the perceptual spatial subtests by control, early AD and late AD groups 138

26. Comparison of performance on the cognitive spatial subtests by control, early and late AD groups 139

27. Comparison of performance on the FSAQ by control, early AD and late AD groups 140

28. Comparison of characteristics and mental status performance by AD patients in stages 3 and 4 of the GDS 225

29. Comparison of performance on the SOS Battery by AD patients in stages 3 and 4 of the GDS 228

30. Frequencies of self-rated responses on items of FSAQ 230

31. Frequencies of proxy-rated responses on items of FSAQ 232

List of Figures

1. Figure-Ground Perception Test 58
2. Position in Space Test 61
3. Spatial Relations Test 62
4. Porteus Maze Test 65
5. Map-Reading Test 67
6. Road-Map Test 70
7. Corsi Block Tapping Test 71
8. Modified Stylus Maze Test 73
9. Mental representation of the familiar environment 76
10. Sketch map of the new environment 78
11. Sketch maps of familiar environment drawn by an AD subject and caregiver .. 102
12. Sketch map of new environment drawn by an AD subject 103
13. Bar graphs of scores on the shortened SOS Battery 126
14. Bar graph of RNL Index 136

Preface

In 15 years, there will be an additional 1,207,200 people over 65 years of age in Canada (McEwan, Donnelly, Robertson & Hertzman, 1991). This represents an increase of 38% over the present (1991) number of 3,169,600. In this group of individuals, dementia is one of the major and perhaps most disabling psychiatric disorders. The current estimated prevalence of dementia in individuals over 65 years of age is 6% which represents 189,985 cases in Canada. As the risk of dementia increases dramatically with age, the estimated increase in prevalence of dementia by 2006 is 7.4% or 324,188 cases (McEwan et al., 1991). Although Alzheimer's disease (AD) is the diagnosis for the majority of dementia cases, the exact proportion is controversial and estimates range from 50% to 60% (Bayles, Kaszniak & Tomeada, 1987). Using the more conservative estimate, it may projected that by the year 2006, there will be 162,094 individuals diagnosed with AD in Canada, an increase of 67,102 cases.

Because there is no medical treatment available to stop the degenerative course of AD, it is not so much the diagnosis that determines the extent of clinical intervention but the social and functional impairments associated with the disease, along with the availability and capabilities of caregivers. Researchers are now turning their attention to the measurement of function in an effort to better understand the intervention needs of people with AD (McEwan et al., 1991). This thesis presents one such study which attempts to further our understanding of spatial orientation and to gain some insight into possible intervention strategies in the management of AD patients who are spatially disoriented. As the theoretical basis of spatial orientation is still in its infancy and the psychometric properties of many well-used research and clinical measurement tools have yet to be determined, it was necessary to explore these issues before intervention strategies could be examined.

The Introduction of this thesis begins by examining the clinical challenges of Alzheimer's disease, the concept and theory of spatial orientation, and what is known about spatial abilities in individuals with Alzheimer's disease. This section then continues with the topic of test development and examines the psychometric properties required of a test or battery of tests used to study a group of Alzheimer's patients. It concludes with a rationale for the study. The Methods section contains the body of the study by describing the inclusion criteria for subjects in the control and Alzheimer groups, the methodology of the study, the mental status examinations and the 13 subtests of the Spatial Orientation Skills (SOS) Battery. The findings of this study are reported in the Results section. The Discussion section examines the theoretical and clinical implications of these results as well as the limitations of the study. The Conclusion section summarizes the original contributions of this project. These contributions include normative data and information on the reliability and validity of the battery of spatial tests when administered to AD patients and to normal control subjects. The use of several types of spatial tests concurrently was a unique aspect of this study. This section also offers an appraisal of the current status of research in spatial orientation and suggestions for further work in this area.

I am indebted to my thesis supervisors, Professor Louise Gauthier and Dr. Lynette Jones for their ongoing patience, support and availability during the entire course of my graduate studies. Thanks are due to Dr. Serge Gauthier who referred all of the Alzheimer patients, many of the control subjects and for providing ongoing feedback and consultation. Dr. Sharon Wood-Dauphinee graciously offered her expertise on the methodological issues of the study. Isabelle Gélinas, Jasmine Cooper and Constance Lalinec put in many hours for the administration of tests.

Nicole Paquet made the telephone calls, arranged the numerous appointments and was invaluable in helping with data entry. Dr. Michal Abrahamowicz advised me on some aspects of statistical analyses and Shan Shan Wang did some of the computer programming and analyses in SAS.

This research project has been fortunate to have the volunteer contributions of many experts and professionals. The Map-Reading Test was constructed by Yung-Chang Liu and the Corsi Block and Stylus Maze tests were constructed by Alan Hammaker. The Golden Age Societies of Montreal and the Alzheimer Society of Montreal generously gave of their time and energy by recruiting control subjects for the study through mailings and word of mouth. Of course, this project could not have been realized without the participation of each of the individuals tested and his or her accompanying family member. The number of hours of testing they endured, their energy and, sometimes, the great distances they travelled to reach the University attested to their enthusiasm about and dedication to the study.

This study was funded by a grant from the Alzheimer Society of Canada from 1990 to 1991 and a grant from the National Health and Research Development Program (NHRDP) from 1991 to 1992. The author of the thesis was supported by a fellowship from the Fonds pour la Formation de Chercheurs et l'Aide a la Recherche (FCAR) from 1989 to 1991.

I cannot put a value on the emotional and psychological support I have received while undergoing this challenge. I thank my mother and father who always believed in me, my sister, Yung-Yung Liu, who nurtured me and my three brothers who entertained and distracted me when it was necessary. Finally, I am grateful to my husband, Andrew Nisbet, who gladly illustrated the figures in this thesis, and who continues to believe in, nurture and entertain me.

1.0 INTRODUCTION

Dementia is a clinical condition in which there is general cognitive deterioration of sufficient severity to interfere with social and/or occupational functioning (American Psychiatric Association, 1987). Degenerative dementia is the most frequent of the dementing diseases accounting for up to 85% of the dementias (Katzman, 1986). It consists of those dementias that are progressive and not reversible (Consensus Conference, 1987). Alzheimer's disease (AD) (see Appendix A for a list of abbreviations) constitutes the diagnosis in the majority, or 50%-60%, of degenerative dementia cases (Bayles, Kaszniak & Tomeada, 1987).

AD is age-dependent, that is, the risk increases rapidly with advancing age. Severe dementia is found in 1% of persons over the age of 65, whereas it is present in more than 15% of persons over the age of 85 (Katzman, 1986; Van Hoesen & Damasio, 1987). The prevalence and incidence rates are unknown for many countries, and known estimates vary considerably even within the same country. Epidemiological data have been compared for Great Britain, Italy, Russia, Scandinavia, North America, Australia, China and Japan (Jorm, 1991; Mass et al., 1987; Rocca, Amaducci & Schoenberg, 1986). In general, these data show that the prevalence of AD is slightly higher in women then men, although there is considerable variation among the studies. Sulkava, Wikstrom and Aromaa (1985) reported a prevalence estimate of 4.2% for women compared to 2.6% for men in a sample of Finnish individuals who were 65 years of age or older. In Great Britain, higher estimates were obtained by Broe, Akhtar and Andrews (1976) but the difference between men and women was similar (6.3% for women and 5.0% for men). In contrast, Kaneko (1975) reported a higher prevalence of AD in Japanese men (2.2%) compared to Japanese women (1.6%) who were 65 years or older. As mentioned in the Preface, current estimates of the prevalence of AD in Canadians over the age of 65 is 6%, however, the incidence rate is unknown. Prevalence estimates of AD cases in the United States are approximately the same at 5% to 5.5% (Katzman, 1986; Schoenberg, Anderson & Haere, 1985). Incidence rates for the United States were collected in a longitudinal study using white males who were categorized into five age groups (Sluss, Gruenberg & Kramer, 1981). These rates range from 430 (60 to 64 year old catetory) to 3,248 (over 80 year old category) cases per 100,000 people per year. The incidence rates of AD decrease to 148 per 100,000 people per year for males in the 65 to 69 year old category but then increase progressively with each age category. Therefore, incidence rates may be distributed bimodally.

AD begins insidiously and may not be recognized until the disease has progressed for two or three years. In the initial stage, the patient usually complains of changes in memory for recent events and may present with signs of depression. This is followed by a gradual and steady progression of cognitive deterioration. The multiple cognitive changes include disturbances in memory, language use, spatial orientation, perception, praxis, the ability to learn skills, solve problems, think abstractly, and make judgements (McKhann et al., 1984). Personality traits may be intensified or disturbed (Huff et al., 1987). Other behaviours that may emerge include paranoid symptoms and delusions, irritability, agitation and verbal or physical agression (Eichelman & Phelps, 1992).

1.1 Diagnosis and Pathology of Alzheimer's Disease

The clinical differentiation of the degenerative dementias begins with ruling out reversible causes such as metabolic disorders or infections. There is no biochemical marker for AD and definitive diagnosis is made histologically by cerebral biopsy or autopsy (Katzman, 1986). In vivo, the diagnosis of AD relies on clinical criteria (McKhann et al., 1984).

Currently, the clinical diagnosis of AD is made according to the Revised Diagnostic and Statistical Manual of Mental Disorders (DSM III-R) (American Psychiatric Association, 1987) and the criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) (McKahn et al., 1984). Kukull et al. (1990) have reported the sensitivity of the DSM-III to be .76 and its specificity to be .80, and those of the NINCDS-ADRDA to be .92 and .65, respectively. The NINCDS-ADRDA criteria seem to be more adequate than those of the DSM III-R because they recognize the problems inherent in diagnosing AD and so specify three categories of diagnosis: possible, probable, and definite, with the last category requiring histopathological confirmation. A "possible" AD diagnosis is used when the course of the disease is abnormal. A clinical diagnosis of "probable" AD is used when the disease has an insidious onset with progressive deterioration and when other systemic or neurological diseases are excluded. In one longitudinal study which used the inclusion and exclusion criteria of the NINCDS-ADRDA for "probable AD", the criteria were found to be 100% accurate in diagnosing 26 AD patients as verified at autopsy (Morris, McKeel, Fulling, Torack & Berg, 1988).

Histopathologically, the disease is characterized by neurofibrillary tangles, neuritic plaques and granulovacuolar degeneration, although this pathology is also evident in other disease states (Van Hoesen & Damasio, 1987). All types of cortex are affected in AD. The primary sensory and motor areas are usually spared, however, and the degree of damage to the association cortices can vary from complete involvement in some patients to damage in only one or two lobes in other patients (Van Hoesen & Damasio, 1987). Subcortical structures are also affected. For example, the number of neurons is diminished in the nucleus basalis of Meynert and neurofibrillary tangles and neuritic plaques are found in the amygdala and basal ganglia (Morris & Kopelman, 1986). These pathologies are also found in the hippocampus which could account for some of the memory impairments observed on neuropsychological testing (Morris & Kopelman, 1986). The cause of these pathological changes in AD is unknown.

Martin and Bryant (1988) have provided evidence to support a genetic basis. These authors report an 8% concordance rate for dizygotic twins versus 42.8% concordance rate for monozygotic twins. Other possible factors currently under investigation are neurotransmitter or neurochemical deficits, viral agents, excessive accumulation of toxins in the brain such as aluminum, and changes in the autoimmune system (Bayles et al., 1987; Cohen, 1983). Neurochemical changes include reduction in acetylcholine-related parameters (Van Hoesen & Damasio, 1987; Whitehouse & Kellar, 1987).

Whether or not AD constitutes a heterogeneous disease is subject to debate. The common findings of neurofibrillary tangles, neuritic plaques and neuronal loss in the cortex at autopsy suggest to some researchers that it is a single disease, however, this argument is not convincing since these pathological markers are also found in other degenerative dementias. The idea that AD is a heterogeneous disease has been based on variation in the age of onset, labelled "senile" if the onset is after 65 years and "presenile" if onset is before this age (Gottfries, 1985). This view has been challenged and it has been suggested that the two should be grouped together based on their similar neuropathology (Alexander & Geschwind, 1984; Gottfries, 1985) as well as descriptive epidemiological data (Rocca, Amaducci & Schoenbert, 1986). Rocca et al. (1986)

reported that age-specific rates for AD show a smooth exponential increase after age 40, and are not bimodal which would be expected if presenile AD and senile dementia were two diseases.

Mayeux, Stern and Spanton (1985) reviewed the records of 121 AD patients whose age at onset of symptoms ranged from 43 to 86 years. The frequency of onset at a particular age was distributed bimodally and it was reported that patients who were younger at onset did not differ from those who were older with regard to duration of symptoms, education, functional activity performance, and scores on the modified Mini-Mental Status Exam. However, based on records of neurologic, neuropsychological and psychiatric assessments they were able to distinguish four subgroups of AD: extrapyramidal, myoclonic, benign and typical. Most of the patients were in the typical group where dementia and functional impairment progressed over the four-year period of the study.

The progression of AD is commonly evaluated using the seven stages of the Global Deterioration Scale (GDS) for Age-Associated Cognitive Decline and Alzheimer's Disease (Reisberg, Ferris, De Leon & Crook, 1982). In stage 1 there is no cognitive decline detectable and in stage 2 there is a very mild cognitive decline, that may not be detected by formal testing. There is objective evidence of cognitive decline in stages 3 and 4, with a mild decline in cognitive abilities and the earliest clear evidence of impairment present in stage 3. At this point, the individual begins to manifest decreased performance in demanding employment or social situations. In stage 4, also called the late confusional phase, there is moderate cognitive decline and the person experiences difficulty with concentration and remembering current events. However, an individual in stage 4 is still capable of travelling to familiar locations. In stage 5 and 6, or the phases of early and middle dementia, the patient can no longer live without assistance. At stage 7, or late dementia, the person loses verbal abilities and is almost always in institutional care. The stages of the GDS have been used in research to classify people into healthy control subjects (GDS 1 and 2), mildly to moderately impaired AD subjects (GDS 3 and 4), severely impaired AD subjects (GDS 5 and 6) and very severely impaired AD subjects (GDS 7) (Flicker, Bartus, Crook & Ferris, 1984).

In summary, the diagnosis of AD is based on excluding other degenerative and reversible dementias and can be confirmed only with biopsy or at autopsy. Several potential causes of AD are under investigation and there is clinical evidence to suggest that subtypes of AD exist. Cognitive deficits vary, further supporting the hypothesis that the disease is heterogeneous. The following is a review of cognitive abilities that decline in AD patients.

1.2 Cognitive Functions in Persons with Alzheimer's Disease

There are large individual differences in the early signs of AD (Breitner, Foldi, Rabins, Sunderland & Butler, 1987). Some patients may present with predominantly visual-spatial problems, such as getting lost in previously familiar environments, whereas other patients may complain of language difficulties, such as finding words for common objects or names of relatives. Although any aspect of higher cortical function can be compromised, memory deficits are always present (Kaplan & Sadock, 1981). The clinical features of AD can be cognitive, functional and behavioural in nature. Descriptions of the changes are based on the patient's self-reported history and information obtained from the patient's relatives. As well, signs of AD can be detected through a clinical evaluation.

1.2.1 Primary and working memory

Alzheimer's disease affects most aspects of memory functioning including primary or working memory (Morris & Baddeley, 1988) and secondary, or long-term memory (Martin et al., 1985). Patients with AD show deficits of primary memory as tested by the serial recall of digits (Kopelman, 1985), letters (Morris, 1984) or words (Corkin, 1982). Kopelman (1985) compared the performance of patients with AD, Korsakoff's syndrome and healthy control subjects on a digit-span task and on the Brown-Peterson task. The latter task requires subjects to retain three items in memory during variable periods of distraction. The AD group differed from the other two groups in showing diminished performance on both tasks. Performance on the Corsi block span test, a nonverbal (spatial) analogue of the verbal span test, is also impaired in AD patients (Corkin, 1982).

The working memory model has been useful in explaining the primary memory deficits seen in AD (Baddeley, Logie, Bressi, Della Sala & Spinnler, 1986; Morris & Baddeley, 1988). This framework assumes that there is a central executive system (CES), with limited capacity, which is responsible for the maintenance, rehearsal and storage of information and which coordinates and controls two slave systems (Morris & Baddeley, 1988). Although the slave systems, named the visuospatial scratch pad (VSSP) and articulatory loop, are regulated by the CES, they function relatively independently in terms of processing and storage (Baddeley, 1992; Farmer, Berman & Fletcher, 1986; Morris & Baddeley, 1988; Smyth, Morris, Levy & Ellis, 1987). The articulatory loop consists of a passive phonological input store and an active articulatory rehearsal process (Farmer et al., 1986). The VSSP is analogous to the articulatory loop, providing temporary storage and maintenance of visuo-spatial rather than verbal material

(Farmer et al., 1986). Dysfunction of the CES leads to an impaired ability to coordinate and perform two concurrent tasks (Morris & Baddeley, 1988).

Baddeley (1992) and Morris and Baddeley (1988) have provided evidence that AD selectively involves the CES. These researchers have shown that with the progression of AD, performance on individual phonological tasks remains intact whereas performance on combined tasks deteriorates when compared to that of young and old control subjects. An impaired CES could also affect performance of the slave systems. Hence, the impaired primary memory of AD patients as assessed by the Corsi-block span or the Brown-Peterson test may be attributed to a dysfunctional CES rather than an impairment of either slave system.

1.2.2 Attention

Adequate attention is central to the cognitive assessment of AD patients, since virtually all behavioural testing is heavily dependent on this (Mohr, Claus, Mann & Chase, 1991). Memory deficits have been attributed to impaired attentional abilities (Baddeley, 1992; Morris and Baddeley, 1988). In addition to the maintenance, rehearsal and storage of information, the CES has also been proposed to be responsible for attention.

Attention has been tested, using the working memory model, by having AD patients and old and young control subjects perform combined tasks (Baddeley et al., 1986). The experiment involved performing a primary tracking task concurrently with each of three secondary tasks which were articulatory suppression, reaction time to tones and memory span. The primary task required subjects to use a light-sensitive pen to track a moving white square on a coloured monitor screen. No differences were observed between the old and young control groups. The

AD patients were impaired in their abilities to perform two of the three secondary tasks, namely, the simple reaction time task concurrently with the tracking task and the digit span task concurrently with the tracking task (Baddeley et al., 1986). In another study it was found that as the disease progressed, performance on the individual tracking and digit span tasks was unaffected, however, performance on combined tasks deteriorated, lending support to the idea that there is a deficit in the attentional capacity of the CES in AD (Baddeley, 1992).

Attentional skills have also been examined independently of working memory. The tasks studied include visual and auditory selective attention, visual pursuit tracking, visual search and divided attention. Foster, Behrmann and Stuss (1992) have documented visual attention deficits in patients with AD. These investigators used two visual search tasks in which a specified target had to be detected in a field of distractors. One task called for automatic, non-effortful processing while the other required controlled, effortful processing. In the former, subjects searched for a target which was a filled circle among a number of distractors which were empty circles. The task requiring effortful processing involved searching for the same target among background distractors consisting of empty circles and filled squares. The number of distractors varied in each task. The AD patients showed impaired performance on both tasks. Performance on the automatic task was slowed to the same relative extent regardless of the number of distractors, which suggested that the impairment was probably due to psychomotor slowing. However, the speed of performance on the controlled task decreased with increasing number of distractors, which indicates that AD patients have impaired visual selective attention on tasks requiring effortful processing (Foster et al., 1992).

Clinicians often employ quick methods to assess attention. For screening purposes, Mohr et al. (1991) recommend the digit span test. This test can be made more difficult and sensitive by having patients recall digits backwards. Sustained attention can be assessed by a test such as "serial sevens", where the patient is asked to subtract 7 serially starting with 100 (Mohr et al., 1991). Attention can also be tested using cancellation tests where subjects are required to scan a sheet of paper containing line drawings of geometric shapes or common objects identified as either targets or distractors, and the task is to cross out the targets. Foldi, Jutagir, Davidoff and Gould (1992) administered nine cancellation tests of varying density (number of items on page) and complexity (number of different distractor shapes) to groups of AD patients, elderly control subjects and depressed patients. All of these tests contained line drawings of geometric shapes. The investigators found that the AD group was significantly impaired in comparison to the other groups, and that the performance of all groups declined with increasing density. The AD patients' deficits were partially attributed to an impaired ability to encode and process an increasing amount of information (Foldi et al., 1992).

In short, AD patients are impaired on tasks of primary memory. This deficit can be attributed to a deterioration of the CES in working memory. The CES is responsible for the attentional processing required to perform dual tasks. The performance of patients on cancellation tests has also confirmed attentional deficits in AD patients.

1.2.3 Secondary memory: Declarative memory

Secondary or long-term memory can be categorized into declarative (memory for facts and episodes or "knowing what") and procedural memory (memory for skills or "knowing how") (Squire, 1987). Declarative memory includes episodic memory which can be further divided into implicit and explicit memory (Graf & Schacter, 1985). Both involve memory over longer periods of time than primary or working memory. Implicit memory, also called "activation" or "priming", refers to the extent to which prior exposure to stimulus material facilitates the recall of other material. This facilitation need not involve conscious recollection (Morris & Kopelman, 1986). Implicit memory has been assessed using word-stem completion tasks where the subject completes a word after being given the first three letters. Morris, Wheatley and Britton (1983) found that while AD patients were impaired on a recognition memory task, performance on the word-stem completion task was normal. The findings of other studies have also suggested that implicit memory, as assessed by repetition-priming tasks, is normal in early AD patients. Repetition-priming tasks require the subject to judge whether a string of letters forms a word (Scarborough, Cortese & Scarborough, 1977) or to judge whether or not a word has been presented previously in the experiment (Moscovitch, 1982).

Explicit memory, also called "episodic" or "recent" memory, refers to memory for personally experienced events (Morris & Kopelman, 1986). While implicit memory appears to be spared in the early stages of AD, explicit memory is profoundly impaired. Deficits in explicit memory have been observed on verbal tests. For example, Pappas et al. (1992) conducted an experiment in which subjects listened to 25 short sentences. The subjects were then asked to listen to the sentences again and this time, they were required to complete the sentence by providing the last word. AD patients showed significant impairments on this test in comparison to healthy control subjects. Due to its early deterioration, explicit memory, as tested on verbal tasks, is often included in clinical screening tests. For example, the Mini-Mental State Examination (MMSE) (Folstein, Folstein & McHugh, 1975) contains one item which asks patients to repeat three words and to recall these words after the next test item, which serves as a distractor.

A decline in explicit memory has also been reported on nonverbal tests in subjects with AD. Impaired performance has been found on tests of recognition memory for spatial and visual patterns (Sahakian et al., 1988), recognition memory for spatial location (Aldelstein, Kesner & Strassberg, 1992; Kesner, Aldelstein & Crutcher, 1989) and recall for location of objects (Crook, Ferris & McCarthy, 1979) after various delay intervals. Similar impairments have been found on a visually-guided maze test. This test involves using a stylus on a 13 by 13 or 10 by 10 matrix of bolt heads to learn a route by trial and error (Barker, 1931; Milner, 1965). Brouwers, Cox, Martin, Chase and Fedio (1984) reported that AD patients were significantly impaired on the maze test (10 x 10 array) in comparison to normal control subjects and Huntington's disease patients. Patients with AD also showed a significant impairment on a smaller version of the maze test (5 x 5 matrix) when compared to normal control subjects (Liu, Gauthier & Gauthier, 1991a).

1.2.4 Secondary memory: Procedural memory

Although working and explicit memory are impaired in early AD, there is evidence to suggest that procedural memory may be relatively intact at this stage. Using a visual finger maze test, Grosse, Wilson and Fox (1991) reported an impairment in explicit memory in a group of AD patients, compared to normal control subjects. However, procedural memory was not impaired in these AD patients because they learned the maze and showed transfer of training when presented with a new maze (Grosse et al., 1991). These investigators attribute this finding to a relative sparing of corticostriatal circuits.

1.2.5 Language

According to Flicker, Ferris, Crook, Bartus and Reisberg (1986), word finding difficulty is the most prominent characteristic of language dysfunction in patients who are in the early stages of AD. Functionally, it has been observed that AD patients lack specific content words in their speech and tend to use words of indefinite reference such as "they", "thing" and "it" (Cummings & Benson, 1986).

The Boston Naming Test of visual confrontation naming (Kaplan, Goodglass, & Weintraub, 1983) is frequently used to assess language in this clinical population (Baum, Edwards, Leavitt, Grant & Deuel, 1988; Knesevich, LaBarge & Edwards, 1986; Williams, Mack & Henderson, 1989). In this test, the subject is asked to name 60 line drawings of common objects. Equivalent 15- (Mack, Freed, Williams & Henderson, 1992) and 30-item versions have been validated for use with AD individuals and developed in order to decrease the demands on attention and concentration. In one study, investigators found the Boston Naming Test to be sensitive not only to age and AD but also to the severity of AD (Flicker, Ferris, Crook & Bartus, 1987). Similar impairments have been found in AD subjects when using verb responses rather than nouns. The Action Naming Test (Obler & Albert, 1979) requires subjects to provide one-word responses to describe what is happening in pictures which depict actions. When Bowles, Obler and Albert (1987) administered this test to AD patients and young and old control subjects, they found that there was a significant difference, with the AD group performing the worst. The results indicated that the AD patients made more errors by providing unrelated (as opposed to near synonyms or related) words suggesting an impairment with word concepts as well as naming.

Using another naming test, Rissenberg and Glanzer (1987) provided healthy young and old adults and AD patients with definitions of concrete and abstract words. The subjects were asked to supply the appropriate words. The AD patients showed a significant impairment on this task and performed worse on abstract items. Language impairments have been found on tests of verbal fluency where the subject is required to produce spontaneously words from a specified category (Martin & Fedio, 1983). Such a test is included in the Modified Mini-Mental State Examination (3MS) (Teng & Chui, 1987) where the patient is asked to name animals with four legs within a 30-second time limit.

Investigators have also examined other aspects of language such as reading comprehension, writing, and semantic abilities. Murdoch, Chenery, Wilks and Boyle (1987) assessed language, retention, reading and writing using the Neurosensory Centre Comprehensive Examination of Aphasia (Spreen & Benton, 1977) and spontaneous speech using the fluency subtest of the Western Aphasia Battery (Kertesz, 1980). AD patients were impaired on verbal expression, auditory comprehension, repetition, reading and writing in comparison to control orthopedic patients. The findings indicated that while syntax and phonology were intact in the AD group, semantic abilities were impaired (Murdock et al., 1987). Similarly, Cummings, Houlihan and Hill (1986) reported that early AD patients, but not advanced AD patients, had intact abilities to read letters, words and commands. However, reading comprehension declined with increasing dementia severity.

There is a possibility that visuospatial deficits confound the impairment of language abilities as tested on reading tasks or on the Boston Naming Test. However, the ability to read correctly without fully comprehending the meaning of what is read suggests that the disturbance of comprehension in AD is a linguistic deficit rather than a product of visuoperceptual disturbances (Cummings et al., 1986). There have been reports of patients presenting with either language or visuospatial deficits (Becker, Huff, Nebec, Holland & Boller, 1986; Eustache, 1992; Martin, 1987) suggesting that the two deficits are probably independent of each other.

In summary, it appears that language production during early AD is fluent but word retrieval, semantic access and comprehension difficulties are characteristic early signs of AD. These language impairments in AD patients do not appear to be related to a decline in visuospatial function although deficits in both functions can occur concurrently.

1.3 Spatial Orientation in Alzheimer's Disease

1.3.1 Definition

"Spatial orientation" describes a broad spectrum of spatial abilities. Tolman (1948) described spatial orientation as a person's cognitive ability to represent space accurately, to map environmental information and to determine the position of the person concerned within that representation. According to Tolman (1948), to be oriented is to have an accurate representation or cognitive map of the surrounding area. Unlike Tolman's definition which focuses on cognitive processes, other terms used to describe spatial orientation include a functional aspect such as "sense of direction" (Kozlowski & Bryant, 1977), "way-finding ability" (Passini, 1984a, b; Weisman, 1987) and "travel behaviour" (Martino-Saltzman, Blasch, Morris & McNeal, 1991). In this section, the term "spatial orientation" encompasses the cognitive as well as the functional aspects of spatial abilities. As with memory deficits in AD, spatial disorientation may be partially attributed to an impairment of selective attention (Bayles et al., 1987).

1.3.2 Spatial skills in Alzheimer's disease

Most patients with AD experience visuospatial deficits at some stage during the disease process (Bayles et al., 1987). Although the early clinical manifestations of AD may vary between individuals, spatial deficits are common in the early stages of AD (Branconnier & DeVitt, 1983; Cogan, 1985; Cummings & Benson, 1986; Tariot et al., 1986), and seem to be an inevitable feature as the disease progresses (Henderson, Mack & Williams, 1989).

Impaired visuospatial abilities do not seem to result from poor visual acuity because the impairments can exist with good visual acuity (Cogan, 1985, 1987; Mendez, Mendez, Martin, Smyth & Whitehouse, 1990; Nissen, et al., 1985). Further, although visual acuity declines with normal aging, the decline is not accelerated in the early stages of AD (Fozard, Wolf, Bell, McFarland & Podolsky, 1977). Various visuospatial deficits have been reported in AD patients. Many of these deficits can be classified into six categories which will be examined here. These are figure-ground discrimination, personal orientation, extra-personal orientation, visuospatial constructional ability, mental representation and the functional aspect of spatial orientation.

1.3.3 Figure-ground discrimination

Cohn, Wilcox and Lerer (1991) recommend an assessment of visual perception in AD using tests involving discrimination of a simple figure or word "embedded" within a more complex visual figure. Figure-ground discrimination is typically assessed using the Embedded

Figures Test (Whitkin, 1950) or the Figure-Ground Perception Test, a subtest of the Southern California Sensory Integration Test (Ayres, 1972). Scores on the Figure-Ground Perception Test have been shown to be adequately correlated with those on the Embedded Figures Test (Spearman's rho correlation coefficient = -.67) (Petersen & Wikoff, 1983). The negative coefficient is explained by the different scoring methods used for the two tests. A lower score on the Embedded Figures Test indicates a better performance whereas a better performance on the Figure-Ground Perception Test is indicated by a higher score.

In a study comparing the performance of subjects on the Embedded Figures Test, Loring and Largen (1985) found that early-onset AD patients performed worse than control subjects, and late-onset AD patients. In another study, Mendez et al. (1990) administered a battery of "complex visual tests" to AD and control subjects. The battery included a test for figure-ground analysis which consisted of identifying three hidden figures (house, hand, hat) and three items from the Figure-Ground Perception Test. Other tests for skills such as visual recognition, form discrimination and spatial localization were also administered. The results indicated that AD patients and control subjects differed most on the figure-ground analysis. An impairment in figure-ground discrimination was found in all AD patients including those who were the least impaired on the MMSE and who had the shortest duration of the disease (six months) (Mendez et al., 1990). Other studies have shown that early AD patients are impaired on the complete Figure-Ground Perception Test (Liu et al., 1991a) and that the test is sensitive to the progression of the disease (Gauthier, Liu & Gauthier, 1990).

1.3.4 Personal orientation

Tests of personal orientation examine "personal or egocentric space" because the spatial relations are made with reference to one's own body (Butters & Soeldner, 1972; Ratcliff, 1982). Two common tests of personal orientation are left-right orientation and the Road-map test. Left-right orientation or discrimination tests commonly use verbal commands or pictures (Ayres, 1972; Benton, 1959; Culver, 1969). The commands generally involve pointing to positions on one's own body or to those of an examiner who sits opposite the subject (Ayres, 1972; Benton, 1959). In addition to commands, the test designed by Benton (1959) uses a full-length drawing of a boy and of the head and torso of a man. The test of Culver (1969) uses pictures of hands or feet in different positions.

Left-right discrimination ability has been tested in mild (MMSE: 16-23), moderate (MMSE : 6-15) and severe (MMSE: less than 6) AD patients and their performance compared to that of patients with multi-infarct dementia and normal control subjects (Fischer, Marterer & Danielczyk, 1990). These investigators used an abbreviated form of the verbal components of the Benton test and a doll was used in place of the examiner. Only the severe AD group was impaired on items referring to the subjects' own body. AD patients at every stage were significantly impaired on items relating to pointing to the doll. Liu et al. (1991a) reported similar findings in another study where early AD patients and control subjects were tested using the left-right discrimination subtest of the Southern California Sensory Integration Test (Ayres, 1972). This subtest contains ten verbal items, six referring to body parts on the subject and four referring to those on the examiner. Although there was no difference in the total scores of the AD and control groups, it was observed that most of the AD subjects' errors were made on items referring to the examiner. However, an error type analysis was not conducted to verify this (Liu et al., 1991a).

The ability to discriminate left and right has also been tested on a more difficult level using the Road-Map Test (Alexander, Walker & Money, 1964; Money, 1976). This test requires the subject to verbalize changes in the direction of a route drawn on a map. The Road-Map Test has a linguistic requirement (De Renzi, 1982) and involves mentally rotating one's body position (Brouwers et al., 1984). In a recent study, Bylsma, Brandt and Strauss (1992) reported that early Huntington's disease patients were significantly slower than normal control subjects, but there was no difference between the groups in the number of errors made. These investigators attributed the deficit to damage in frontal-striatal pathways that link the frontal cortex with the head of the caudate nucleus, which is affected in early Huntington's disease (Bylsma et al., 1992).

Patients with AD have been reported to be impaired on the Road-Map Test in comparison to control subjects (Brouwers et al., 1984; Liu et al., 1991a). Brouwers et al. (1984) have also found that female AD patients perform significantly worse than male AD subjects. Performance on the Road-Map Test has also been shown to deteriorate with aging and with the progression of dementia as measured on the GDS (Flicker, Ferris, Crook, Reisberg & Bartus, 1988).

In order to clarify the verbal, mental rotation and right-left discrimination requirements of the test, Flicker et al. (1988) administered the Boston Naming Test and the Road-map test concurrently to young normal, aged normal, early and advanced AD subjects. Both tests have linguistic requirements, and half of the Boston Naming Test items were rotated 180 degrees to introduce a rotation requirement to the test. The performance of both the aged control and early

AD groups declined similarly with the rotation requirements of both tasks, suggesting that early in the disease, AD produces no further impairment of spatial-rotation abilities than is produced by normal aging (Flicker et al., 1988).

It appears that a disorder of personal space with respect to parts of one's own body does not occur until the late stages of AD. However, distinguishing right from left, with respect to another person's body, or on the Road-Map Test, does deteriorate early in people with AD.

1.3.5 Extrapersonal orientation

Extrapersonal orientation has been defined as orientation in space using external objects as reference co-ordinates (Aubrey & Dobbs, 1989; Semmes, Weinstein, Ghent & Teuber, 1955). To study extrapersonal orientation, Semmes et al. (1955) devised the Map-Reading Test, a locomotor task requiring the subject to follow routes on a series of maps by walking on a larger floor version of the maps. The test can be visually- or tactually-guided. Deficient performance on the Map-Reading Test is associated with injury to either parietal lobe (Semmes et al., 1955), or to the posterior areas of the brain (Ratcliff & Newcombe, 1973; Semmes, Weinstein, Ghent & Teuber, 1963).

There appears to be several processes involved in extrapersonal orientation. Healthy elderly subjects have been found to perform with more errors and take longer to complete the Map-Reading Test than young control subjects (Aubrey & Dobbs, 1989). This age difference in performance has been attributed to decreased working memory (central executive) and mental rotation abilities. Aubrey and Dobbs (1989) argue that these cognitive functions are required in order to keep track of one's position while following the path, to hold in memory one's current location and the direction of path leading to it, to rotate mentally the map in order to make a correct left or right turn and to select the correct line to follow when more than one path converges on the dot.

Performance on the Map-Reading Test has also been studied with respect to personal orientation comparing subjects with localized brain lesions with control subjects who had peripheral nerve injuries (Semmes et al., 1963). Personal orientation was assessed by having subjects point to parts of their body which corresponded to those labelled on diagrams of ventral and dorsal views of a human figure. In this study, both personal and extrapersonal orientation were impaired in subjects with lesions of the posterior part of the left hemisphere. In addition, anterior lesions (particularly those of the left hemisphere) resulted in an impairment of personal but not extrapersonal orientation, whereas the converse was the case for right posterior lesions (Semmes et al., 1963). It is unlikely that the impaired performance of subjects with brain lesions could be attributed to problems of vision or language, however, interpretation of performance on the Map-Reading Test should also take into account the attentional requirements of the test (Semmes et al., 1963).

Bylsma et al. (1992) also studied the performance of Huntington's Disease patients on the Map-Reading Test with respect to personal orientation as measured on the Road-Map Test. Results of the latter were discussed previously. These investigators used two conditions for the Map-Reading Test. In the TURN condition, subjects were required to face the direction they were walking whereas in the NO-TURN condition, subjects always faced a wall identified as the north wall. The Huntington's Disease patients were impaired on the TURN conditions, which was attributed to a deficit in mental rotation, as verified by their performance on the Road-Map

Test. However, the patients were not impaired in the NO-TURN condition which the authors argured was a true measure of extrapersonal orientation. Based on these findings, Bylsma et al. (1992) hypothesized that mental rotation is affected by frontal cortex lesions while extrapersonal orientation is compromised by parietal lesions.

Only one study to date has involved AD patients in a study of extrapersonal orientation using the map-reading test (Liu et al., 1991a; Liu, Gauthier & Gauthier, 1991b). Early AD subjects performed significantly worse than healthy control subjects in this study. Although an error-analysis was not done, the AD patients were observed to perform errors relating to the following: finding the starting circle on the hand-held floor maps, discerning which route to take on converging circles, concentrating on the task, remembering instructions, remembering which routes had already been taken and remembering to use the external reference point. One could speculate that the impaired performance of the AD subjects could be attributed to any number of deficits including those related to mental rotation, extrapersonal orientation, memory, attention and visuospatial abilities.

1.3.6 Visuospatial constructional ability

Visuospatial constructional ability is required in many activities of daily living such as drawing, putting together a puzzle or assembling a bicycle or household item (Spreen & Strauss, 1991). It can be assessed using free-hand drawing, copying, drawing from memory and copying schematic drawings using three-dimensional blocks. These tests are sensitive to injury of either the left or right cerebral hemisphere in adults. Patients with left-hemisphere damage accurately depict the spatial relations among objects but tend to omit detail (De Renzi, 1982), whereas those with right-hemisphere damage include detail, but fail to maintain a coherent spatial organization among elements in their drawings (Ratcliff, 1982). As with the other spatial abilities discussed, drawing tasks are complex and multidimensional measures which involve visuospatial as well as executive abilities (Barr, Bilder & Kaplan, 1990), and a memory component if the subject is asked to draw from memory.

Initial free-hand drawing and copying of two- and three- dimensional figures, as well as clock drawing, are useful in the diagnosis of dementia (Henderson et al., 1989; Mohr et al., 1991; Strub & Black, 1985). Free-hand drawing, but not copying a cube, is sensitive to normal aging (Plude, Milberg & Cerella, 1986), whereas copying two intersecting pentagons (two-dimensional) has been found to be useful in screening AD patients (Teng & Chui, 1987).

Standardized qualitative and quantitative scoring methods have been developed in order to understand better the skills involved in drawing (Brantjes & Bouma, 1991; Kirk & Kertesz, 1991). These methods have been applied to drawing tasks such as spontaneous copying or drawing figures such as a house, clock, tree, person, geometric shapes or other common objects from memory. Brantjes and Bouma (1991) have categorized drawing skills into three components: spatial-perceptual, conceptual and executive-motor. When examining spatial-perceptual components, the drawings of AD patients display characteristics similar to those seen in patients with either left- or right-hemisphere lesions in that they may be scattered, poorly related spatially and oversimplified. However, drawings of AD patients differ in that there is no evidence of unilateral neglect (Brantjes & Bouma, 1991; Kirk & Kertesz, 1991). Conceptually, the drawings of AD patients exhibit a reduced number of ideas and an increase in the number of repetitions. The perseveration frequently seen in the drawings of persons with AD has been identified as a failure of the executive-motor component.

Recently, clock drawing and copying have gained popularity as useful screening tools for AD (Wolf-Klein, Silverstone, Levy, Brod & Breuer, 1989), and for the assessment of visuospatial construction (Henderson et al., 1989). Scoring methods vary but have been clearly documented (Sunderland et al., 1989; Wolf-Klein et al., 1989; Rouleau, Salmon, Butters, Kennedy & McGuire, 1992). Clock drawings can be used to distinguish people with early AD from healthy control subjects (Doyon, Bouchard, Morin, Bourgeois & Côté, 1991; Mohr et al, 1991; Sunderland et al., 1989; Wolf-Klein et al., 1989) and they demonstrate high sensitivity and specificity (Doyon et al., 1991; Shulman, Shedletsky & Silver, 1986; Wolf-Klein et al., 1989). Sunderland et al. (1989) have reported high inter-rater reliability for clock drawing tests and good validity coefficients when scores on this test are correlated with three global measures of dementia. In addition, Shulman et al. (1986) found the clock drawing task to correlate well with scores on the MMSE. Qualitative analysis of clock drawing tasks provides useful information concerning the underlying processes of visuospatial skills such as the ability to organize spatially the numbers (Rouleau et al., 1992; Shulman et al., 1986).

1.3.7 Mental representation

The information one has about one's environment has been referred to as a "mental representation" or "cognitive map" (Kozlowski & Bryant, 1977; Passini, 1984a; Tolman, 1948). Knowledge of an environment has been indirectly measured using a variety of techniques including pointing to directions and locations of landmarks out of view, estimating the distance and travel time between landmarks (Kozlowski & Bryant, 1977), performing way-finding tasks in real environmental situations (Kozlowski & Bryant, 1977; Passini, 1984b), viewing and identifying slide projections of the environment in question (Allen, Siegel & Rosinski, 1978; Weber & Brown, 1978) and by having subjects sketch maps (Passini, 1984a, 1986; Rovine & Weisman, 1989). These methods have been used with healthy elderly and young adult subjects as well as with brain-injured subjects to understand the use of environmental information for way-finding. Inaccurate sketch maps (McFie, Peircy & Zangwill, 1950; Ross, 1980) and floor plans (Ross, 1980) have been reported in association with topographical disorientation in patients with focal right-hemisphere lesions and bilateral posterior damage. Cogan (1985) described an AD patient who had normal visual acuity but whose disorientation caused him to stop driving. This patient was able to draw an outline of a map describing an area with which he had a life-long association, but he was found to be unable to place the major cities correctly. The use of sketch maps has been used only recently in studies involving normal subjects and controlled studies involving AD patients are scant.

Sketch map tasks are easy and quick to administer and can provide relevant information about a patient's mental representation that other tests cannot. Visuospatial constructive and memory skills are components of the ability to represent mentally features of the environment. The characteristics of sketch maps have been examined by Appleyard (1970) and Rovine and Weisman (1989). Rovine and Weisman (1989) conducted a study which controlled for different environmental experiences. In that study, university students were taken on a tour of a 12-block area of a small town. The subjects learned a common set of elements consisting of 20 buildings. Subsequently, they performed a sketch map task which required them to draw a map of the area and to include all the buildings pointed out along the tour. This was followed by a way-finding task where the participants were asked to take the investigator to eight of the target destinations using the most direct route. Upon analysis, the maps were found to be of two levels of complexity: sequential or spatial. According to these investigators, "sequential maps connected buildings by a continuous path, thus ignoring the gridiron structure of the area traversed. Spatial maps, by contrast, showed at least some knowledge of inter-relationships among paths travelled" (p. 223). The sketch maps were rated according to 1) the frequency of landmarks, path segments and path intersections, and 2) the topological accuracy of placement of the 20 target buildings, or their relationships to each other. Way-finding was also measured in terms of distance travelled, total turns and whether it was the best route. When compared to other pencil-paper measures of self-efficacy, visualization, orientation and sense-of-direction, sketch map measures, particularly the accuracy score, were the best predictors of way-finding performance. Based on these findings, it was suggested that sketch maps could contribute relevant information on mental representation in other populations.

In contrast to studies that examine mental representation in a macrospatial context, such as cities and neighborhoods, Moeser (1988) studied cognitive mapping of the interior of a building. The building was a hospital and was selected based on its unique configuration and on reports that it was very difficult to navigate in it. Student nurses who had worked in the hospital for various periods of time were asked to draw a floor plan of four floors of the hospital and to include as many details as possible. These sketch maps were scored by comparing them to floor plans supplied by the hospital administration which resembled survey maps. The subjects drew landmark maps (placing rooms in relation to each other without indicating connecting corridors) or route maps (drawing a few corridors and rooms that were located along these routes). No one drew survey maps with a complex set of corridors.

The maps were scored by giving one point to each label of an item that was placed appropriately in relation to the others. There was a total number of 85 items which included rooms, entrances and elevators. The general findings were that students with 25 months' experience performed better on the labelling task than students with 4 months' experience but they were no different from those with 11 months' experience. In addition, none of these students formed survey-type cognitive maps as would be expected of people who are experienced with a setting. Moeser (1988) concluded that either mental representations of survey maps do not develop automatically or that the building layout was too complex, and that people automatically develop cognitive maps of the survey type for simpler environments.

To the author's knowledge, only one similar study has been conducted with AD patients (Liu et al., 1991a). In a laboratory setting, early AD and healthy control subjects were asked to sketch maps of their homes or a "familiar environment". Subsequently, they were taken on a tour of the main level of a university building denoted as a "new environment". The subjects were then required to perform a way-finding task by leading the investigator to each of five target rooms in the new environment. This was followed by sketching maps of the route taken to the five target rooms. Sketch maps of familiar environments were verified by the investigator during home visits at which time the subjects also led the investigator to each of the rooms in the familiar environment. The sketch maps were rated according to the number of correctly identified rooms and the relationships of these rooms, much like the frequency and accuracy measures used by Rovine and Weisman (1989). The AD patients were impaired in their way-finding ability in the new environment, but not at home. These patients were also impaired in mentally representing both environments and obtained lower scores for the relationship between rooms than for the number of rooms in both sketch maps. These data suggest that the impairment of mental representation observed in AD patients may be a result of impaired recall whereas intact way-finding at home may be attributed to intact recognition memory. More comprehensive scoring and analyses of sketch maps could provide helpful information about the processes underlying this impairment in AD.

1.3.8 Functional spatial skills: Driving ability

The initial signs of AD are usually first noticed by the patient him-/herself and the immediate family members. As a result of changes in memory for recent events, word-finding abilities, attention span or spatial orientation, activities of daily living (ADL) which require these abilities are inevitably affected. Driving, way-finding and wandering are three aspects of functional spatial orientation that will be addressed.

The capacity to drive is of particular concern to health professionals and caregivers of patients who are in the early stages of AD (Chenoweth & Spencer, 1986; Weintraub & Kapust, 1989). One retrospective survey involving 30 AD and 20 healthy control subjects revealed that 47% of AD patients, in contrast to only 10% of control subjects, had experienced at least one accident in the previous five years (Friedland et al., 1988). An odds ratio of 7.9 was estimated for accidents involving patients with AD. These accidents occurred with patients in the early or middle stages of the disease and were not related to driving experience or alcohol intake. Although tests of memory, attention, visuospatial construction and reaction time were administered to these subjects, the results were not reported. A greater accident rate among persons with AD compared to control subjects has also been found in a more recent study which also used a retrospective survey (Dubinsky, Williamson, Gray & Glatt, 1992).

In another survey involving 72 outpatients with dementia, 53% of whom were diagnosed with AD, 30% had had at least one accident since the onset of dementia symptoms (Lucas-Blaustein, Filipp, Dungan & Tune, 1988). This survey found that demented patients who had had accidents since the onset of illness were as likely to still be driving as those who did not have accidents, but were more likely to become lost (Lucas-Blaustein et al., 1988).

While some investigators advocate that all AD patients stop driving (Friedland et al., 1988), others alert clinicians to the legal and moral implications of such a severe action (Drachman, 1988; Lucas-Blaustein et al., 1988). The capacity to drive depends on complex abilities including decision-making, memory, visuospatial skills, judgment and attention (Friedland et al., 1988), as well as on having a quick reaction time and adequate psychomotor skills. Further prospective studies using randomly selected samples, including the general population, are needed to establish accurately the relative risk of accidents in AD patients (Kokmen, 1989), and to determine the predictors of declining driving capacity in early AD patients.

1.3.09 Functional spatial orientation: Way-finding and wandering behaviour

Way-finding has been introduced in the previous section in the context of mental representation. Disorders in way-finding, also called topographical or environmental disorientation, usually occur as a result of global impairments of cerebral functioning, but have also been reported in patients with localized brain lesions (Habibi & Sirigu, 1987).

Topographical disorientation is defined as the inability to recognize, navigate or find one's way in an environment that is familiar to the person (De Renzi, 1982; Habib & Sirigu, 1987; McFie et al., 1950). This disorder has been observed in patients with lesions in the right hippocampal region, right occipital lobe with thalamic involvement (Habib & Sirigu, 1987), as well as in a patient with Dyke-Davidoff-Masson Syndrome which is characterised by developmental atrophy and porencephaly involving the right frontal and parietal lobes (Fine, Mellstrom, Mani & Timmins, 1980). Ratcliff and Newcombe (1973) reported that topographical disorientation results from bilateral lesions involving the posterior aspects of the parietal, temporal and occipital lobes. Using the Map-Reading Test, these investigators found that subjects with bilateral lesions were significantly more impaired than subjects with unilateral lesions involving the same areas of the brain. The Map-Reading Test involves topographical orientation in that it requires a subject to "monitor changes in one's position relative to the environment" (p. 453) (Ratcliff & Newcombe, 1973). Although most authors report lesions involving the posterior right hemisphere in patients with topographical disorientation, the exact anatomical correlate of this disorder is not known. Most of these cases involve patients with hemianopia or quadrantopia which are not generally seen in AD patients. Thus, environmental disorientation in AD patients is probably a result of different underlying processes.

Way-finding is similar to driving in that decision-making, attention, memory and visuospatial skills are all required in order to succeed. In studies involving healthy subjects, way-finding has been treated as a "problem-solving" skill (Downs & Stea, 1981; Passini, 1984a, b; Rovine & Weisman, 1989; Smyth et al., 1987). Downs and Stea (1981) propose four sequential stages that one uses for way-finding. These are: 1) to orient oneself by establishing a mental representation of the environment; 2) to select a destination and establish a connection between where one is and where the destination is in space; 3) to maintain the necessary direction or relationship between the mental map and the environment as well as execute necessary actions and decisions, and 4) to discover the destination by recognizing it as such. Passini (1984a) offers a similar description of the components involved in way-finding, however, he proposes that these components are iterative and not sequential.

The use of sketch maps to measure way-finding ability was discussed in the previous section. Based on results obtained from sketch maps, Rovine and Weisman (1989) suggest that there are differences in strategies used for assimilating the vast amount of environmental information required for navigation. They found that successful way-finders attend to the spatial nature of the environment as opposed to memorizing the route. In another study involving elderly persons residing in a nursing home, Weisman (1987) observed four way-finding strategies. These are: 1) the use of "perceptual access", also called landmarks or goals, to guide one's trip; 2) "blind navigation" or following a trail such as coloured lines running along corridor floors or walls; 3) the use of signs and landmarks to update information and clarify choices where decisions are to be made, and 4) developing a mental image or cognitive map of the environment. This last strategy would appear to be the most effective and comparable to attending to the spatial nature of the environment.

In addition to intact cognitive abilities, successful way-finding is also largely dependent on environmental factors. For example, way-finding is facilitated if landmarks are available and if the destination is visible (Carpman, Grant & Simmons, 1985; Weisman, 1987). Weisman (1981) reported that the complexity or layout of 10 university buildings accounted for 56% of the variance in way-finding, whereas experience or familiarity with the building accounted for only 9%. Even interior decorating strategies seem to affect way-finding. Evans, Fellows, Zorn and Doty (1980) found that way-finding is facilitated if walls are painted in distinct colours as opposed to a monochromatic beige. Environmental factors that affect way-finding ability contribute to the "legibility" of that environment (Weisman, 1981). This refers to how easily the environment is spatially understood.

Way-finding in healthy individuals can be considered a purposeful activity. Severe difficulties in way-finding may result in nonpurposeful activity called wandering. Wandering is most often observed among residents in institutions where dementia tends to be more severe. By filming nursing home residents Martino-Saltzman et al. (1991) identified three types of ineffective travel behaviour: lapping, pacing and random wandering. As yet, no controlled study has examined way-finding in AD patients using the technique of navigation on streets or in buildings or observation through filming. Such a study could measure locomotor skills which, in addition to cognitive skills, are also important in acquiring spatial knowledge (Allen et al., 1978). Further, the method would be "ecologically" valid, which refers to how well test results reflect one's abilities in everyday life (Plude et al., 1986; Wilkins, 1986).

An indirect method of evaluating way-finding is by using a questionnaire as has been done for obtaining information on driving abilities. In the study on environmental legibility, Weisman (1981) administered a self-report questionnaire containing 10 items. These items were specific to the university buildings under study and ranged from how frequently a person has been lost, to how confident a person would be in giving directions to a stranger or in drawing a sketch map of a building. It is not known how these self-reports relate to more objective methods of measuring way-finding. Kozlowski and Bryant (1977) found self-reports of one's own "sense of direction" to be accurate in predicting performances in pointing to unseen buildings, to the direction north, and map-completion. A visual analogue scale was used for these self-reports.

To summarize, way-finding, as studied using navigation tasks and observing travel behaviour, is a complex activity requiring the integration of many cognitive abilities in addition to locomotor skills. These methods are ecologically valid but are difficult to use with AD patients unless factors such as aspects of the physical environment, individual way-finding strategies and spatial impairments are concurrently addressed. In addition, they are time consuming. A questionnaire would be a more efficient way of obtaining information on the way-finding abilities of AD patients as has been done in studies on driving abilities in this population (Lucas-Blaustein et al., 1988).

The use of self-reports with AD patients raises the issue of judgement and the reliability of the responses. Using patient and proxy responses, McGlynn and Kaszniak (1991) report that AD patients substantially under-estimate their difficulties with cognitive tasks in everyday life. They also found that AD patients are inaccurate when predicting their performance on cognitive tasks and tend to over-estimate their memory abilities. However, Rocca et al. (1986) compared responses from AD patients and proxy responses using a questionnaire covering several topics that ranged from life habits and family history to medical history and found contrary results. These investigators found that the responses of surrogates were generally in excellent or good agreement with those obtained from the patient. Hence, the accuracy of responses may depend on the variable being measured. The accuracy of self- and proxy-ratings may also be affected by the type of report the subjects are asked to use. Baddeley, Sunderland and Harris (1982) compared responses from a questionnaire concerning memory with responses from diaries concerning the same items on the questionnaire. Self- and proxy-rated responses were compared using patients with acute and chronic head injuries and a control group. Baddeley et al. (1982) found that self-rating by questionnaire was not correlated with objective measures of memory, but proxy-ratings by questionnaire were. Also, the diaries of the chronic head injury and control groups were significantly correlated with objective measures as were the diaries of the proxies, but to a lesser extent. As yet, it is not known whether an indirect measurement of way-finding by a questionnaire would provide accurate information on actual way-finding abilities. In addition, although one may question the accuracy of self-ratings by patients with AD, a better understanding can be gained if self-ratings are compared with proxy-ratings and if both self- and proxy-ratings are compared with objective measures.

1.3.10 Summary

Spatial orientation is a multi-dimensional concept involving many aspects of cognitive function. It also involves environmental factors which are related to way-finding ability and driving capacity. Many tests have been developed to study spatial disorientation. Studies of the spatial abilities of patients with brain lesions have provided many neuropsychological tests that are useful for studying the AD population. In addition, studies involving healthy subjects have provided other methods and information about the processes involved in successful way-finding. Together, these studies serve as a framework for studying spatial orientation in AD. These skills can be categorized into different aspects of spatial function depending on their place within a hierarchy of skill. On a basic level, one could test left-right orientation and on a complex level, way-finding ability could be examined. For all capacities, memory, attention and language need to be considered as factors potentially associated with a spatial deficit.

On some clinical scales, such as the Dementia of the Alzheimer Type Inventory (Cummings & Benson, 1986) and the Brief Cognitive Rating Scale (Reisberg et al., 1983), spatial orientation is frequently classified with orientation to time and other ADL skills. In addition, many of the items on these scales involve only verbal responses from the subject, thereby possibly confounding the spatial skill with verbal abilities. There is no standard set of spatial tests one can use to understand the level of spatial impairment of an AD patient. Many of the spatial tests that have been used have unknown psychometric properties, despite their long standing use in neuropsychology. The next section addresses factors that need to be considered in the development or use of such tests.

1.4 Test Development

An important aspect of an instrument is whether it is a standardized test or a rating scale. A standardized test has a specified population of individuals as its interpretive frame of reference, that is, the norms reflect the characteristics of the samples used. These characteristics may be age, sex, years of education or cultural background. The norms are used to determine the level of performance in another group of individuals. For example, the Wechsler Adult Intelligence Scale (WAIS) is frequently used to compare older brain-injured individuals with each other and with the norms from a healthy aged population.

When interpreting results from a standardized test, it is important to consider the normative sample for which the test was designed (Anastasi, 1982). If the patient is not comparable to the original normative sample on variables such as age, sex, language, culture and socioeconomic status, the interpretation of the patient's score would not be a valid one. For example, the Porteus Maze Test (Porteus, 1959) was originally designed for and standardized on children. Although two subsequent adult test items were added, the test ages only reach a maximum of 17 years and normative data are only available for a maximum chronological age of 15 years. Although many of the spatial tests reviewed in the literature have standardized administration procedures, they are not standardized tests because normative data do not exist. Moreover, the standardized procedures developed for one type of population may not be appropriate for a population with dementia. For example, establishing time limits may result in floor effects in the scores, and excessively long instructions that require intact memory and attention may be too difficult for a person with dementia to comprehend.

An alternative to selecting a test that has been standardized using a general population is to find a test that has "specific" norms, that is, the test has been standardized on a narrowly defined population (Anastasi, 1982; Golden, Sawicki & Franzen, 1984). These can be "subgroup" norms, or "local" norms which are even more narrowly defined than subgroups, that is, one may want to use a test standardized on patients in a particular hospital or in all of the hospitals in a city (Anastasi, 1982; Golden et al., 1984). The disadvantage of using these alternative standardized tests is that generalization from the norms to the patient's score is limited, however, the advantage is that interpretation of the patient's performance may be more accurate and valid. Even if the test scores are normalized to subgroup or local norms, interpretations must be made with caution in heterogeneous geriatric populations.

The rating scale is more readily available to the clinician or researcher who is interested in differentiating AD patients from other geriatric patients. The rating scale is usually more descriptive and easier to construct and use than the standardized test. The Global Deterioration Scale (GDS) (Reisberg et al., 1982) is an example of a rating scale commonly used to discriminate AD patients from other patients and one stage of the disease from another. Unfortunately, rating scales tend to lack the rigorous psychometric evaluations that standardized tests are usually subjected to, and therefore reliability and validity are often major issues when the scales are used with an AD population. Since AD is a progressive disease, part of diagnosis involves reevaluation to detect disease progression. Therefore, it may also be desirable for the chosen rating scale to be examined with respect to its ability to detect significant change reliably over time. An instrument's ability to detect minimal clinically important differences is referred to as responsiveness by Guyatt, Walter and Norman (1987) who offer a formula to calculate such an index.

Prior to administering any measurement instrument, its basic properties of reliability and validity must be examined. When a battery of tests is developed, there is also the issue of internal consistency of the battery.

1.4.1 Reliability

Reliability refers to the performance consistency of the test as administered by the same user or by different users (Feinstein, 1987). Reliability involves two sources of variability: external and internal variability (Feinstein, 1987). Assessment of external variability involves intra-rater and inter-rater reliabilities. A reliable instrument is one in which the scores from subjects are reproducible over time and between raters. Establishment of external reliability should be done with raters who represent those who will eventually use the instrument, whether they be a clinician, caregiver or a patient. Usually, inter-rater reliability is measured during the same testing session whereas intra-rater reliability is measured over two sessions, hence the latter is also called test-retest reliability.

The test-retest method also assesses temporal stability or variance due to the passage of time because it involves more than one test session (Golden et al., 1984). Its limitation is the possibility of carry-over effects, that is, if the time interval is too short, the reliability coefficient may be an overestimate of the true stability due to memory or learning effects (Kline, 1979), for example, the subject may remember his or her responses thereby inflating the reliability. This method does not deal with the reliability of the instrument itself because the same instrument is used on the second administration. Instead, the focus is with subject variability and consistency of the users (Feinstein, 1987). When the test-retest method is used for cognitive tests, the time interval between tests should not be too long because, as the time interval increases, the reliability coefficient decreases (Anastasi, 1982). Nunnally (1978) suggests allowing two weeks to one month between tests in order to account for day-to-day differences.

Internal variability or internal consistency is the assessment of variance due to content sampling and heterogeneity of the test items (Golden et al., 1984). It is the expected correlation between an actual test and a hypothetical alternative form of the same length (Nunnally, 1978). Cronbach's Alpha (Cronbach, 1951) is the usual index of internal consistency. The value of an acceptable coefficient depends on the heterogeneity and number of subtests. An extremely high coefficient could suggest that the battery may be too narrow and too specific, which would reduce the validity of the test because the subtests are almost identical to each other. The internal consistency of a battery of tests depends on the inter-correlations between the subtests comprising such a battery. If the subtests measure the same construct, one would expect the battery to be internally consistent. However, if the construct involves heterogeneous skills, as reflected in the subtests, then this heterogeneity would lower the internal consistency of the battery. Another factor to consider is that Cronbach's alpha increases with the number of subtests of a battery (Nunnally, 1978).

1.4.2 Validity

Validity is concerned with what the instrument intends to measure. The instrument is validated in relation to the purpose for which it is used (Carmines & Zeller, 1979). Face, content, criterion and construct validity should be considered in selecting or developing an instrument to measure spatial skills in AD. Face validity concerns judgements about a test either on the part of the rater or the patient (Nunnally, 1978). This is important to consider when testing the elderly in order to minimize factors that may affect performance. The test should appear appropriate to the subjects taking it. If the test does not appear appropriate to the subjects, they may resist performing or they may not perform at their optimum levels. On the other hand, a researcher may wish to disguise the purpose of the test, or eliminate face validity in order not to bias the subjects' responses (Payton, 1980). Face validity also relates to how the examiner perceives the test. The criterion is seldom based on empirical evidence but on the subjective judgement of one or more experts as to whether or not a test looks reasonable (Streiner & Norman, 1989).

Content validity is closely related to the concept of face validity. Content validity consists of a judgement of whether the test samples all the important areas of the domain of interest. An instrument designed to measure spatial orientation should sample all of the important and stable aspects of this construct as applied to the AD and elderly populations. Since patients should perform at their optimum level, performance should not be timed unless it is actually speed that is of interest, such as reaction time measures. Usually the assessment of content validity involves consulting panels of individuals considered as "experts" (Spitzer et al., 1981; Thorn & Deitz, 1990). These experts may be patients, caregivers, relatives or health professionals depending on the purpose of the test (Spitzer et al., 1981). The experts provide feedback on the completeness, clarity and applicability of the instrument (Wood-Dauphinee, Opzoomer, Williams, Marchand & Spitzer, 1988). In the development of a quality of life index for cancer patients, Spitzer et al. (1981) used an a priori criterion for establishing content validity. The criterion used was that a majority of a panel (51%) should respond affirmatively to all five questions about each item of the index being developed.

Criterion validity is evaluated by correlating scores on the scale with some other measure of the variable under study (Streiner & Norman, 1989). It is preferable if the other measure is a "gold standard". In general, the rationale for criterion validity is to replace one measure, usually the gold standard, with a shorter, less expensive or less invasive one. Concurrent validity refers to the relationship between scores on a new measure and scores on the criterion measure which is administered at approximately the same time as the new measure. Predictive validity refers to the ability of the new measure to predict performance on a criterion measure which is assessed in the future. Although there are many tests of spatial skills, particularly in neuropsychology, few of these have been applied to the AD population. Those that have been used with AD patients have been used to gain knowledge about spatial deficits associated with AD and none has been referred to as the "gold standard". As it is not possible to compare a measure of spatial orientation to a gold standard, other standards must be used. A commonly accepted standard of rating the severity of dementia is the GDS. This scale uses descriptions of general behaviours to determine the level of disease progression, and these behaviours include spatial orientation skills. Thus, one method that could be used to assess the concurrent validity of a new measure of spatial orientation skills would be to compare performance on this new measure to scores on the GDS, not in terms of the new measure's ability to rate the severity of dementia, but in terms of its ability to rate the severity of spatial disorientation. The predictive validity of a new measure could be assessed, for example, by asking whether or not current scores on the new measure predict one's ability to drive an automobile in six months, or if the new measure predicts institutionalization of the individual in one year. A longitudinal study would be required in order to address the predictive validity of such a new measure.

Construct validity refers to the extent to which the new measure relates to other measures in a manner that is consistent with the theoretically derived hypothesis regarding the behaviour or ability being measured. Construct validity is related to content validity in that the test items must represent adequately the behaviour or ability of interest (Payton, 1980). A new measure possesses construct validity if it correlates to a certain extent with other measures that evaluate the same function or behaviour that it claims to evaluate. The rationale for construct validity is to use the underlying theory of an existing measure to help develop an instrument that better explains a certain behaviour under study. Given that the new measure could replace the use of the GDS for the purpose of describing spatial orientation, it could be argued that the method suggested for establishing criterion validity would also contribute to establishing construct validity.

1.5 Rationale for the Study

The literature contains many studies that have examined and measured various aspects of spatial orientation. Some of these have used tests that are designed to examine simple, perceptual spatial skills, while others measure more complex skills. Many studies suggest that the right posterior region of the brain is critically important for performance on spatial tasks. There is also evidence that indicates that spatial orientation depends on an intact ability to plan, concentrate, remember and integrate information, which implicates the frontal, temporal and parietal areas of the brain. When spatial skills are studied in healthy young and elderly subjects, it is apparent that spatial orientation involves processing at many levels, and that there are age and individual differences. Some skills decline with age and subjects of all ages use different strategies to orient themselves.

Interpretation of the results from tests of spatial performance in AD patients is challenging because the disease is global and affects the brain bilaterally. There is evidence to suggest that although both AD and other clinical populations, such as individuals with localized brain lesions or cerebrovascular accidents, experience spatial disorientation, the processes underlying the disorientation appear to be different. In order to study these processes, reliable and valid tests are needed. Although many spatial tests appear valid, most studies have not specifically addressed this issue. The same problem exists with respect to the reliability of spatial tests.

Finally, few studies have examined several levels of spatial function concurrently. This would permit a better understanding of the underlying processes contributing to an observed deficit.

A battery of spatial tests with established psychometric properties could serve several purposes. Clinically, the data from these tests could be used to support or question current strategies used in the management of disoriented AD patients (Bleathman, 1987; McGrowder-Lin & Bhatt, 1988; Munson et al., 1991 & Ohta & Ohta, 1988). Stable and valid subtests could potentially be included in diagnostic screening tests, clinical drug trials and, in future studies, as a predictor variable for more complex spatial behaviours such as the capacity to drive. The data could also help answer some theoretical questions. These questions pertain to what is the underlying process of spatial disorientation in AD patients, and whether there is a pattern of decline in certain skills as the disease progresses.

A preliminary study has been conducted to determine whether specific spatial orientation skills are affected by AD (Liu et al., 1991a). In that study, spatial orientation was operationally defined as performance on tests of basic (perceptual), higher cognitive and functional spatial skills. Tests were selected to examine skills in each of these categories. Basic spatial skills were assessed using tests of figure-ground discrimination, shape recognition, size discrimination, recognition of position in space, and left-right discrimination. Higher-order cognitive spatial skills were evaluated using pencil and paper mazes, the Map-Reading Test, the Road-Map Test, a test of spatial memory span and a visually-guided maze. Mental representation was tested by having subjects draw floor plans of their home or familiar environment and of a new environment, that is, the first floor of an unfamiliar building. Functional spatial skills in the familiar and new environments were tested by observing the subject navigate in both places.

The performance of a group of 15 early AD patients was compared to that of a group of 15 control subjects matched with respect to age, sex and years of education. The results indicated that the AD group showed deficits on tests of higher cognitive spatial orientation, including mental representation, but that certain basic spatial orientation skills were intact. On tasks of functional spatial orientation, the AD group was impaired in the new environment but their performance in familiar environments was normal. These findings suggest that in the early stages of AD, disorientation in new environments is related to deficits of higher-order cognitive spatial skills. These results provided a basis for the development of a test battery for assessing spatial skills in persons with AD. Based on the findings from this study, tests that were not sensitive to the presence of AD were eliminated, namely tests for shape and size discrimination.

1.6 Objectives of the Study

The global objective of this study was to determine the reliability and validity of a battery of subtests for the assessment of spatial orientation skills in persons with AD. This battery of subtests was based on those used in the preliminary study. Most of these tests have been described in the literature and are well known in research on spatial skills. For most of these subtests, normative data and information concerning their reliability and validity are lacking. For those that have psychometric data, the data pertain to other populations such as children, young normal adults or adults with focal brain lesions. Thus, the specific objectives of this study were to: 1) obtain normative data for these tests using healthy elderly subjects, 2) determine the external reliability or the test-retest and inter-rater reliability of each of these spatial tests when administered to patients with AD and to healthy elderly subjects, 3) determine the internal reliability or internal consistency of the battery, and 4) determine the content, construct and, to a limited extent, the criterion validity of this battery of tests as applied to AD patients. Since the global objective was to evaluate the subtests and not the battery as a whole, total scores were not analyzed except for the internal consistency study.

In order for the results from this study to have clinical relevance, another objective was to develop a preliminary shortened battery for possible use in the clinic. The criteria used for selection of subtests were that they possessed good psychometric properties, they had clinical relevance and that taken together, they represented the same range of spatial skills tested using the original battery. Since the purpose of the shortened battery was for clinical use, total scores were tabulated and then used for determining construct validity. The internal consistency of the shortened battery was also examined.

Given that AD involves changes in many aspects of cognitive function which may influence spatial abilities, a concurrent assessment of some of these aspects of cognitive function may help explain the performances on spatial tests. Thus, a related objective of the study was to determine whether AD patients were also impaired on objective tests of attention, naming ability, verbal memory and visuospatial constructive abilities. At the same time, normative data were collected for these tests because they were not available in the literature. Test-retest and inter-rater coefficients were also obtained for the clock drawing test which was used to evaluate visuospatial constructive abilities. This information was considered to be relevant given its frequent use with AD patients. Since AD affects global cognitive functioning, it was expected that AD patients would also be impaired on these other tests of cognitive function.

2.0 METHODS

This study examined the performance of 3 groups of subjects on 13 spatial orientation skills subtests and 4 nonspatial tests. This section describes the control, AD subjects in GDS 3 and 4, and AD subjects in GDS 5 who volunteered to be tested in this study. The data from the GDS 5 AD subjects were used in the analysis for establishing criterion validity. The mental status examinations and each of the 13 subtests of the SOS Battery are also described in detail. Adaptations were made to these subtests so that the AD patients could perform them. These adaptations are reported in addition to their known psychometric properties. This is followed by a description of the four nonspatial tests used to assess attention, language, verbal memory and visuospatial construction. This section then presents the procedures and statistical analyses used for determining content validity, test-retest and inter-rater reliability, internal consistency as well as construct and criterion validity of the battery.

2.1 Subjects

2.1.1 Normal Control Subjects

A total of 103 English-speaking, healthy subjects were initially recruited for the collection of normative data. Some of these volunteers were referred from a group of healthy volunteers being followed longitudinally at the Douglas Hospital Center. Volunteers were also recruited through community seniors groups and various chapters of the Golden Age Society. The spouses of some of the AD subjects were also invited to participate. The sample size of the AD group used for the validity study was intended to be 25, and so that of the normative sample was set at four times this, namely 100 subjects, in order to have sufficient numbers to describe the variation in the performance of the normal aged population.

Subjects were selected to be between the ages of 50 and 80 years. This range was chosen based on the report that general cognitive function in healthy men and women, as measured on the MMSE, appears stable between the ages 50 and 80 years but shows a decline after the age of 80 (Bleecker, Bolla-Wilson, Kawas & Agnew, 1988). Furthermore, this range was comparable to that of the AD group.

The normal control subjects had no history of neurological or psychiatric deficits, such as head injury or psychosis, that would have affected cognitive function. This was verified at the initial visit using a health status questionnaire (Furrie, 1987). Both the normal control and AD subjects were asked to answer socio-demographic questions on a questionnaire entitled "General Information" (Appendix B). The normal control subjects did not demonstrate any visual field deficits as tested by confrontation and visual acuity, with or without visual aids, was at least 20/50 when tested with the pocket Snell card. Each subject signed an informed consent form (Appendix C).

Each individual was administered the MMSE (Folstein et al., 1975) and 3MS (Teng & Chui, 1987). A minimum of 26 on 30 was set as the cut-off score for the MMSE. Subjects in this normative sample also had to be right-handed as assessed by the Edinburgh Handedness Inventory for which normative data, based on 1100 young and adult individuals, exist (Oldfield, 1971). The five-item version of the original 20-item questionnaire was used in order to decrease administration time (Bryden, 1976; Raczkowski & Kalat, 1974). This involved asking whether the subject used the left hand only, the right hand only or either hand to perform the following activities: writing, drawing, throwing (a ball), cutting with scissors and using a toothbrush.

Factor analysis has demonstrated that these five items load on the common factor of what is normally meant by handedness (Bryden, 1977).

Of the 103 control subjects who attended the initial visit, one man was excluded because he was 48 years old and therefore did not meet the criterion for age. An additional five subjects, one man and four women, were excluded because their MMSE scores were below the 26 cut-off score. The final control group was composed of 97 subjects and consisted of 36 men and 61 women. The group's mean age (standard deviation) was 67.7 (SD = 7.1) years with a mean of 13.7 (SD = 3.7) years of education. The mean scores on the mental status examinations were 29.2 (SD = 1.1) for the MMSE and 97.3 (SD = 2.8) for the 3MS.

From every block of three healthy elderly subjects evaluated in the normative study, two subjects were randomly selected. The first person drawn was used for the test-retest reliability and the second one for inter-rater reliability studies. All subjects were informed at the beginning of the study that there was a possibility that they would be asked to return for a second visit. If a subject did not agree to return to complete the reliability study, he or she was dropped from the reliability study and replaced by the next new subject. This selection process yielded a group of 33 subjects for the test-retest study and a group of 27 subjects for the inter-rater reliability study.

The group used for the test-retest reliability study consisted of 11 men and 22 women. The mean age was 65.7 (SD = 7.2) years and the mean number of years of education was 13.7 (SD = 3.2). The mean MMSE and 3MS scores were 29.4 (SD = 0.9) and 98.2 (SD = 1.7) respectively. The inter-rater reliability group was composed of 14 men and 13 women with a mean age of 67.2 (SD = 6.8) years and mean number of years of education of 12.9 (SD = 3.6).

The difference in sex distribution of the two groups was due to a random effect. The mean MMSE and 3MS scores were 29.6 (SD = 0.7) and 98.2 (SD = 1.8) respectively for the inter-rater reliability group.

2.1.2 Alzheimer Subjects

Subjects in the early stages of AD were recruited by referral from the McGill Centre for Studies in Aging at the Montreal General Hospital. The term "AD" refers to the early AD group. It had been estimated that 25 was a realistic sample size to achieve over the period of 18 months. These subjects were diagnosed by a neurologist as having probable primary progressive dementia of the Alzheimer type according to the Revised Diagnostic and Statistical Manual of Mental Disorders (DSM III-R) (American Psychiatric Association, 1987) and the criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's and Related Disorders Association (NINCDS-ADRDA) (McKahnn et al., 1984). The sensitivity and specificity of these criteria have been reported in the Introduction.

In order to help exclude multi-infarct dementia, the AD subjects had to have a score of less than five on the Ischemic Scale (Hachinski et al., 1975). The validity of the Ischemic Scale has been established using patients with AD, dementia of the vascular type, Pick's disease (Gustafason & Nilsson, 1982), depression (Wagner, Oesterreich & Hoyer, 1985), dementia associated with Parkinson's disease and normal pressure hydrocephalus (Molsa, Paljarri, Rinne, Rinne & Sako, 1985). The Ischemic Scale is used as an exclusion criterion in the diagnostic criteria of the NINCDS-ADRDA (McKahnn et al., 1984).

The 25 AD subjects were classified as being in stages 3 or 4 of the GDS (Reisberg et al., 1982). An inter-rater reliability coefficient of .82 has been reported for the GDS using 43 patients with AD (Gottlieb, Gur & Gur 1988). The GDS has been validated using memory tests and validity coefficients range from .50 to .63. Correlations with behavioural and clinical assessments vary widely, the highest being between .57 and .66 (Reisberg et al., 1982).

Subjects in GDS 3 had significantly higher MMSE and 3MS scores than those in GDS 4 (see Appendix F). Despite this, it is not unusual for AD subjects to display symptoms of both stages 3 and 4 from one day to another, depending on their level of fatigue or stress. Stress may be associated with going on a trip or visiting the hospital. In the literature, patients in stages 3 and 4 are usually combined and labelled the "early stages" of AD (Baum et al., 1988; Flicker et al., 1984; Flicker et al., 1988).

Other inclusion criteria for the AD subjects were that they be community-residing, with at least one primary caregiver who could attend the testing session, be English-speaking and right-handed as tested on the five-item version of the Edinburgh Handedness Inventory (Bryden, 1976; Oldfield, 1971; Raczkowski & Kalat, 1974). The subjects were required to perform adequately on the confrontation test for visual fields and to show visual acuity of a minimum of 20/50 using the pocket Snell card. Subjects were also required to be independent in walking and able to give consent to participate in the study (Appendix C).

All AD subjects were aged between 50 and 80 years. The upper limit of 80 years was chosen based on reports that severe dementia increases dramatically after that age in persons with AD (Katzman, 1986). Although there was an inclusion criterion for the performance of the normative group on the MMSE, there was no MMSE criterion for the AD subjects.

Each AD subject was required to attend three sessions for an initial, test-retest and inter-rater evaluation. Those who could not return for a second or third visit were replaced by a new AD subject, however, the data collected from the initial subject were used in the analyses. For example, if a person could not return to participate in the inter-rater reliability study, the test-retest data were kept and another subject was selected to provide data for the missing inter-rater assessment.

The selection procedure yielded a total of 25 AD subjects (9 men, 16 women). Twelve subjects were classified as being in stage 3 and 13 subjects were in stage 4 of the GDS scale. The characteristics of the AD and control groups are described in Table 1. Descriptive statistics were calculated for age, education, MMSE and 3MS. All distributions were unimodal. Age distributions for the control (median = 67, mode = 64) and AD (median = 69, mode = 73) groups approached normality although both were slightly negatively skewed (skewness = -0.3 for control, and -0.5 for AD). The distribution of the number of years of education was close to normality for the control group (median = 14, mode = 16, skewness = 0.2), and was positively skewed for the AD group (median = 11, mode = 12, skewness = 2.8). Negative skewness was also found in the distributions of the control group for scores on the MMSE (median = 30, mode = 30, skewness = -1.4) and 3MS (median = 98, mode = 100, skewness = -1.6), as well as in the distributions of the AD group for MMSE (median = 21, mode = 19, skewness = -0.7) and 3MS (median = 68, mode = 28, skewness = -0.8).

The groups were comparable with respect to the number of women and men. The average age and years of education of the AD group were not significantly different from those of the control group. As expected, the mean MMSE and 3MS scores for the early AD group were significantly lower than those of the control group.

Table 1

Characteristics of AD and control groups on the initial visit

	AD (GDS 3 & 4) n=25	CONTROL n=97	Chi-square(df)
Males (n)	9	36	
Females (n)	16	61	.01(1)
	M(SD)	M(SD)	t(df)
Age (years)	66.6(8.1)	67.7(7.1)	-0.7(120)
Education (years)	11.5(5.8)	13.7(3.7)	-2.3(120)
MMSE[a]	20.4(5.7)	29.2(1.1)	-17.7(24)*
3MS[b]	65.1(15.8)	97.30(2.8)	-10.1(24)*

*p<.0001, Student's two-tailed t-test.

[a]Mini-Mental State Examination.

[b]Modified Mini-Mental State Examination.

Performance on mental status examinations can be influenced by normal aging and by education level (Kittner et al., 1986). These possible influences should be considered in the interpretation of the results. The mental status scores were correlated with each other and with age and education using Pearson correlation coefficients. For each group, the MMSE and 3MS were highly correlated with each other as expected (AD: $r = .88$, $p < .0001$; control: $r = .58$, $p < .0001$). There were no other significant correlations for the AD group. However, for the control group, age was negatively correlated with MMSE ($r = -.24$, $p < .05$) and 3MS ($r = -.35$, $p < .001$). Mental status and age were not related to years of education in either group.

All except two of the AD subjects attended the three sessions. These two subjects, one woman and one man, attended the initial session and the test-retest session but did not attend the inter-rater session due to fatigue and inconvenience. When compared to the initial visit, the AD subjects performed slightly better on the mental status exams during the test-retest visit (MMSE: $t(23) = 2.2$, $p < .05$; 3MS: $t(23) = 2.2$, $p < .05$), but there was no significant change in mental status on the inter-rater visit (MMSE: $t(22) = 0.9$, $p > .05$; 3MS: $t(22) = 1.2$, $p > .05$).

As there is no gold standard for the measure of spatial orientation, the stage of AD was used as a criterion based on the assumption that spatial orientation declines with the progression of AD. The GDS is frequently the standard method used for staging the progression of the disease. Therefore, a late stage AD group was recruited so that stages 1 and 2 (control), 3 and 4 (early AD) and 5 (late AD) were represented. This GDS 5 group consisted of 10 subjects (5 males and 5 females), and met the same inclusion criteria as those in the other AD group. The mean age of this sample was 68.6 (SD = 7.3) years, the mean scores on the MMSE were 11.3 (SD = 6.3), and 35.9 (SD = 20.0) on the 3MS. These subjects were asked to attend only one session.

The medical charts of all GDS 3 and 4 and GDS 5 AD patients were reviewed in order to see whether there were any physical impairments or motor disorders that could affect their speed of performance on the tests. In addition, results of Computed Tomography (CT) scans were noted, where available.

2.2 Materials

2.2.1 Mental Status

Two mental status examinations were used concurrently, the MMSE (Folstein et al., 1975) and the 3MS (Teng & Chui, 1987). Researchers have made slight modifications to mental status exams in order to make them more appropriate to the setting (Watkins, Gouvier, Callon & Barkemeyer, 1989). In this study, the item "state" was replaced by "province" to make it more appropriate for Canadians. Three categories of words were used for recall and each category contained three words: something to wear: shoes, shirt, socks; a colour: blue, black, brown; and a good personal quality: honesty, charity and modesty. The first word of each category was used for the initial visit, the second word for the second visit and the third word for the inter-rater visit for the AD patients. The three-stage command used was "Take this paper with your left hand, fold it in half and hand it back to me". The pentagon figure was similar to that found in Folstein, Anthony, Parhad, Duffy and Gruenberg (1985). The cut off score of 26 out of 30 on the MMSE was selected based on the findings of Kay et al. (1985) who determined that the 25-26 cutting point has a sensitivity of 89.7% and specificity of 73.8% in identifying mild, moderate and severe cases of dementia according to the DSM-III.

Test-retest reliability as assessed at 24 hours, 4 weeks, 6 weeks and 16 months has revealed reliability coefficients ranging from .75 to .90, using 24 to 115 subjects with dementia (Anthony, Le Resche, Niaz, Vonkorff & Folstein, 1982; Fillenbaum, Hayman, Wilkinson & Hanes, 1987; Thal, Grundman & Golden, 1986; Uhlmann, Larson & Buchner, 1987). The validity of the MMSE has been established by correlating it with a psychiatrist's assessment of 97 patients (Anthony et al., 1982). The MMSE was 87% sensitive in detecting dementia or delirium at a cutting-point of 23-24, and 82% specific in determining the absence of both diagnoses (Anthony et al., 1982). Contrary to Bleecker et al. (1988) who found no correlation between the MMSE and education, Anthony et al. (1982) reported that the positive predictive value of the MMSE was 61% and that the 39% of subjects who were false positive all had fewer than nine years of education.

An alternate form of the MMSE, the 3MS, developed by Teng and Chui (1987) was used concurrently in this study to determine if the 3MS is more sensitive to differences between AD subjects. The 3MS appears to sample a wider range of cognitive functions than the MMSE, while maintaining brevity in its administration. Using 249 patients with eight different types of dementia, Teng and Chui (1987) examined the inter-rater reliability, but only for the pentagon drawing item, and reported a correlation coefficient of .98. The data from the current study were used to determine the test-retest and inter-rater reliability coefficients for the total scores of both mental status examinations.

2.2.2 Spatial Orientation Skills (SOS) Battery: Perceptual Spatial Skills Subtests

The SOS Battery consists of 13 subtests categorized into three types of spatial skills.

Four of these subtests examine basic perceptual skills, seven subtests assess more complex spatial skills, called cognitive skills. A third, even more complex level, involves using a questionnaire to evaluate functional spatial orientation. As described in the Introduction, these subtests were used in a preliminary study and were selected for the current study based on their ability to discriminate between the performance of normal control subjects and a group of early AD patients. The questionnaire replaced the way-finding task used in the preliminary study which was extremely time-consuming and thus, not practical. For most of these subtests psychometric data are not available, particularly when used with the normal elderly population or patients with AD. Thus, the first step was to establish their reliability and validity using subjects that represent these populations. Although many of these tests have been adapted from their original versions, these adaptations were minor and necessary in order for the AD patients to complete the battery of tests. In this section, these adaptations are clearly described for each subtest and, generally, they involve removing time limits or decreasing the number of items in order to shorten the subtest.

2.2.2.1 Left-Right Discrimination

The ability to discriminate left from right, or personal orientation, was tested using a ten-item test modified from a subtest of the Southern California Sensory Integration Test (Ayres, 1972). The test contained six commands that referred to the subject's body, such as "take this pencil with your left hand" and four that referred to the examiner's body, such as "put it in my left hand". The method of administration was modified from that used by Ayres so that the scores were not dependent on time limits, thus, a point was allotted for each correct response regardless of the response time. In the original version, one point was also allotted for each correct response within a one-minute time limit.

Inter-rater reliability of the original version has been established with a sample of adult head trauma patients and the correlation coefficient obtained was .93 (Siev, Freishtat & Zoltan, 1986). There are no data on the validity of this test. Although there was no significant difference between the mean scores of the AD and control groups in the preliminary study (Liu et al., 1991a), errors made by the AD group tended to be on items pertaining to the examiner's body. This test was kept because it contained information that was not covered by other spatial tests in the battery, namely, orientation to one's own body as well as to a confronting person. It was hypothesized that the test might be more sensitive if the number of items relating to extrapersonal orientation was increased to six and the number of items relating to personal orientation was decreased to four. Therefore, the items "Touch your right eye with any hand" and "Show me your left foot" were replaced by "Point to my right shoulder" and "Touch my left hand".

2.2.2.2 Figure-Ground Discrimination

This spatial skill was assessed using the Figure-Ground Perception Test of the Southern California Sensory Integration Test (Ayres, 1972). The Figure-Ground Perception Test contains 18 plates of black and white line drawings. Each plate consists of a complex figure at the top and six simple figures at the bottom, three of which are contained in the upper complex figure (see Figure 1). The subject is asked to indicate which three of the six simple figures are contained in the complex figure. The first set of nine plates contains figures of familiar objects

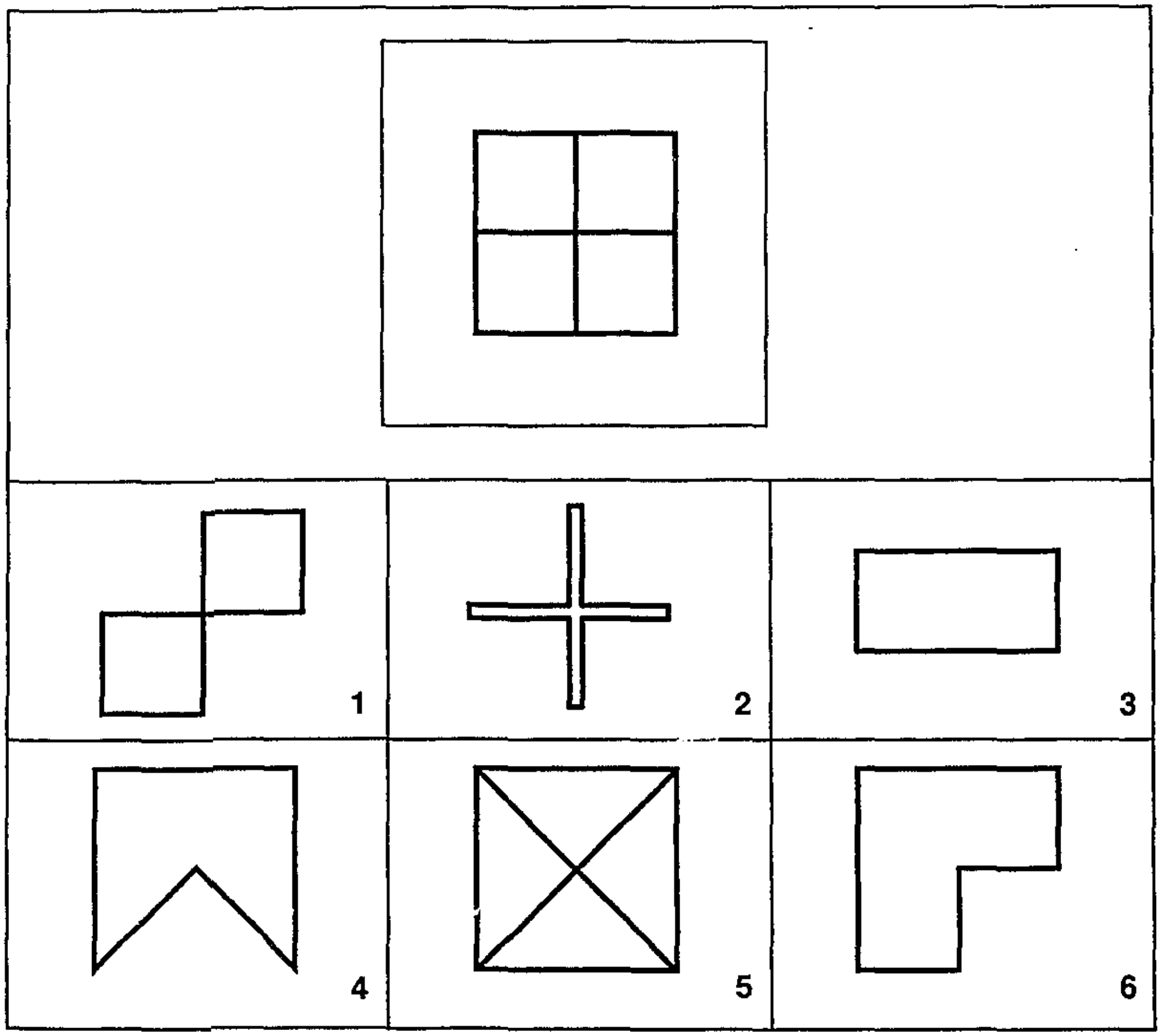


Figure 1. This is not an actual test item but a close replica of the trial item for geometric figures on the Figure-Ground Perception Test.

and the second set of nine plates contains geometric figures. The score on the Figure-Ground Perception Test is the number of simple figures correctly identified, and there is a maximum score of 24 points on each set or 48 for the complete test. In the original test, administration is terminated after five consecutive errors and there is a one-minute time limit per plate. In this study, the entire test was administered regardless of the number of consecutive errors and the time limit was not implemented. As in the original format, one point was allotted to each correct response. The adapted format for administering the complete test has been used in another study (Petersen, Goar & Van Deusen, 1985).

The Figure-Ground Perception Test was initially designed for use with children, however, normative data exist for adults (Petersen et al., 1985; Petersen & Wikoff, 1983). The test-retest reliability coefficients are .71 for the "first-five", that is, test scores based on terminating after five consecutive errors, and .90 for the entire test (Petersen et al., 1985). The Figure-Ground Perception Test shows adequate construct validity when correlated with the Embedded Figures Test (r = -.67), (Petersen & Wikoff, 1983). Although scores on Figure-Ground Perception Test are not significantly correlated with age or education for men (Petersen & Wikoff, 1985), significant correlations have been reported for women between Figure-Ground Perception Test scores and age (r = -.41) and education (r = .41) (Petersen et al., 1983).

2.2.2.3 Position in Space

The Position in Space Test is also a subtest of the Southern California Sensory Integration Test (Ayres, 1972). It evaluates concepts such as in-out, up-down and front-behind (Siev et al., 1986). The test consists of 30 items and is divided into three parts, the first two involving perceptual skills and the third involving memory. This study uses only the first two parts. Each of these two parts consists of eight pairs of plates depicting geometric configurations (Figure 2). The first (stimulus) plate depicts one or more geometric figures which are also represented in the second (target) plate in a variety of spatial orientations. The task of the subject is to select from the second plate the set of geometric figures that are of the same orientation as those depicted in the first plate. In part one, the stimulus plate is movable and so can be placed under each selection on the target plate to facilitate matching and comparison. In part two, the stimulus plate is permanently located on the left side of the figure. The number of figures in the stimulus plate begins with one and increases to three in part one, and to four in part two. The number of choices range from two to three in part one, and from three to four in part two. The total score for this test is 16. As specified by Ayres (1972), one point was allotted for each correct answer.

Mahoney and Siev (reported in Siev et al., 1986) found that a group of normal adults, aged 20 to 49 years, scored perfectly on the first two parts of the test. Using this test, Liu et al. (1991a) demonstrated that part II of this test discriminated between AD and control groups but part I did not. Inter-rater reliability was established by Baum (reported in Siev et al., 1986) using a group of adult patients with head trauma and the correlation coefficient obtained was .89.

2.2.2.4 Spatial Relations

The ability to perceive two or more objects in relation to oneself was assessed by using the Spatial Relations test, also a subtest of the Southern California Sensory Integration Test (Ayres, 1972) (see Figure 3). The test consists of two formboards, one with an egg-shaped hollow and one with a diamond-shaped hollow, two pegs and eight blocks for placement in the

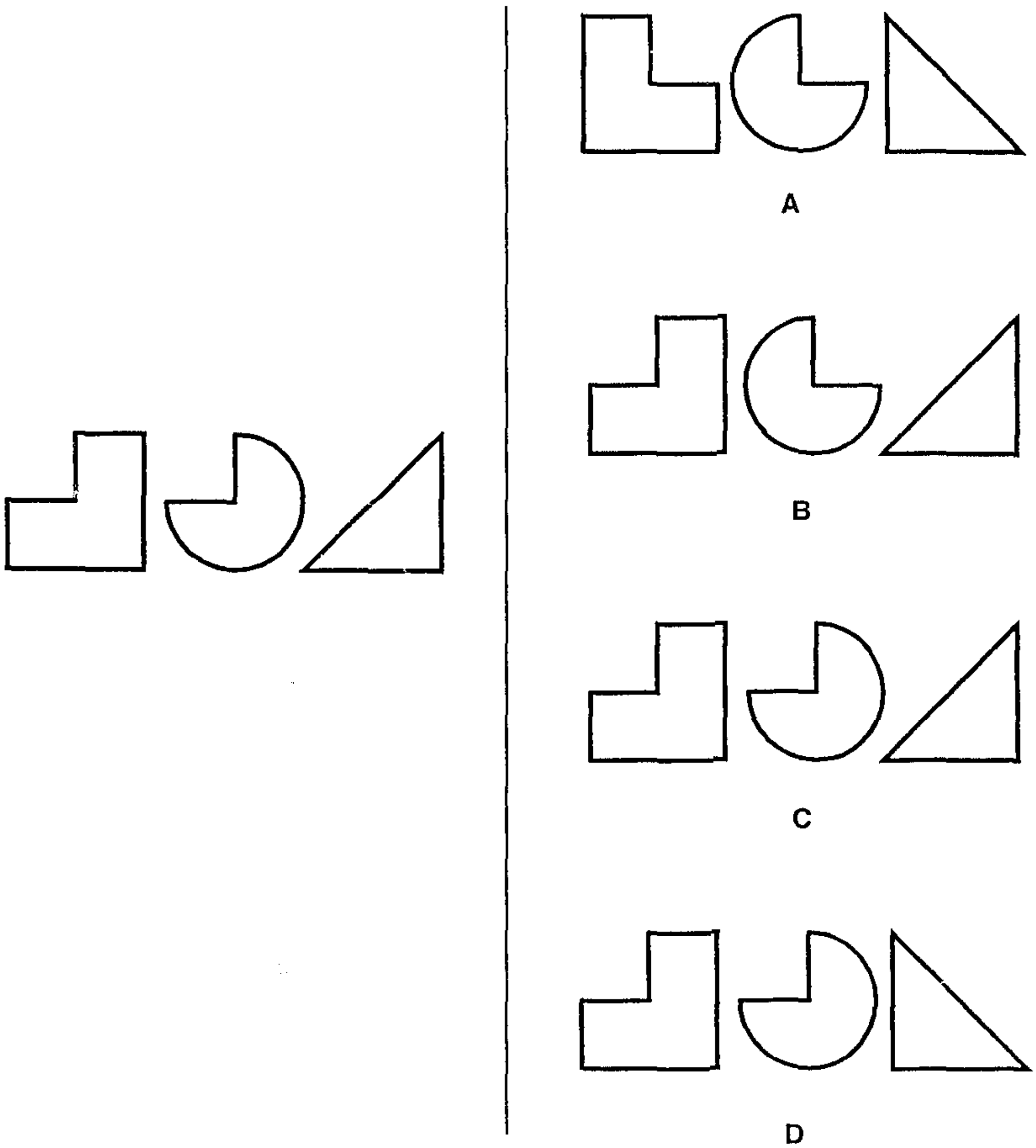


Figure 2. This is a close replica of a test item in part two of the Position in Space test. The subject is asked to select the set of figures on the right that match the stimulus set on the left.

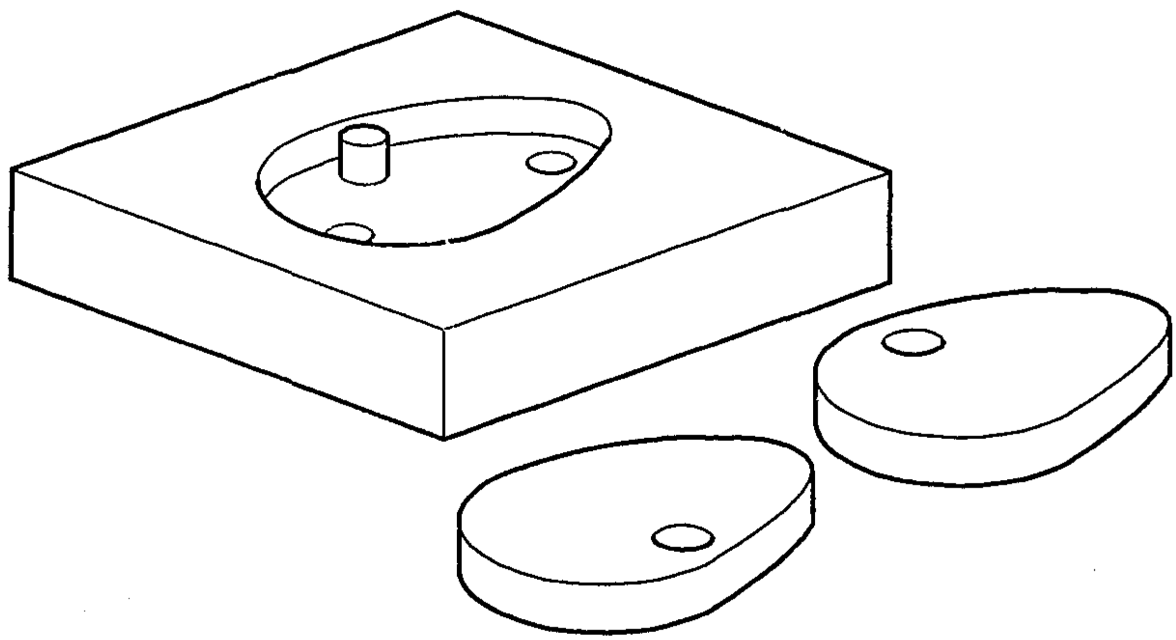


Figure 3. This is a close replica of a test item using the egg-shape forms on the Spatial Relations Test.

formboards (four egg-shaped blocks and four diamond-shaped blocks). A peg is inserted into one of the formboards by the examiner and the subject is asked to select the block that fills the hole by choosing one of the two blocks given to him or her. There are a total of 60 test items, 25 for the egg-shaped formboard and 35 for the diamond-shaped formboard. In order to shorten the test, 10 items were selected based on a previous study (Liu et al., 1991a). These items were 5, 7, 11, 13, 15, 21, 23, 25 and 27. Item 1 was used as a trial. The total score for this test is 10. One point was allotted for each correct response as specified by Ayres (1972).

Although standardized for children, normative data have also been collected for 90 subjects aged 50 to 64 years and 60 subjects aged 65 to 74 (Taylor, 1968). Correlations between odd and even items given one week apart revealed a reliability-stability coefficient of 0.95 (Ayres, 1962). Test-retest correlations ranged from 0.28 to 0.77 for children aged 4 to 11 years (Ayres, 1972). Discriminant validity of this test has been evaluated using 50 children with brain pathology. These children were significantly impaired on the test compared to a general population of children (Ayres, 1968).

2.2.3 Spatial Orientation Skills (SOS) Battery: Cognitive Spatial Skills Subtests

Cognitive spatial skills are described as those skills requiring problem-solving, decision-making, extrapersonal orientation, spatial memory and mental representation. Many of the following tests are well described in the literature on spatial skills. These are the Porteus Maze, Map-Reading, Road-Map, Corsi Block, Stylus Maze and sketch map tests.

2.2.3.1 Porteus Maze Test

The Porteus Maze (Porteus, 1959) Test was selected to assess spatial planning and decision-making (Figure 4). It consists of 11 mazes with a ceiling score of 17 test years. In the present study subjects were given the complete set of mazes instead of starting with the maze corresponding to test age 11 as specified by Porteus for adults. This decision was based on the results of preliminary study which indicated that early AD subjects performed at a lower level than test age 11 (Liu et al., 1991a). As specified by Porteus (1959), subjects were given two trials for mazes up to 11 years and four trials were allowed for mazes for years 12 and above. The scoring method used differed from those detailed by Porteus. Points were not taken off for qualitative errors such as lifting the pencil, touching a line, cutting corners or wavy lines. Subjects were reminded of the instructions during the test if they made these qualitative errors. The score on this test was the highest test year the subject completed within the allotted number of trials.

The validity of the Porteus Maze test has been studied by comparing the pre- and post-operative performances of patients who underwent frontal lobotomies (Porteus, 1959). Patients with frontal lobotomies were significantly impaired on the Porteus Maze Test (Porteus, 1959). Its reliability has not been reported.

2.2.3.2 Map-Reading Test

The Map-Reading Test assesses extrapersonal orientation, spatial problem-solving and mental rotation (Semmes et al., 1955). The original version of the test consists of 15 diagrams of paths: five visual maps and ten tactile maps. Only the five visual maps were used in this

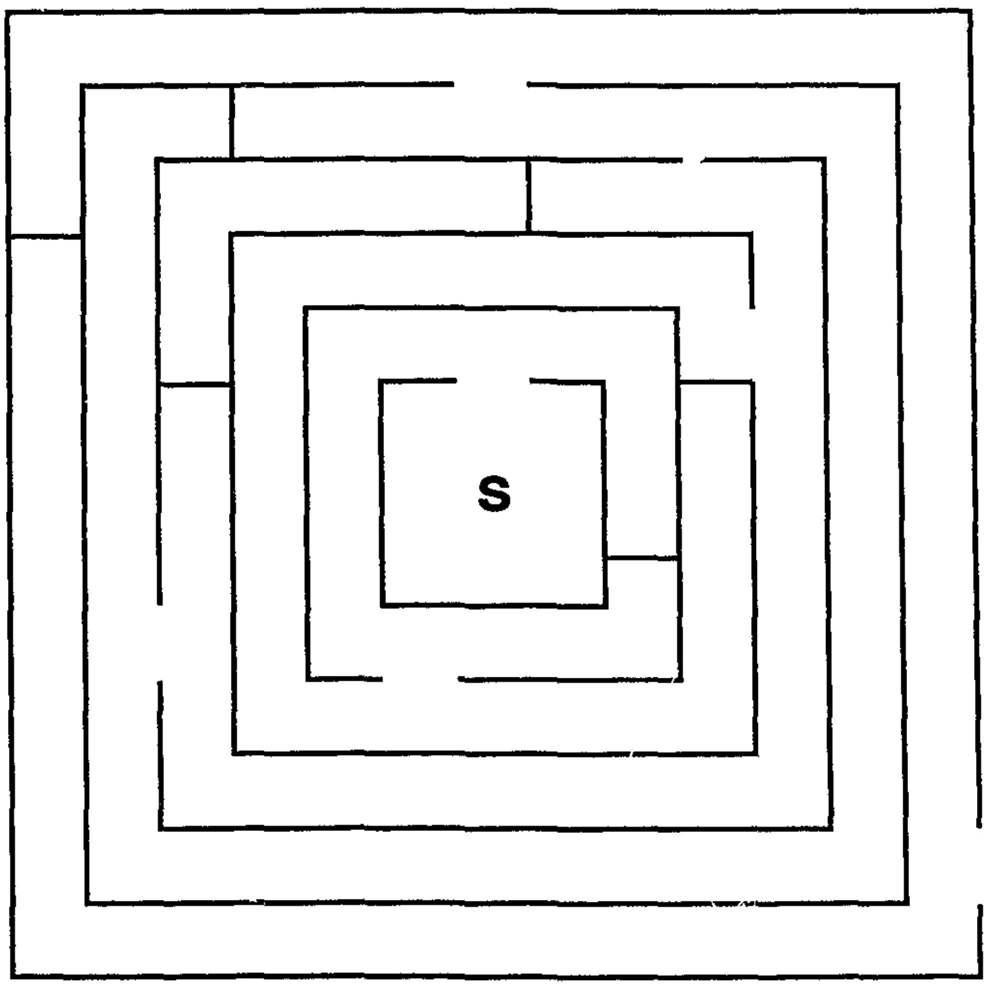


Figure 4. In this replica of a test item on the Porteus Maze Test, the subject is asked to take a pencil and, starting in the centre indicated by the "S", draw the route to the exit.

study and they were preceded by a trial map (see Figure 5). Each map was drawn on a 32 cm square card which contained nine red circles, one of which was indicated as the beginning circle by a black ring. The direction North was also indicated on each map. The nine circles were represented on the floor of the examining room by nine red circular spots, 15.5 cm in diameter sewn onto a 155 cm square piece of white vinyl fabric. Adjacent spots were 60 cm apart from centre to centre. The reduction in the size of the test was the only modification made to the original map. This adaptation makes the test more practical for clinics or laboratories and minimizes fatigue that can result from taking many steps. In general, subjects only needed to take one step to reach each of the circles. In the preliminary study AD subjects were able to perform this test.

The five maps were graded in difficulty; the first map contained five turns and each successive map involved an additional turn. One wall was indicated as North by a sign. The subject was instructed to hold each card with the direction North opposite to him-/herself at all times. A revised, shorter version of the original instructions was used: "I will give you some maps to follow (the trial map was shown to the subject). The North indicated on the map represents the North wall (the wall facing the subject). I would like you to follow the route on the map by walking on the appropriate circles on the floor. You begin on the spot that is marked by a black ring. You must keep the map in front of you so that you are always facing the direction you are walking and that you are never walking sideways or backwards". Any part of the instructions was repeated as frequently as needed and the participants were not timed in order to minimize stress. Subjects received one point for each correct turn and the maximum score for the entire test was 35.

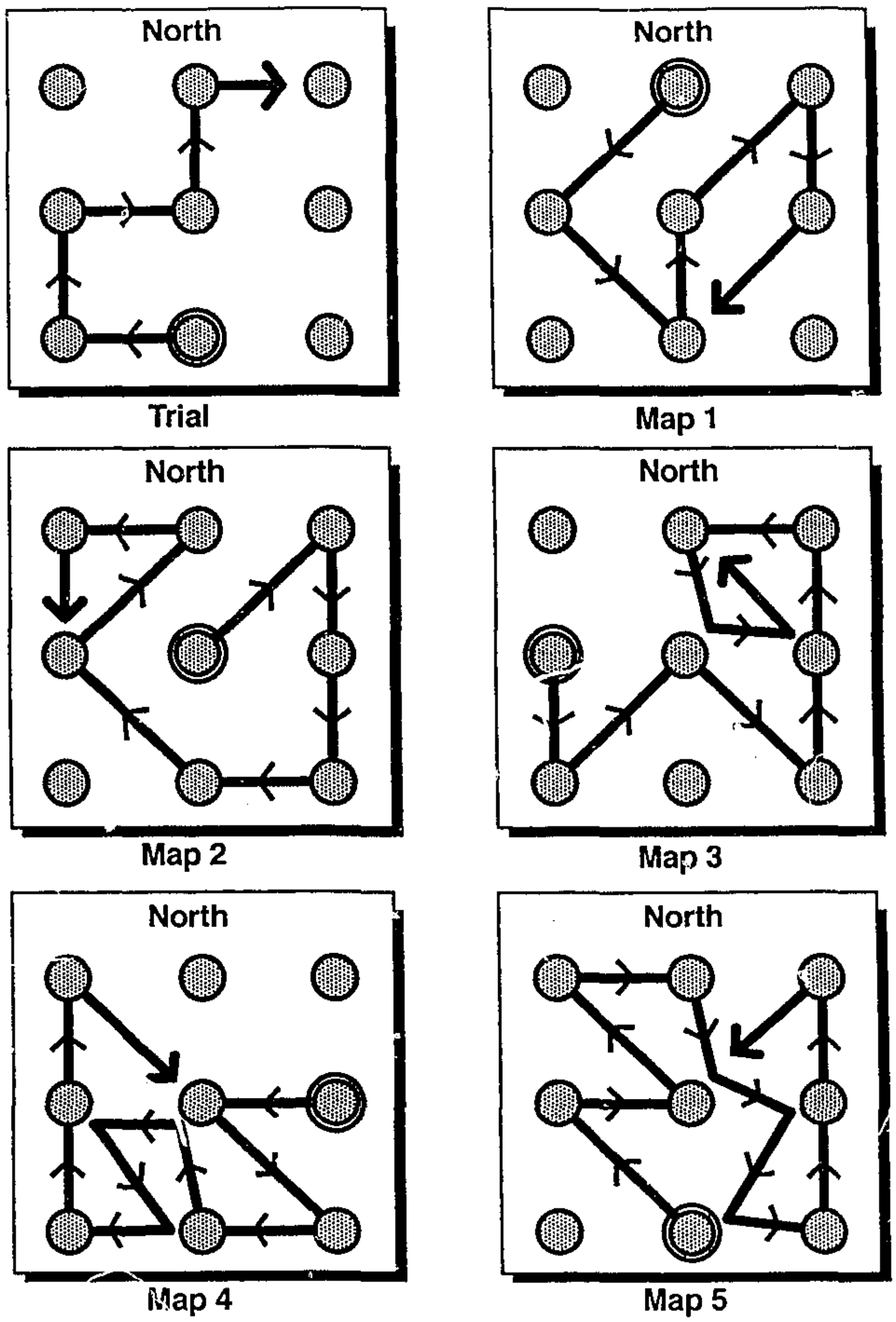


Figure 5. The Map-Reading Test assesses extra-personal orientation.

The scoring method used in the original version involves a total score based on all of the maps (Semmes et al., 1955). A score of 0 is given for any map on which the subject makes an error. If a map is successfully completed, the subject receives a score equivalent to the number of turns on the map. Thus, a higher weighting was placed on the successful completion of the longer maps. Other authors have implemented different scoring methods. For example, Ratcliff and Newcombe (1973) allocated a point for each disc visited in its correct ordinal position. Aubrey and Dobbs (1989) analyzed the mean number of correctly completed segments to the first error as well as the total times taken to perform the maps. They point out that the number of segments of a route does not necessarily correspond to the number of turns, that is, the distance between two discs corresponds to a segment. For this study, a point was given for each correct turn as was done in the preliminary study in which the primary concern was with the subject's ability to determine the direction rather than keeping track of the number of segments.

Although no data are available regarding its validity or reliability, the Map-Reading Test is well-known and has been used extensively in studies of patients with focal brain lesions (De Renzi, 1982; Rudel, Teuber & Twitchell, 1974). Aubrey and Dobbs (1989) have shown that the Map-Reading Test appears to be sensitive to the effects of aging by demonstrating poorer performance in a healthy elderly group in comparison to a healthy younger group of undergraduate students. The influence of education was not examined in that study.

2.2.3.3 Road-Map Test

This test, designed by Alexander at al. (1964), involves presenting the subject with a route drawn on a map and instructing the subject to pretend that he or she is walking along the route

(Figure 6). The Road-map test has been used as a test of personal orientation. It is considered to contain more cognitive than perceptual components because it requires mental rotation and more concentration than the left-right discrimination test.

The map is kept upright at all times and the subject must verbalize at each turn whether it is right or left. One point is allotted for each correct turn for a maximum score of 32. It was observed in the preliminary study that the original map was difficult to follow because it lacked contrast (Liu et al., 1991). Therefore it was slightly adapted by colouring the route blue, by indicating the start with a red dot and by including hatch lines in spaces not part of the route (see Figure 6). The validity and reliability of this test were not assessed by Alexander et al. (1964), although it is widely used by neuropsychologists and has been used with various clinical populations such as patients with AD, multi-infarct dementia and localized cerebral lesions (Brouwers et al., 1984; Butters, Barton & Brody, 1970; Bylsma et al., 1992).

2.2.3.4 <u>Corsi Block Tapping Test</u>

The Block Tapping Test was designed by Corsi (Milner, 1971, 1980) to study immediate spatial memory and non-verbal memory span. It consists of nine black wooden cubes, measuring 3 cm square, affixed in an irregular pattern on a black wooden board, measuring 20.5 cm by 25.5 cm (see Figure 7). The blocks are numbered on the side facing the examiner so that the subject's performance can be numerically coded. In this study, the test served only as an evaluation of immediate spatial memory span and not as an index of learning. Subjects were not asked to perform the 24 trials as specified by Corsi (Milner, 1971). The instructions were, "I am going to tap some of these blocks and I would like you to tap the blocks in exactly the same sequence

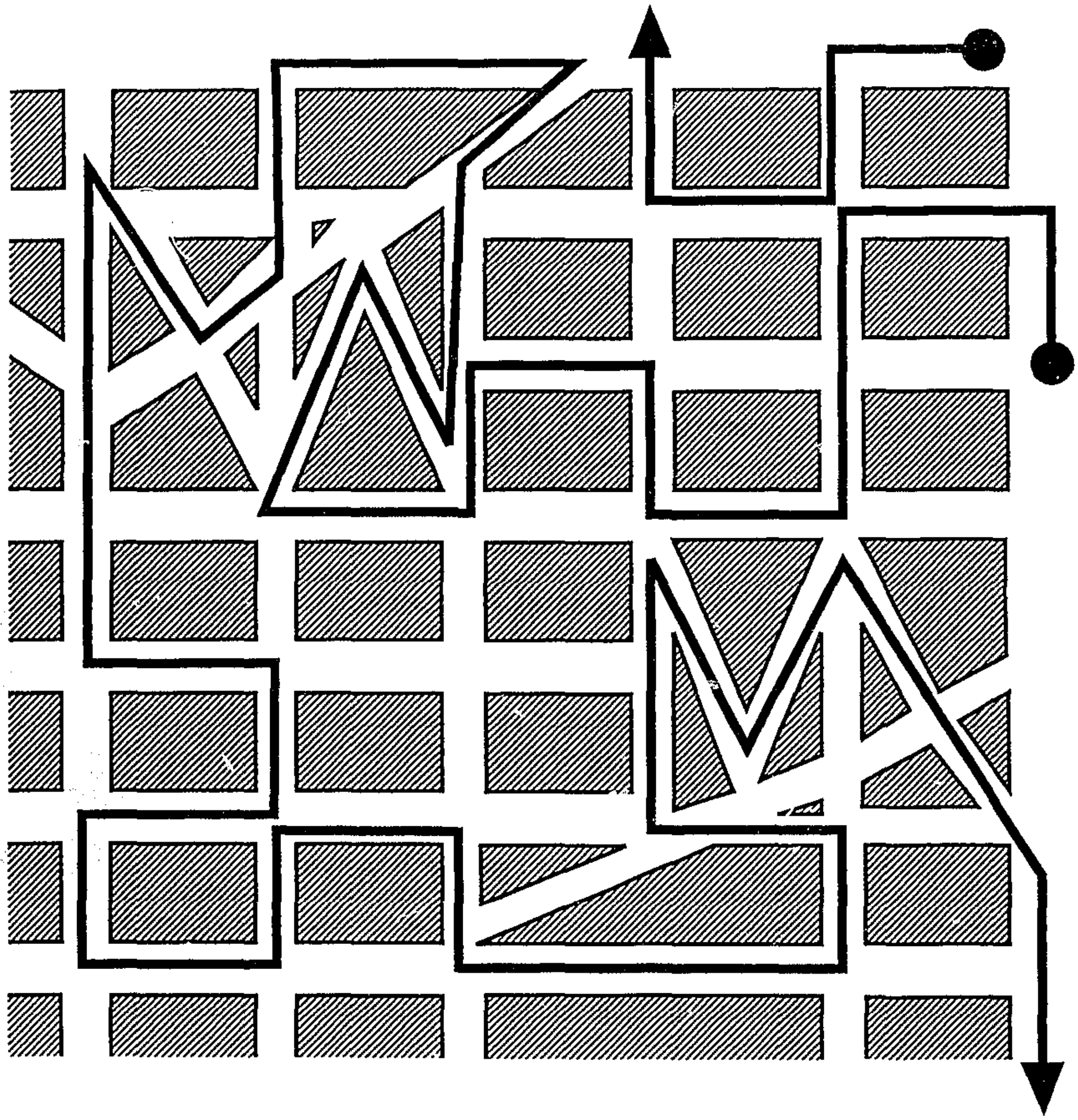


Figure 6. The instructions for the Road-Map Test are demonstrated using the smaller route. Subjects are asked to verbalize at each turn whether he/she would be making a left or a right turn.

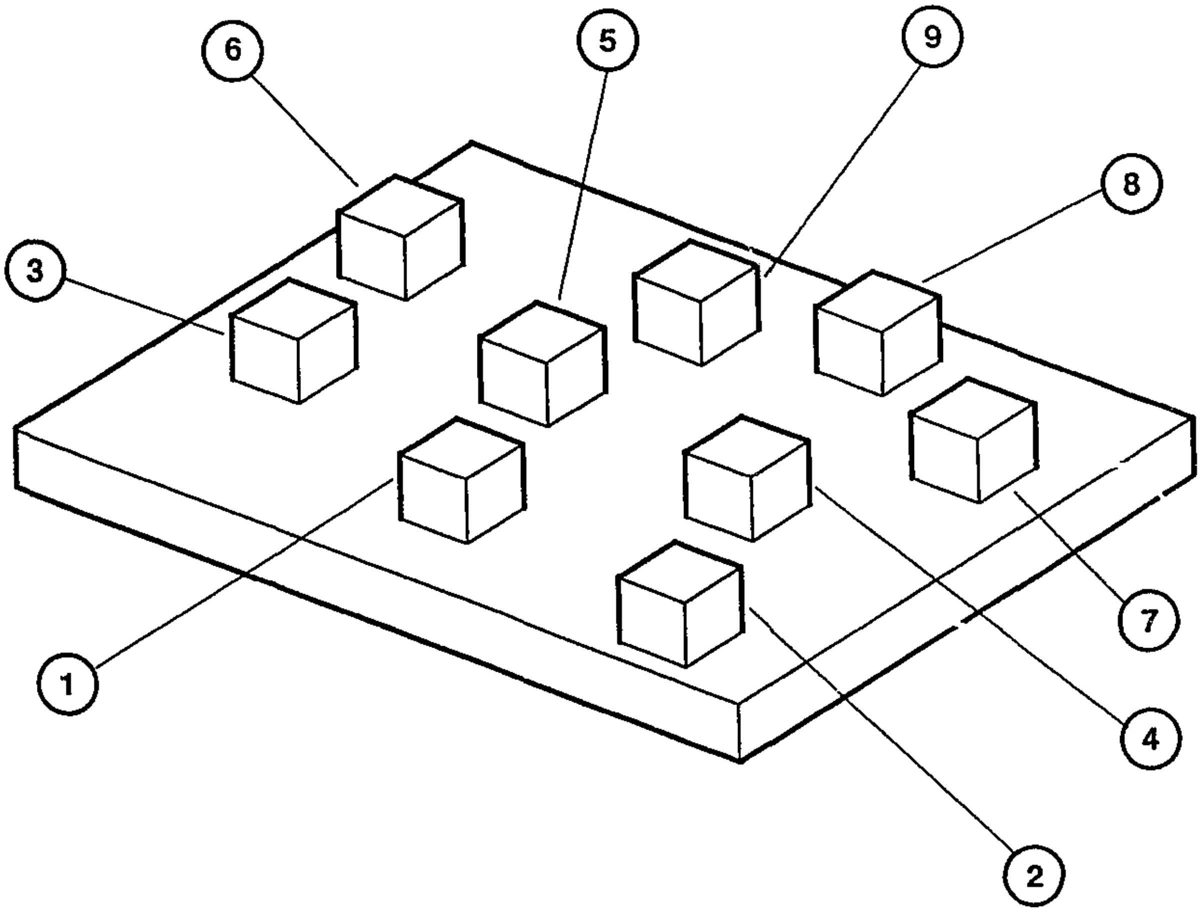


Figure 7. This figure of the Corsi Block Tapping Test shows the subject's view. Numbers appear on the examiner's side.

as I do". If the subject was unable to repeat the sequence correctly, the examiner did not present the same sequence. Instead, another sequence of the same span was presented. If the subject was unable to repeat this sequence, the test was terminated.

With correct performance, the span progressively increased to a maximum of nine, each span containing two sequences or trials as on the original version. The second sequence was not used if the subject was able to repeat the first sequence correctly and the score was the span repeated, regardless of whether it was the first or the second trial. The reliability and validity of this test have not been established although it is commonly accepted as a nonverbal version of the digit span test (Milner, 1971, 1980).

2.2.3.5 Modified Stylus Maze Test

Spatial learning was assessed using a modified version of the visually-guided Stylus Maze Test used by Milner (1965) and Brouwers et al. (1984). The apparatus consists of a black wooden board with 100 brass screws forming a 10 by 10 array as on the original version. The only modification made was that a quadrant of the original test was used. This quadrant measured 33 cm by 33 cm and had a five-by-five array (see Figure 8) of brass screws that were spaced 2.5 cm apart centre to centre. In the preliminary study, some AD subjects experienced extreme difficulty performing the test even with the five-by-five array, thus, administering the original version would not have been possible.

The task of the subject was to discover and learn, by trial and error, a hidden path from the lower left corner to the upper right corner of the array. The subject was instructed to use a metal stylus and to proceed horizontally or vertically but not diagonally. Screws lying outside

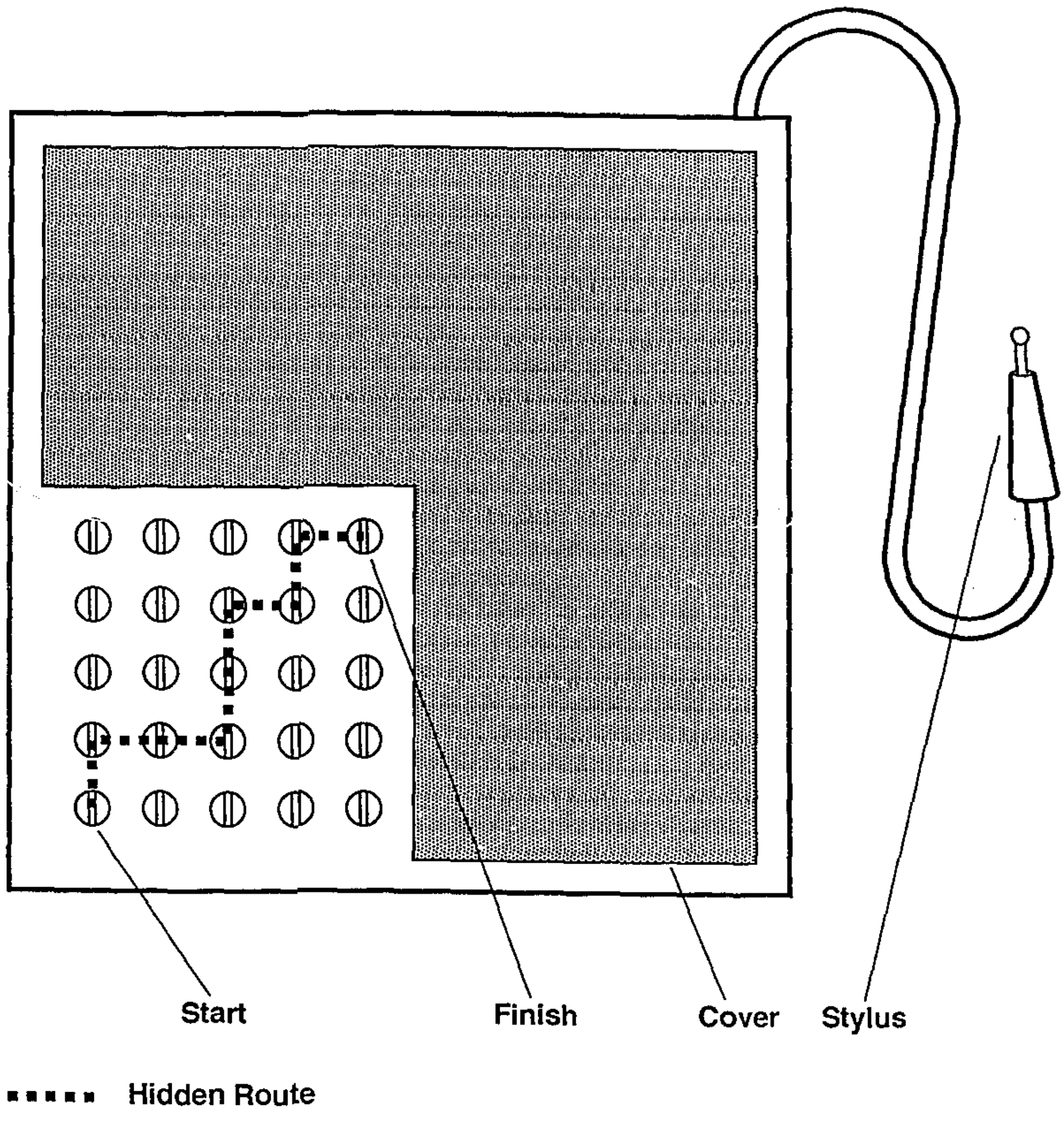


Figure 8. The modified Stylus Maze Test used a quadrant of the original test.

of the path were connected to a battery and buzzer which provided auditory feedback each time the subject touched a screw not on the correct path. When an incorrect screw was touched, the subject was instructed to return to the previous screw and to try another direction. As chance influenced the number of errors made on the first trials, the number of errors made on only the fifth trial was used to calculate the score.

The Stylus Maze Test has been used extensively in neuropsychological evaluations (De Renzi, 1982), however, it has not been evaluated for reliability. Pothig, Pogelt and Roth (1985) recommend the Stylus Maze Test for assessment of "orientational capacity", problem-solving and memory.

Before analyzing the data from the Stylus Maze Test, the scores were transformed. Unlike the other subtests, a higher raw score indicated poorer performance. If higher Stylus Maze scores also reflected better performance, it would facilitate comparisons with the other tests and also permit using these data in factor analysis and logistic regression. The formula would have to take into consideration that there was no limit to the number of errors and therefore there was no denominator for the scores. In the preliminary study (Liu et al., 1991a) it was observed that as the number of errors increased, there was a tendency in the AD group to perseverate. Following consultation with a statistician, it was decided that the following square-root function would be an appropriate method of treating the data:

$$[1/(\#\text{ errors} + 1)]^{1/2} \times 10$$

This function put a heavier penalty on the first few errors and less penalty as the number of errors increased. It also provided a maximum score of 10 which was comparable to the other subtests of the SOS Battery.

2.2.3.6 Mental Representation of the Familiar Environment

Mental representation was assessed by having the subject draw a sketch map of the main level of his or her home after a sample floor-plan (see Figure 9) was shown to the subject. The subject was also asked to label each room. The control subjects were then given a sheet of paper and a stamped envelope addressed to the University, and asked to draw another floor plan when they returned home. This floor plan was subsequently returned to the University by mail and used to score the initial floor-plan sketched in the laboratory.

Each AD subject and accompanying person (caregiver, spouse or friend) was also asked to draw a floor plan of the main level of the AD subject's home. In addition, the caregiver was asked to complete and return another floor plan by mail as was done for the control subjects. This second floor plan was used to confirm the accuracy of a caregiver's first floor plan which was used as a standard for scoring the AD patient's sketch map. The control subjects and caregivers of the AD subjects were instructed to verify that the maps they drew at home were accurate and complete.

Scoring was based on the average of two ratios. The first ratio was the number of rooms (frequency) correctly identified over the actual number of rooms. The second ratio was the number of rooms correctly placed in relation to the others (accuracy) over the actual number of rooms. The actual number of rooms referred to those only on the main level of the homes.

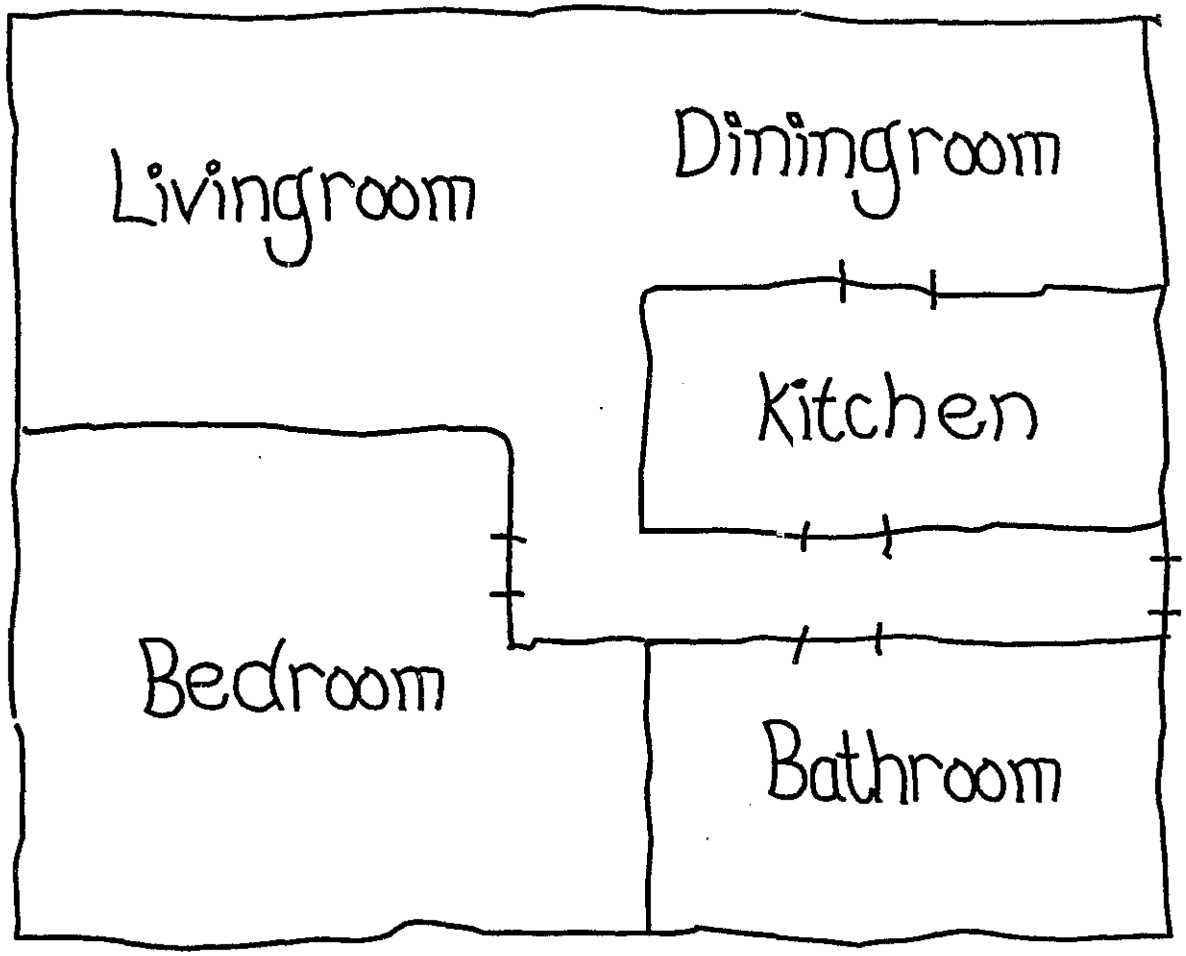


Figure 9. This sketch map was shown to subjects as an example of a hand-drawn floor plan of a home.

2.2.3.7 Sketch Map of the New Environment

A memory task was designed to evaluate visuospatial construction skills. All control and AD subjects were shown a floor plan of a "new environment" which contained five rooms labelled as: living room, kitchen, dining room, bedroom and bathroom (see Figure 10). Subjects were asked to study the floor plan for one minute and were warned that they would be asked to reproduce the floor-plan. When one minute had elapsed, the floor plan was removed and the subject was given a white sheet of paper and asked to draw as much as he or she could recall. The score was the average of the frequency and accuracy scores as described for the sketch map task of a familiar environment. Given that a standard floor-plan was used for this test, all of the scores were based on a denominator value of six. Subjects were not required to identify and label the "entrance" and "hallway".

2.2.4 Spatial Orientation Skills (SOS) Battery: Functional Spatial Abilities

A 12-item Functional Spatial Abilities Questionnaire was developed to measure functional spatial skills or way-finding ability (see Appendix D). Items for the questionnaire were based on a survey conducted prior to the study. The Alzheimer Society of Montreal sent out a bilingual questionnaire (Appendix E) with its regular mailings to approximately 200 caregivers of AD patients. The questionnaire asked caregivers to describe behaviours and actions that they had seen in persons with AD who had difficulty with orientation to place. A total of 30 caregivers responded to the questionnaire. The content of these responses was categorized into spatial behaviours relating to indoor and outdoor aspects of familiar and new environments.

Caregivers reported that in familiar environments many AD patients mistook doors and

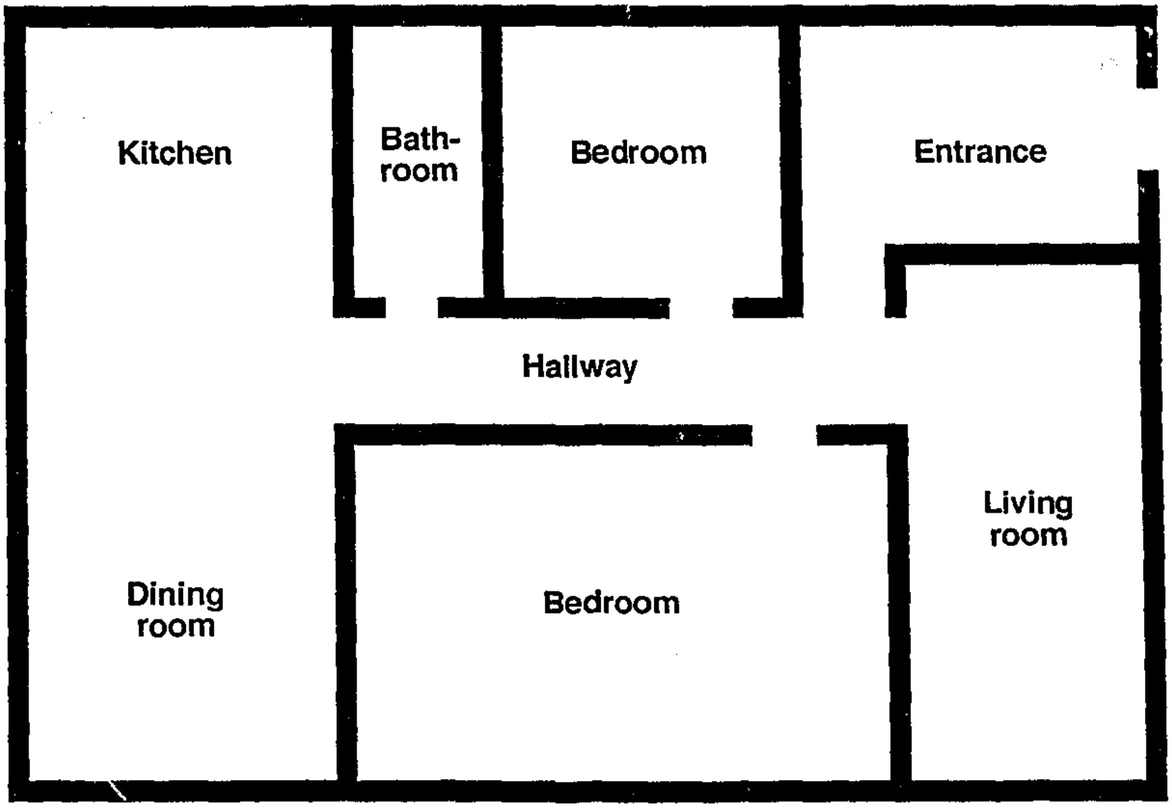


Figure 10. Subjects studied this map for one minute after which they drew the map from memory.

entered the wrong room. Some AD patients were also reported to be confused as to what floor of their homes they were on by looking for bathrooms or laundry rooms on the wrong floor. Other behaviours observed were pacing in corridors, fidgeting with doorknobs, and appearing indecisive about entering or exiting a room. When outside in familiar environments, caregivers reported observing excessive walking in circles, following strangers in the neighbourhood, difficulty in navigating in a familiar supermarket, taking an unusually long period of time to reach a destination, getting lost in the metro and appearing distracted. These patients also tended to walk either very slowly, take smaller than normal steps, or walk extremely fast. Behaviours reported in unfamiliar environments were similar. Caregivers reported that many AD patients were unable to continue driving or to travel to a new environment alone because they became lost. The AD patients were noted to show difficulty in navigating to and from a parking lot, washrooms or various rooms in new environments such as in a doctor's office or nursing home. Emotions associated with these behaviours were anxiety, distress, tearfulness, hesitancy, denial and annoyance.

The content of the Functional Spatial Abilities Questionnaire was based on these responses. The first six questions related to functional abilities in new environments, and the last six related to functional abilities in familiar environments. The Functional Spatial Abilities Questionnaire was given to control and AD subjects to complete at the end of the testing session. The control subjects were asked to identify a healthy relative or friend who could complete the questionnaire on the control subjects' behalf. Similarly, the caregivers were asked to complete the questionnaire with respect to the performance of the AD patients. This allowed the comparison of self-rated and proxy-rated responses.

Items on the Functional Spatial Abilities Questionnaire were rated as 1 for "yes", 2 for "not applicable" and 3 for "no". The second category was used if the subject was not given the opportunity to perform the activity. For example, some caregivers never permitted the person with AD to travel alone, not necessarily because this person had been lost before, but for fear that this person would get lost. However, if an AD subject was no longer performing an activity because he or she had shown an inability to do so safely in the past, the caregiver was instructed to rate the item "no". The total score for this questionnaire is 36. An item analysis was conducted in order to determine whether each of the items of the Functional Spatial Abilities Questionnaire was able to discriminate between individuals who have AD and those who do not. This analysis gave some indication of the validity of the questionnaire. If many of the items did not discriminate between AD and control groups, then the questionnaire would not be a useful tool in its present format.

2.2.5 Other Tests of Cognitive Function

Other tests were also administered with the SOS battery in order to measure attention, language, memory and other visuospatial constructive skills. These were cancellation tests, a confrontation naming test, a verbal memory span test and the clock drawing test. Because the battery was time consuming to administer, an attempt was made to select tests that would not add much more time to a session. These measures were administered following the completion of the SOS battery.

2.2.5.1 Attention

Two cancellation tests were used, the line-crossing test, also known as Albert's Test, and the Bells Test. The line crossing test (Albert, 1973) does not have any distractors and requires the subject to cross out all 40 lines randomly placed on a 21.5cm x 28cm sheet of paper. The Bells Test (Gauthier & Joanette, 1992; Gauthier, Dehaut & Joanette, 1989) is more difficult in that it uses silhouette figures of bells (targets) surrounded by figures of 14 types of common objects (distractors). The subjects were asked to circle all the bells they could find on the sheet. The maximum scores are 40 for Albert's Test and 35 for the Bells Test.

Normative data exist for Albert's Test (Vanier et al., 1990) and the Bells Test (Gauthier et al., 1992; Vanier et al., 1990). In a study using both tests with control subjects and patients with right hemisphere cerebral vascular accidents, Vanier et al. (1990) found that the Bells Test was more sensitive in detecting unilateral neglect. Neither test has been used with an AD population.

2.2.5.2 Language

Language ability was assessed using the Boston Naming Test which is a 60-item test of visual confrontation naming (Kaplan et al., 1976). The items are line drawings of common objects. In order to reduce the duration of the test, and to decrease the demands on attention and concentration, only odd items were used, which resulted in a maximum score of 30. As described in the Introduction, 30-item versions have been validated (Mack et al., 1992). No reliability data have been reported by the original authors.

2.2.5.3 Verbal memory

The Hebb verbal memory test (Hebb, 1961; Milner, 1971) involving repetition of digits was used to test short-term memory span. This is a verbal version of the Corsi Block Test. As this test was not used to measure learning, only the maximum verbal span was assessed, as was done with the Corsi Test. The maximum span tested was nine.

2.2.5.4 Clock drawing

Given that there have been relatively few studies of visuospatial constructive ability in AD patients, clock drawing was also included in the study. The subject was given a sheet of paper with a circle drawn on it. The instructions were, "Here is the outside circle of a clock. I would like you to fill in the numbers of the clock starting from one". After the subject had done so, he/she was instructed as follows: "Please draw in the two hands of the clock to indicate ten minutes past eleven." The clocks were scored according to the method developed by Doyon et al. (1991) which provides a maximum score of 20. This method has been shown to have high sensitivity and specificity when used to distinguish a group of AD patients from a group of normal control subjects. Although this test has been reported to have high reliability by Doyon et al. (1991), it not certain whether this refers to test-retest or inter-rater reliability. As mentioned, a related objective of the study was to obtain these reliability coefficients because of the current widespread use of clock drawing tasks in the clinic and in research.

2.3 Procedure

All assessments were administered in a quiet laboratory at the School of Physical and

Occupational Therapy at McGill University and each subject was tested alone. At the beginning of each session, subjects were informed that their performance was not timed and that accuracy was more important than speed. A 10-minute rest period was provided after 40 minutes to avoid fatigue. Scores were recorded on a summary sheet (Appendix G).

In the preliminary study (Liu, 1991a), the SOS subtests were block randomized, that is, the perceptual subtests were administered before the cognitive subtests which were followed by the functional subtests. Within each category the subtests were administered in a random order. For the functional spatial skills category, navigation in the new environment was tested before navigation in the familiar environment which was tested during a home visit. Random presentation of the tests in the perceptual and cognitive spatial skills categories was not successful for the AD patients. Due to anxiety, many patients were reluctant to perform certain subtests in the sequence that they were presented. Since the priority in the present study was to obtain as much data as possible by encouraging AD patients to perform all of the subtests, it was decided that the SOS subtests would be administered in the same sequence for all subjects. When a subject expressed reluctance to perform a subtest, the next subtest was presented. At the end of the testing session, subtests that were not administered were presented once again to the subject. This method of administration was also used for the control subjects although fewer control subjects expressed anxiety or reluctance to perform a subtest.

2.3.1 Content validity

In order to evaluate the adequacy of the subtests for assessing spatial orientation, a panel of experts was asked to comment on the choice and completeness of skills assessed by the SOS battery. The panel consisted of ten experts: two researchers, two neurologists, three OTs and three psychologists. They received a summary of the SOS battery, an operational definition of spatial orientation and a description of the spatial skills tested. They were then asked to answer a content validation form (Appendix H) and to return it by mail. A 51% positive response ("yes", "agree" or "strongly agree") was set, a priori, as acceptable for content validity for questions one to five. This criterion has been accepted and used for content validation of other indices (Spitzer et al., 1981).

2.3.2 Test-retest and inter-rater reliability

The 33 control subjects, described under the section "Subjects", and all 25 early AD subjects were used to estimate test-retest reliability. The time interval between all sessions was two weeks. The effects of practice and learning or carry-over effects would have increased with shorter time intervals. Longer time intervals would have been less practical and may have increased the chances of detecting actual cognitive deterioration.

For the control group, inter-rater reliability was determined by retesting a group of 27 healthy individuals who were different from those used for the test-retest study. It was anticipated that it would be very difficult to recruit enough AD subjects for both reliability studies over the time period the study was being conducted. Therefore, in addition to the initial and the test-retest reliability visits, the AD subjects were asked to return a third time for the inter-rater reliability study. The interval between the two sessions was two weeks. In addition to the investigator, two occupational therapists (OTs) and a psychologist were recruited to be raters. The three raters were selected to represent the types of clinicians and researchers who are likely to use the SOS Battery. A copy of the SOS Battery manual was given to each rater. This was followed by a training session where the raters were shown the method of administering and scoring the subtests. The investigator performed most of the initial assessments, that is, 60 of the 97 control subjects, 20 of the 25 early AD subjects and 9 of the 10 late AD subjects. The investigator also performed most of the test-retest assessments: 19 of the 33 control subjects and 18 of the 25 early AD subjects. She acted as the second rater in inter-rater reliability study for five control subjects and two early AD subjects. All other test sessions were equally distributed among the three other raters.

2.3.3 Construct validity

Two validation procedures were used to obtain evidence for construct validity. First, the original SOS battery was evaluated for its ability to discriminate between persons with AD and those who were healthy by comparing the mean scores of the two groups on all 13 subtests. In order to justify using the SOS Battery with AD patients, it should, at the very least, be able to differentiate individuals who have AD from those who do not. An Odds Ratio value that was above one would be considered as indicative of a positive association between higher scores and the likelihood of being a control subject.

Construct validity was also assessed using the Reintegration to Normal Living Index (Wood-Dauphinee et al., 1988) and the shortened SOS Battery. The rationale for using this index was that performance on the SOS Battery could reflect an individual's quality of life. A person who is spatially disoriented should experience a poorer ability to integrate into normal living. The Reintegration to Normal Living Index was designed to assess global function or, specifically, the degree of reintegration to normal living in individuals after incapacitating illness. It has been defined as being related to the physiological, psychological and social components of quality of life but, in addition, includes the functional status of the patient (Wood-Dauphinee et al., 1988). In theory, spatial disorientation should be associated with a decline in global function particularly as it relates to the psychological and social components of quality of life and the functional status of an individual. The construct of global functioning, as measured by the Reintegration to Normal Living Index, includes skills that relate to spatial orientation. This is seen in the first three items which are 1) I move around my living quarters as I feel is necessary, 2) I move around my community as I feel is necessary, and 3) I am able to take trips out of town as I feel are necessary. It is an 11-item questionnaire, and each item is a statement accompanied by a visual analogue scale. These scales are anchored by phrases reflecting whether or not the statements describe the situation of the subject. Subjects are asked to mark an X within the scale to indicate how this statement applies to them at that point in time. Each visual analogue scale is a 10 cm premeasured line which is converted to a score from 1 to 10 points, depending on where the individual marks the scale. The 11 items sum to a total of 110 points but are proportionately converted to 100 points for ease of interpretation. It was hypothesized that if the shortened SOS battery measures an aspect of spatial orientation that relates to global functioning, then one would expect that total scores of the shortened SOS Battery would be positively correlated with the total score of the Reintegration to Normal Living Index.

The Reintegration to Normal Living Index is presently being used with AD subjects and preliminary results indicate that the caregivers of AD persons are able to identify the reintegration to normal living pattern (A. Opzoomer, personal communication). In this study, this index was answered by all AD and control subjects as well as a person they identified as a proxy. In most cases, the proxy was the spouse or offspring of the control subject, and the spouse or caregiver of the AD subject.

The accuracy of self-ratings by AD patients has been discussed in the Introduction. This was examined by comparing self-ratings with proxy responses by caregivers. The questionnaire given to the caregivers was adapted by changing the subject pronouns so as to apply to another person. The Reintegration to Normal Living shows high internal consistencies (Cronbach's alpha above .90 for each item), adequate proxy-patient response correlations (r = .65), responsiveness to change and good construct and criterion validities (Wood-Dauphinee et al., 1988).

2.3.4 Criterion validity

By definition, criterion validity refers to the correlation of a measure with another well accepted measure, preferably a "gold standard". Although there is currently no gold standard for measuring spatial orientation, the GDS could be used as a criterion based on its wide acceptance as a criterion for grading the progression of AD. Each stage is accompanied by general descriptions of behaviours and abilities that characterize that particular stage. These behaviours include those that pertain to spatial disorientation. Thus, the GDS was adopted as a criterion for evaluating the validity of the SOS Battery. A group of 10 GDS stage 5 (late AD) subjects was recruited to represent a more advanced state of the disease. Performance on the SOS Battery was compared using subgroups of 10 GDS stages 3 and 4 (early AD) subjects and 10 GDS 1 and 2 (control) subjects who were matched with respect to sex, age and years of education to the GDS stage 5 (late AD) group. Rather than correlating the scores on the SOS battery with the criterion, the performances of the three groups were compared using the "known groups" approach (Bohrnstedt, 1983). Criterion validity was also evaluated for the shortened battery by comparing the total scores for each of the three groups.

2.4 Data Analysis

All of the statistical analyses in this study were conducted using the SAS/STAT (SAS Institute Inc.) statistical software. For all Student's t-tests, the SAS program used the formula for Satterthwaite's (1946) approximation for the degrees of freedom (Steel & Torrie, 1980). All subtest scores were treated as being at the interval level of measurement. Descriptive statistics calculated for subtests and total scores included means, medians, modes, skewness and standard deviations.

Although most studies group AD subjects in GDS 3 and 4, Adelstein et al. (1992) did separate subjects in the two stages and found a significant difference in spatial performance as measured on tests of spatial recognition and memory for spatial order. In this study, prior to grouping subjects in GDS 3 and 4 in the analyses, the data from the groups were analyzed separately and then compared to verify that there were no major differences.

Comparative analyses of the means of the early AD and control groups were done using the two-tailed Student's t-tests for independent samples. Item analysis of the Functional Spatial Skills Questionnaire was performed using the Wilcoxon signed-rank test. This nonparametric test was chosen because the responses to each item were categorical in nature. In other statistical analyses, the total scores on the Functional Spatial Abilities Questionnaire were assumed to be interval in order to permit the use of factor analyses. The total time taken to complete the battery was also compared between the two groups.

Pearson's correlation coefficients were obtained between the SOS Battery subtest scores and age, years of education and mental status scores. These analyses were used to determine to what extent age, education and mental status were related to performance on each subtest. If performance on a subtest was affected by normal aging or years of education, then these influences need to be considered in the interpretation of the data. These variables also need to be considered in selecting subtests for a shortened version of the battery.

The intra-class correlation coefficient (ICC) was calculated as an index of test-retest and inter-rater reliability. The ICC is a ratio of the variance of interest over the sum of the variance of interest plus error (Kay, et al., 1979). The ICC analyses were based on the two-way random effects model proposed by Bartko (1966):

$$R = (MSP - MSE) / [MSP + (k - 1) \times MSE + (k/n) \times (MSR - MSE)],$$

where MSP denotes estimates of the variance (mean squares) due to subjects, MSR denotes estimates of the variance due to rater effects and MSE denotes the estimate of the error variance. The number of subjects is denoted by n and the number of raters is denoted by k. This model does not require any assumptions about the presence or absence of the rater-subject interaction. In addition, while it is conceptually similar to the model proposed by Ebel (1951), it takes into account the rater effects when estimating error variance. In contrast, Ebel's formulae implicitly ignore these effects and accordingly tend to over-estimate the actual value of the ICC. An ICC value was deemed acceptable if it was statistically significant at $p < .05$ and if its point estimate was equal to or greater than .70. This cut-off would ensure that the proportion of variance corresponding to systematic differences between subjects was not lower than 50% of the total variance in observed scores.

Internal reliability or internal consistency of the battery was examined using four types of analyses. These analyses were subtest-to-subtest correlation coefficients, subtest-to-total score correlation coefficients, Cronbach's coefficient alpha and factor analysis. For subtest-to-subtest correlational analysis, correlation matrices of the subtest scores were generated using Pearson's product-moment correlation coefficients. The correlations were used to provide indications as to how subtests within and between categories related to each other. Subtest-to-total score Pearson's correlation coefficients were calculated by correlating each subtest with the total score omitting that subtest. If the subtest score was not removed from the total score, the correlation would be artificially inflated because the correlation coefficient would include correlation of the subtest score with itself. This procedure was done for standardized scores as well as raw scores. The number of items in each subtest varied considerably, and it was of interest to determine whether this variability in test items would affect the correlations. For the total raw score, subtests were weighted differentially, based on their maximum scores which ranged from 1 to 48. The total standardized score took into account the observed variance and standardized each subtest to have a standard deviation of 1 unit. Therefore, subtests with lower variance would be more heavily weighted relative to their weighting when the raw scores were used. Kline (1986) has proposed that the correlation between an item (in this case, a subtest) and the total score should be above .20 and that subtests with lower correlations should be discarded.

Internal consistency of the SOS battery was also estimated using Cronbach's alpha coefficients (Cronbach, 1951). Theoretically, Cronbach's alpha is an average of all possible split-half reliabilities of the battery. A split-half reliability refers to randomly dividing the battery into two sub-batteries, which are then correlated with each other (Streiner & Norman, 1989). Again, Cronbach's alphas were calculated using raw and standardized scores. The total raw score was simply the sum of the individual subtest scores. An overall alpha coefficient that is greater than .80 is regarded as indicative of high internal consistency (Carmines & Zeller, 1979). After obtaining an overall alpha, Cronbach's alphas were then calculated for the battery after removing one subtest at a time. If an alpha coefficient showed a large increase after the removal of a subtest, this would indicate that its exclusion would increase the homogeneity of the battery (Streiner & Norman, 1989).

A fourth measure of internal consistency was derived from the results of a factor analysis, which was performed to identify groups of subtests that test similar abilities. Although the battery demonstrates high internal consistency as shown by correlation coefficients and Cronbach's alphas, the battery of subtests may be described in terms of several underlying factors that relate to spatial orientation. Factor analysis was conducted with the SAS FACTOR+ procedure (SAS Institute Inc., 1990) and used the principal factors approach with varimax rotation. This approach ensured that factors were orthogonal, thus uncorrelated, while maximizing the variance of loadings (correlations between variables and factors) within factors and across variables. This helps to contrast variables that do load on a given factor with those that do not load, enhancing the interpretability of results. When defining the factor structure, a variable was said to "load" on a factor if its correlation was at least .40. Factor analysis was first done using data from only the control group. The large number of subtests did not permit factor analysis to be conducted using the data from the AD group which was smaller in size. The number of factors was determined using the standard criterion according to which a factor is retained if and only if its eigenvalue exceeded 1.0 (Tabachnick & Fidell, 1989). This implied that all factors that explained more variance than a single variable were accepted.

The results from the first factor analysis were used together with other criteria to develop a preliminary shortened version of the battery. The criteria for selecting subtests were that the subtest demonstrated good external reliability when administered to AD patients and that the scores were not significantly correlated with age or education in either group. Since the shortened version was designed for use with AD patients, the selection of subtests was based primarily on data obtained from the AD group. An attempt was made to select subtests that represented the factors that emerged from the first factor analysis. The creation of a shortened version of the SOS battery served two main purposes. First, the time required to administer the battery in a clinical situation would be considerably reduced. Second, factor analysis could be conducted using data from the smaller AD group. The second reason was important because, as with the other analyses, it could not be assumed that the data from the two groups would behave similarly when subjected to the same analysis.

For the shortened battery, two factor analyses were conducted, one involving the data from the AD group and another involving data from the control group. When performing the second factor analysis for the AD group a fixed number of factors was selected to equal the number of factors of the control group in order to enhance the comparability. Ideally, the recommended maximum number of variables would be five for a sample of 25 subjects (Tabachnick & Fidell, 1989). The internal consistency of the shortened battery was evaluated by calculating subtest-to-total score Pearson's correlation coefficients and Cronbach's alphas for both the AD and control groups. Once again, these calculations were done for raw and standardized scores. The alpha coefficients for the shortened battery were expected to be smaller than those of the original battery because the value of alpha is not only dependent on the average correlation among subtests, but also on the number of subtests (Carmines & Zeller, 1979).

Multiple logistic regression analysis was used to establish construct validity and was based on the hypothesis that spatial performance, as assessed on the SOS battery, is related to the presence or absence of AD. First, if the SOS battery demonstrates construct validity, there should be a marked difference between the performance of normal subjects and AD patients. Second, this difference between the groups should be detected even when adjusted for possible confounding subject characteristics. Multiple logistic regression analysis was used for this purpose. In this analysis, a binary variable indicating whether a subject was normal or had AD was used as the dependent variable and the independent variables were the subtest score, gender, age and education. To adjust for the effects of multiple analyses, a stringent criterion ($p \leq .001$) was used to determine the significance of the effect of each subtest score. Confidence intervals (95%) for the adjusted odds ratio were determined in order to assess the discriminative ability of the battery.

Construct validity was also evaluated by correlating the shortened SOS battery total score with the Reintegration to Normal Living Index scores using the Pearson product-moment correlation coefficient. In order to facilitate comparison between the two measures, the raw scores from each of the subtests in the shortened battery were converted to a score out of 20 and added to obtain a total score on 120. This total score was then converted to a score out of 100 because the total score for the Reintegration to Normal Living Index was also out of 100.

A one-way analysis of variance (ANOVA) using a fixed effects model was performed to determine the criterion validity of the battery using the GDS as the criterion. Post-hoc analyses were conducted using the Neuman-Keuls test. Similarly, ANOVA and post-hoc analyses were conducted for the shortened battery using the total scores for each of the three groups.

3.0 RESULTS

This section begins by presenting evidence for the content validity of the SOS Battery. The distributions of scores and the normative data are compared to the scores of the AD group. Next, subtest-to-subtest correlation coefficients are presented followed by subtest-to-total score correlation coefficients for the full battery. Both sets of data are described for the AD group and then for the control group. Correlations coefficients between the 13 subtests and age, education and mental scores are then listed separately for the AD and control groups as these data were used as part of the criteria for developing the shortened battery. Next, Cronbach's alpha coefficients for the full battery are listed for the AD group and then for the control group. The first factor analysis was conducted using the full battery and data from only the control group. Based on all of the above information, a preliminary shortened version of the SOS Battery is then proposed. Again, the internal consistency of the shortened battery is evaluated using correlational analyses, Cronbach's alphas and two factor analyses, one for the AD group and another for the control group. Construct validity was evaluated using multiple logistic regression with the full battery and by correlating the shortened battery with the Reintegration to Normal Living Index. Criterion validity, specifically concurrent validity of the full battery was assessed using the GDS as a criterion. Finally, the results of other tests of cognitive function are presented.

3.1 Content Validity

A panel of 10 experts was consulted to evaluate the adequacy of the SOS Battery subtests for assessing spatial orientation. The experts were sent content validation forms (Appendix H) to complete and return by mail. Of the experts consulted, six replied and all had positive responses to the first five items on the questionnaire. Many of these individuals could be classified into combined professions. Two were psychologists and researchers completing their doctorates, two were OTs also completing doctorates, one was an OT with a doctoral degree and one was an OT clinician. The two neurologists and one psychologist researcher did not respond. Since this met the 51% criteria, the SOS Battery was accepted in its current format for use in the study.

3.2 Descriptive and Comparative Analysis of the SOS Battery

The performances of patients in GDS 3 and 4 are presented in Appendix I. There were significant differences in the scores of the two groups on the Position in Space and the Road-Map tests. The groups performed similarly on all other subtests of the SOS Battery, and there was no difference between the groups on the total SOS Battery score based on the shortened version, or on the self-rated and proxy-rated Reintegration to Normal Living Index scores. As the groups differed on only two of the 13 subtests, they were combined to form the "early AD group" in all subsequent analyses.

All distributions of scores on the perceptual, cognitive and functional spatial tests were unimodal for the early AD and control groups. These values can be found in Table 2. The data and results from the statistical analyses of the SOS Battery subtest scores are presented for the early AD and control groups in Tables 3, 4 and 5. These tests have been categorized into perceptual, cognitive and functional spatial subtests. The mean time to complete the battery was 103.6 minutes for the AD group (range: 80 to 150 minutes) which was significantly longer than the completion time of 87.2 minutes (range: 60 to 135 minutes) for the control group, ($t(117)$ =

Table 2

Distributions of scores on subtests of the SOS Battery

	AD (GDS 3 & 4)			CONTROL		
Subtest	Median	Mode	Skewness	Median	Mode	Skewness
Left-Right[a]	9	9	-1.1	10	10	-2.6
FGP[b]	25	25	2.0	35	28	0.0
Position in Space	11	11	-0.4	15	16	-1.9
Spatial Relations	7	7	0.2	10	10	-1.7
Porteus Maze	7	5	1.4	15	16	-1.9
Map-Read	2	0	2.5	27	35	-0.7
Road-Map	18.5	16	-0.4	30	32	-1.2
Corsi Block	3	3	-0.8	5	5	0.3
Stylus Maze	5	7.1	0.9	7.1	10	-0.2
Sketch-map: Fam[c]	0.6	0	-0.3	1	1	-2.3
Sketch-map: New[d]	0.3	0	0.9	1	1	-2.2
FSAQ: Self-rated[e]	30	32	-0.3	36	36	-2.2
FSAQ: Proxy-rated[f]	26	22	-0.2	36	36	-3.7

[a]Left-Right Discrimination

[b]Figure-Ground Perception

[c]Sketch-map of the familiar environment

[d]Sketch-map of the new environment

[e]Functional Spatial Abilities Questionnaire: Self-rated

[f]Functional Spatial Abilities Questionnaire: Proxy-rated

Table 3

Comparison of performance on perceptual spatial subtests

	AD (GDS 3 & 4)			CONTROL			
Subtest [maximum score]	n	range	M (SD)	n	range	M (SD)	t (df)
Left-Right[a] [10]	25	4-10	8.3 (1.6)	97	8-10	9.9 (0.4)	**-4.7*** (24.7)
FGP[b] [48]	22	5-39	24.8 (6.2)	97	23-48	35.6 (6.4)	**-7.7*** (117)
Position in Space [16]	25	4-16	11.0 (3.5)	97	10-16	15.1 (1.1)	**-5.7*** (25.3)
Spatial Relations [10]	25	5-10	7.7 (1.4)	97	8-10	9.7 (0.6)	**-7.4*** (26.9)

Note. The degrees of freedom (df) vary due to the unequal variances of the AD and control groups. The estimates of df were based on Satterthwaite's (1946) approximation.

[a]Left-Right Discrimination

[b]Figure-Ground Perception

*$p<.0001$, two-tailed Student t-test for independent samples.

Table 4

Comparison of performance on cognitive spatial subtests

	AD (GDS 3 & 4)			CONTROL			
Subtest [maximum score]	n	range	M (SD)	n	range	M (SD)	t (df)
Porteus Maze [17]	24	5-17	8.3 (3.9)	97	10-17	16.1 (1.9)	**-9.7*** (25.7)
Map-Read [35]	24	0-35	5.1 (8.2)	96	4-35	26.0 (8.5)	**-10.9*** (118)
Road-Map [32]	24	0-32	19.5 (6.9)	95	15-32	28.3 (4.4)	**-6.0*** (28)
Corsi Block [9]	25	0-5	2.9 (1.7)	97	3-7	4.6 (0.9)	**-5.0*** (27.1)
Stylus Maze [10]	22	3-10	5.3 (1.7)	96	4-10	8.0 (2.0)	**-5.9*** (116)
Sketch map: Fam[a] [1]	23	0-1	.51 (.37)	82	0-1	.93 (.12)	**-5.4*** (23)
Sketch map: New[b] [1]	24	0-1	.28 (.30)	95	0-1	.90 (.18)	**-9.7*** (27.3)

[a]Sketch-map of familiar environment

[b]Sketch-map of new environment

*$p<.0001$, two-tailed Student's t-test for independent samples.

Table 5

Comparison of performance on the Functional Spatial Abilities Questionnaire (FSAQ)

	AD (GDS 3 & 4)			CONTROL			
Respondent [maximum score]	n	range	M (SD)	n	range	M (SD)	t (df)
Self-rated [36]	21	22-36	29.7 (4.4)	97	28-36	35.0 (1.9)	**-5.6*** (24.1)
Proxy-rated [36]	23	14-34	25.6 (5.6)	85	25-36	35.3 (1.9)	**-8.2*** (23.4)

*p<.0001, two-tailed Student's t-test for independent samples.

4.5, p < .0001).

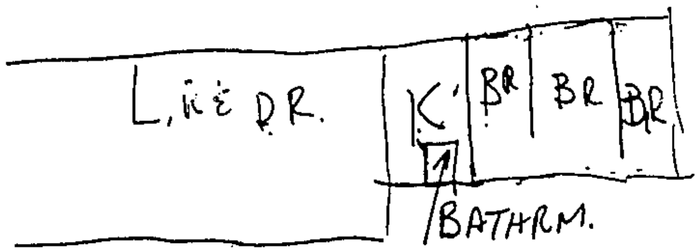
3.2.1 Comparative analysis of subtests of SOS Battery

On all perceptual tests the means of the AD group were significantly lower than those of the control group (see Table 3). The control group obtained almost perfect scores on all of the perceptual subtests, except for Figure-Ground Perception Test. Only three AD subjects were unable to finish the second part of the Figure-Ground Perception Test, which involved disembedding geometric figures. They were able to complete the first part of the test, however, which involved disembedding line drawings of familiar objects.

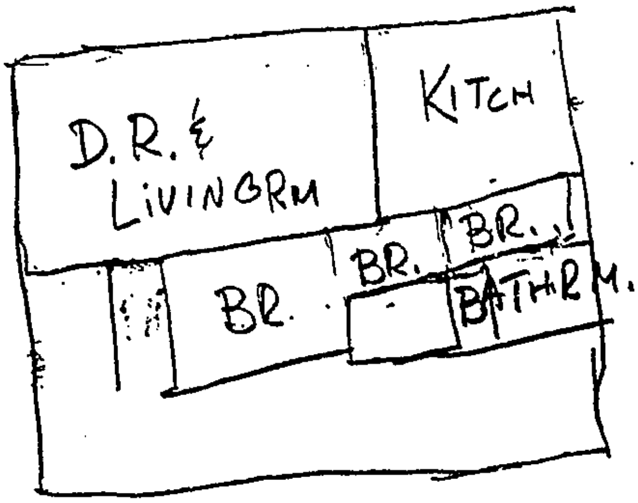
The AD group also performed significantly poorer on all of the cognitive spatial subtests. Table 4 shows the mean scores of the two groups. Tests for which sample sizes were smaller than 25 for the AD group and 97 for the control group indicate that some of the subjects were unable or unwilling to complete the subtest. For the control group the sample size for the sketch map of the familiar environment was considerably smaller than for the other tests. This was due to the lower compliance of the proxies of control subjects in returning the drawing from which the subject's drawing was scored. Examples of sketch maps are shown in Figures 11 and 12.

3.2.2 Functional spatial abilities

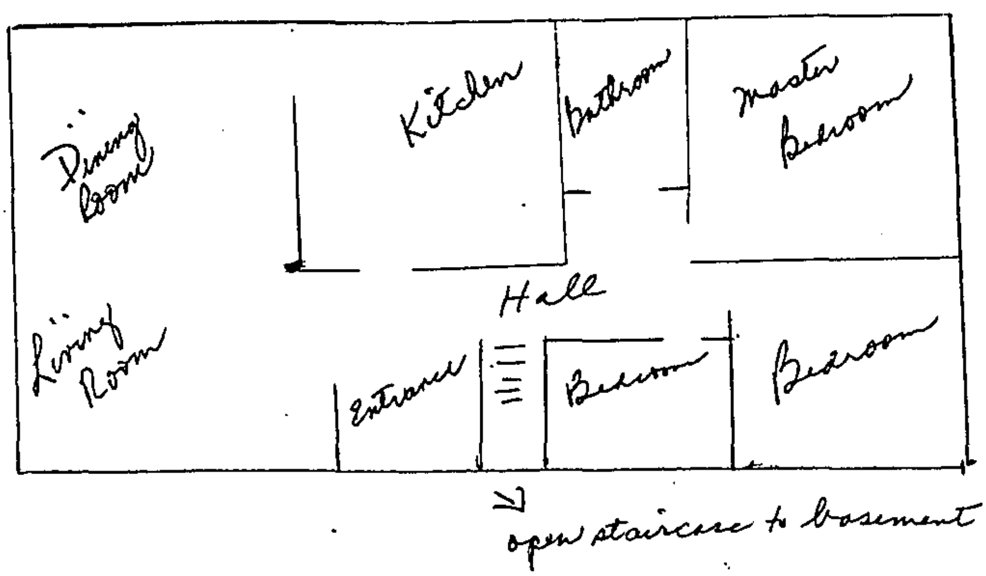
The third type of spatial skills examined was functional spatial skills as measured by the Functional Spatial Abilities Questionnaire. The AD subjects' self-ratings were significantly lower than those of the control group on the Functional Spatial Abilities Questionnaire (see Table 5). Similarly, proxy-ratings of the AD group were significantly lower than proxy-ratings of the



(a)



(b)



(c)

Figure 11. Sketch maps of familiar environment drawn by an AD subject (a and b) and caregiver (c). Sketch map (a) was drawn on the initial visit and sketch map (b) on the second or test-retest visit. These maps were compared to that drawn by the subject's spouse (c). The AD subject was male, 63 years old and in GDS stage 3. This patient had 14 years of education, his MMSE score was 27/30 and his 3MS score was 67/100.

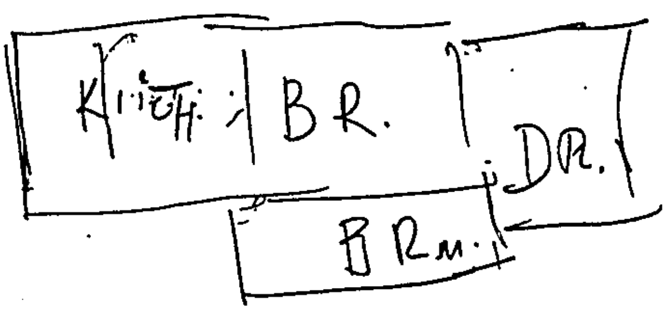


Figure 12. This sketch map of the new environment, drawn by the same AD subject shown in Figure 11, was completed on the initial visit. It was scored by comparing it with the map of the new environment (see Figure 10).

control group. Two of the AD subjects were unable to complete the functional questionnaire due to fatigue. The two caregivers who did not return the questionnaire were not the caregivers of the two AD subjects who did not complete the questionnaire. One AD subject did not consent to giving the questionnaire to her caregiver. All of the control subjects were able to complete the questionnaire, however, the proxies of twelve of these subjects did not return or respond to the questionnaire. Compliance was not a problem with the AD group's caregivers because they attended the testing sessions, and were asked to complete the questionnaires at the test site.

The differences between self- and proxy-ratings on the Functional Spatial Abilities Questionnaire were calculated for each group and were then analyzed using two-tailed paired Student's t-tests. These data were tested for statistical differences from zero. While there was no difference between the self- and proxy-ratings for the control group (mean difference = -0.3, SEM: 0.2, $t(84) = -1.9$, $p > .05$), the AD group's self-ratings were significantly higher than ratings by their caregivers (mean difference = 2.8, SEM: 1.1, $t(20) = 2.5$, $p < .01$).

As mentioned in the Methods section, an item analysis was conducted on the Functional Spatial Abilities Questionnaire in order to determine whether the items could individually discriminate between individuals who have AD and those who do not. The analysis used the Wilcoxon-rank sum test. These results are shown in Appendix J for the self-rated responses and in Appendix K for the proxy-rated responses. There was a significant difference between the two groups on all items except for item 9 when it was self-rated. This item states, "I require supervision when I travel in the neighbourhood". There was a disagreement between the AD subjects and their caregivers on this item.

3.3 Correlation Between SOS Subtests and Demographic/Mental Status Variables

Pearson's correlation coefficients between the 13 subtests and age, education and mental status, are listed in Table 6 for the AD group and in Table 7 for the control group. For the control group, age was negatively correlated with scores on four subtests: Figure-Ground Perception, Corsi Block, Stylus Maze and sketch map of the new environment. For the AD group, only the scores on the Stylus Maze were negatively correlated with age. The AD subjects' scores on the Corsi Block and on the self-rated Functional Spatial Abilities Questionnaire were positively correlated with age. For both groups, scores on the Porteus Maze were positively correlated with years of education, but for the control group scores on the Figure-Ground Perception and Corsi Block tests were also related to years of education whereas this was not the case for the AD group. With respect to the mental status of the control group, scores on the 3MS were related to scores on four of the subtests (Figure-Ground Perception, Spatial Relations, Porteus Maze and Road-Map tests) whereas scores on the MMSE were significantly correlated with only the Figure-Ground Perception Test. For the AD group, scores on both the 3MS and MMSE were related to scores on many more subtests than for the control group. Scores on both mental status exams were significantly correlated with the same subtests for the AD group with the exception of the Stylus Maze Test scores which were significantly correlated with only the 3MS.

3.4 Test-Retest and Inter-Rater Reliability Coefficients of SOS Subtests

The intra-class correlation coefficients (ICCs) were used as indices of test-retest and inter-rater reliabilities for subtests of the SOS Battery. For the AD group, all ICCs, except for those

Table 6

Correlation of SOS subtests with demographic/mental status variables for AD group

	Age	Education	MMSE	3MS
Left-Right[a]	.20 (25)	.23 (25)	**.64**** (25)	**.57*** (25)
FGP[b]	-.02 (22)	.08 (22)	.24 (22)	.28 (22)
Position in Space	.24 (25)	.20 (25)	**.63**** (25)	**.63**** (25)
Spatial Relations	.09 (25)	.25 (25)	.17 (25)	.35 (25)
Porteus Maze	.03 (24)	**.45*** (24)	**.43*** (24)	**.47*** (24)
Map-Read	-.01 (24)	-.05 (24)	.16 (24)	.26 (24)
Road-Map	-.05 (24)	.37 (24)	**.54*** (24)	**.59*** (24)
Corsi Block	**.49*** (25)	-.03 (25)	**.46*** (25)	**.51*** (25)
Stylus Maze	**-.43*** (22)	.13 (22)	.27 (22)	**.45*** (22)
Sketch-map: Fam[c]	.03 (23)	.13 (23)	**.71**** (23)	**.69**** (23)
Sketch-map: New[d]	-.14 (24)	-.03 (24)	**.56*** (24)	**.50*** (24)
FSAQ: Self-rated[e]	**.47*** (23)	-.04 (23)	.07 (23)	.05 (23)
FSAQ: Proxy-rated[f]	.11 (23)	.03 (23)	.24 (23)	.34 (23)

Note. Numbers in brackets indicate the sample sizes.

[a]Left-Right Discrimination

[b]Figure-Ground Perception

[c]Sketch-map of familiar environment

[d]Sketch-map of new environment

[e]Functional Spatial Abilities Questionnaire: Self-rated

[f]Functional Spatial Abilities Questionnaire: Proxy-rated

*p<.05, **p<.001, ***p<.0001, Pearson's product-moment correlation coefficients.

Table 7

Correlation of SOS subtests with demographic/mental status variables for control group

	Age	Education	MMSE	3MS
Left-Right[a]	-.14 (97)	.13 (97)	.18 (97)	.13 (97)
FGP[b]	**-.26*** (97)	**.43***** (97)	**.21*** (97)	**.39***** (97)
Position in Space	-.05 (97)	.09 (97)	-.05 (97)	.08 (97)
Spatial Relations	-.16 (97)	-.02 (97)	.14 (97)	**.22*** (97)
Porteus Maze	-.15 (97)	**.25** *(97)	-.02 (97)	**.20*** (97)
Map-Read	-.35 (96)	.19 (96)	.07 (96)	.08 (96)
Road-Map	-.20 (95)	.18 (95)	-.05 (95)	**.32*** (95)
Corsi Block	**-.29*** (97)	**.22*** (97)	.05 (97)	.17 (97)
Stylus Maze	**-.30*** (96)	.12 (96)	.18 (96)	.17 (96)
Sketch-map: Fam[c]	-.18 (82)	-.05 (82)	-.00 (82)	.21 (82)
Sketch-map: New[d]	**-.24*** (95)	.16 (95)	.02 (95)	.13 (95)
FSAQ: Self-rated[e]	-.07 (97)	.17 (97)	-.08 (97)	.10 (97)
FSAQ: Proxy-rated[f]	-.01 (85)	.08 (85)	.07 (85)	.18 (85)

Note. Numbers in brackets indicate the sample sizes.

[a]Left-Right Discrimination

[b]Figure-Ground Perception

[c]Sketch-map of familiar environment

[d]Sketch-map of new environment

[e]Functional Spatial Abilities Questionnaire: Self-rated

[f]Functional Spatial Abilities Questionnaire: Proxy-rated3

*p<.05, **p<.001, ***p<.0001, Pearson's product-moment correlation coefficients.

Table 8

Test-retest and inter-rater reliabilities for AD group

	Test-retest		Inter-rater	
Subtest	n	ICC	n	ICC
Left-Right[a]	24	**.49***	23	**.51***
FGP[b]	20	**.83*****	15	**.30***
Position in Space	24	**.82*****	23	**.87*****
Spatial Relations	24	.22	23	.28
Porteus Maze	23	**.89*****	21	**.71*****
Map-Read	23	**.96*****	22	**.90*****
Road-Map	21	**.87***	21	**.71*****
Corsi Block	24	**.66****	23	**.68*****
Stylus Maze	18	**.67****	16	**.63***
Sketch-map: Fam[c]	20	**.88*****	17	**.83*****
Sketch-map: New[d]	23	**.49****	20	**.89*****
FSAQ: Self-rated[e]	21	**.84*****	20	**.82*****
FSAQ: Proxy-rated[f]	20	**.85*****	16	**.81*****

[a]Left-Right Discrimination

[b]Figure-Ground Perception

[c]Sketch-map of familiar environment

[d]Sketch-map of new environment

[e]Functional Spatial Abilities Questionnaire: Self-rated

[f]Functional Spatial Abilities Questionnaire: Proxy-rated

*p<.05, **p<.001, ***p<.0001, intra-class correlation coefficient.

for the Spatial Relations Test, reached statistical significance (Table 8). The majority of these significant coefficients met the .70 criterion for adequate reliability. Specifically, 8 of the 13 subtests, demonstrated adequate test-retest or inter-rater reliability. For the control group (see Table 9), only three subtests met the criterion for acceptable test-retest reliability and four met the criterion for adequate inter-rater reliability.

3.5 Internal Consistency

As described in the Methods section, internal consistency was assessed in four ways. First, subtest-to-subtest correlation coefficients were calculated. Next, subtest-to-total score correlation coefficients and Cronbach's alphas were estimated. Last, two factor analyses were performed.

3.5.1 Subtest-to-subtest correlations

Pearson's product-moment correlation coefficients were calculated between the scores obtained on each of the subtests of the SOS Battery. The coefficients for the perceptual subtests are listed in Table 10 for the AD and control groups. Three of the six coefficients reached significance in the AD group whereas two reached significance in the control group. All of these coefficients involved scores obtained on the Position in Space Test and were higher for the AD group (average $r = .56$) than for the control group (average $r = .28$). The coefficients between cognitive subtests are presented in Table 11 for the AD group and in Table 12 for the control group. Ten of the twenty-one between-subtest correlation coefficients reached significance for the AD group in comparison to eight coefficients for the control group. Most of the correlations

Table 9

Test-retest and inter-rater reliabilities for control group

	Test-retest		Inter-rater	
Subtest	n	ICC	n	ICC
Left-Right[a]	33	-.05	27	**.35***
FGP[b]	33	**.71****	27	**.73****
Position in Space	33	**.44***	27	**.39***
Spatial Relations	33	**.43***	27	-.15
Porteus Maze	33	.19	27	.03
Map-Read	33	**.87****	27	**.89****
Road-Map	32	.26	27	**.68****
Corsi Block	33	**.43***	27	.28
Stylus Maze	33	.17	27	.03
Sketch-map: Fam[c]	28	-.02	22	.13
Sketch-map: New[d]	33	**.31***	26	**.62****
FSAQ: Self-rated[e]	33	**.78****	27	**.83****
FSAQ: Proxy-rated[f]	23	**.67****	22	**.81****

[a]Left-Right Discrimination

[b]Figure-Ground Perception

[c]Sketch-map of familiar environment

[d]Sketch-map of new environment

[e]Functional Spatial Abilities Questionnaire: Self-rated

[f]Functional Spatial Abilities Questionnaire: Proxy-rated

*p<.05, **p<.001, ***p<.0001, intra-class correlation coefficient.

Table 10

Correlation between perceptual subtests for AD and control groups

	1		2		3		4	
	AD	C	AD	C	AD	C	AD	C
1. Left-Right[a]	—	—	.33 (22)	.17 (97)	**.53*** (25)	.15 (97)	.33 (25)	.18 (97)
2. FGP[b]			—	—	**.66**** (22)	**.33*** (97)	.32 (22)	.35 (97)
3. Position in Space					—	—	**.50*** (25)	**.24*** (97)
4. Spatial Relations							—	—

Note. Numbers on top represent corresponding subtest in left column. Numbers in brackets indicate the sample sizes.

[a]Left-Right Discrimination

[b]Figure-Ground Perception

*$p<.05$, **$p<.001$, Pearson's product-moment correlation coefficients.

Table 11

Correlation between cognitive subtests for AD group

	1	2	3	4	5	6	7
1. Porteus Maze	—	**.58**** (23)	.36 (23)	**.42*** (24)	**.49*** (21)	**.51*** (22)	.24 (23)
2. Map-Read		—	.27 (23)	.30 (24)	**.54*** (22)	.25 (22)	**.48*** (23)
3. Road-Map			—	.12 (24)	.41 (22)	**.53*** (22)	.40 (23)
4. Corsi Block				—	.14 (22)	**.48*** (23)	.25 (24)
5. Stylus Maze					—	.28 (20)	**.46*** (21)
6. Sketch-map: Fam[a]						—	**.52*** (22)
7. Sketch-map: New[b]							—

Note. Numbers on top represent corresponding subtest in left column. Numbers in brackets indicate the sample sizes.

[a]Sketch-map of familiar environment

[b]Sketch-map of new environment

*p<.05, **p<.001, Pearson's product-moment correlation coefficients.

Table 12

Correlation between cognitive subtests for control group

	1	2	3	4	5	6	7
1. Porteus Maze	—	.13 (96)	**.46***** (95)	.18 (97)	.16 (96)	.18 (82)	.17 (95)
2. Map-Read		—	**.38**** (95)	**.34**** (96)	.06 (96)	**.35*** (82)	.10 (95)
3. Road-Map			—	**.25*** (95)	.02 (95)	**.28*** (81)	.13 (94)
4. Corsi Block				—	.03 (96)	.11 (82)	**.22*** (95)
5. Stylus Maze					—	**.28*** (82)	-.03 (95)
6. Sketch-map: Fam[a]						—	.01 (82)
7. Sketch-map: New[b]							—

Note. Numbers on top represent corresponding subtest in left column. Numbers in bracket indicate the sample size.

[a]Sketch-map of familiar environment

[b]Sketch-map of new environment

*p<.05, **p<.001, ***p<.0001, Pearson's product-moment correlation coefficients.

that reached significance for the AD group were not significant for the control group and vice versa. As with the perceptual subtests, the coefficients between cognitive subtests were generally higher for the AD group (average $r = .50$) in comparison to those for the control group (average $r = .32$).

Table 13 lists Pearson's correlation coefficients between perceptual and cognitive subtests by group. For the AD group, 14 of the 28 between-subtest correlations were significant and 13 correlations were significant for the control group. The majority, or eight of these coefficients represented correlations between the same subtests for the AD and control groups. Again, the significant coefficients were on average higher for the AD group (average $r = .52$) in comparison to those for the control group (average $r = .34$). Specifically, for the AD group scores on the Porteus Maze Test and on the sketch map of the familiar environment were related to scores on each of the four perceptual subtests. Again, scores on the Position in Space Test were related to the majority of the cognitive subtests for the AD patients. For the control group, scores on the Porteus Maze, Map-Reading, Road-Map and Corsi Block were related to scores on the perceptual subtests especially the Figure-Ground Perception, Position in Space and Spatial Relations tests. Scores on the sketch map of the familiar environment were not strongly related to the perceptual subtest scores.

Table 14 presents Pearson's correlation coefficients between scores obtained on the Functional Spatial Abilities Questionnaires and each of the perceptual and cognitive subtests. The correlation between AD subjects' self-ratings and their proxies' ratings was moderate ($r = .39$) but not statistically significant ($p > .05$). The self-ratings of the control subjects were, however, highly correlated with the ratings of their proxies ($r = .68$, $p < .0001$). For both the

Table 13

Correlation between perceptual and cognitive subtests for AD and control groups

	Left-Right[a]		FGP[b]		Position in Space		Spatial Relations	
	AD	C	AD	C	AD	C	AD	C
Porteus	.57*	.06	.51*	.37**	.67**	.23*	.55*	.28*
Maze	(24)	(97)	(21)	(97)	(24)	(97)	(24)	(97)
Map-Read	.32	.08	.31	.57***	.43*	.34*	.31	.38***
	(24)	(96)	(22)	(96)	(24)	(96)	(24)	(96)
Road-Map	.28	.25*	.03	.44***	.47*	.26*	.48*	.47***
	(24)	(95)	(22)	(95)	(24)	(95)	(24)	(95)
Corsi	.42*	.01	.41	.28**	.69***	.29*	.43*	.15
Block	(25)	(97)	(22)	(97)	(25)	(97)	(25)	(97)
Stylus	.28	.08	.11	.18	.18	.03	.34	.08
Maze	(22)	(96)	(21)	(96)	(22)	(96)	(22)	(96)
Sketch-map:	.45*	-.01	.45*	.23*	.70**	.03	.45*	.18
Fam[c]	(23)	(82)	(20)	(82)	(23)	(82)	(23)	(82)
Sketch-map:	.35	.09	-.08	.12	.32	-.03	.04	.19
New[d]	(24)	(95)	(21)	(95)	(24)	(95)	(24)	(95)

Note. Numbers in brackets indicate the sample size.

[a]Left-Right Discrimination

[b]Figure-Ground Perception

[d]Sketch-map of familiar environment

[e]Sketch-map of new environment

*p<.05, **p<.001, ***p<.0001, Pearson's product-moment correlation coefficients.

Table 14

Correlation between FSAQ[a] and perceptual/cognitive subtests for AD and control groups

	FSAQ Self-rated		FSAQ Proxy-rated	
	AD n=20-23	C n=82-97	AD n=20-23	C n=75-97
Left-Right[b]	.34	-.09	.25	-.10
FGP[c]	-.11	.08	.00	.01
Spatial Relations	.08	.04	.11	-.01
Position in Space	.00	.01	.19	-.04
Porteus Maze	.07	.09	.27	.07
Map-Read	.37	**.29***	.36	.20
Road-Map	.11	.18	.32	.10
Corsi Block	.20	.13	.33	.21
Stylus Maze	.06	-.06	.40	.09
Sketch-map: Fam[d]	.04	.04	.14	.12
Sketch-map: New[e]	.18	.15	.17	-.03

[a]Functional Spatial Abilities Questionnaire

[b]Left-Right Discrimination

[c]Figure-Ground Perception

[d]Sketch-map of familiar environment

[e]Sketch-map of new environment

*p<.005, Pearson's product-moment correlation coefficients.

AD and control groups, scores on the Functional Spatial Abilities Questionnaire were not significantly correlated with those on any of the subtests, except for the Map-Reading Test, which was correlated with the self-rated Functional Spatial Abilities Questionnaire for the AD group (Table 14).

In summary, the subtest-to-subtest correlation coefficients were generally higher for the AD group than for the control group. These coefficients presented a pattern which suggests that perceptual subtests scores were related to cognitive subtest scores to a similar extent that subtests within these categories were related to themselves. With respect to functional spatial orientation, scores on the Functional Spatial Abilities Questionnaire were clearly not correlated with performance on the other subtests of the SOS Battery.

3.5.2 Subtest-to-total score correlation coefficients

Table 15 presents subtest-to-total score Pearson correlation coefficients by group for raw and standardized scores. As mentioned, these total scores did not include the score of the subtest with which it was being correlated. Both raw and standardized data provided similar results. For the AD group, all of the subtests met the $r = .20$ criterion. In contrast, three subtests did not meet the criterion for the control group. These were the Left-Right Discrimination, Stylus Maze and sketch map of the new environment.

3.5.3 Internal consistency: Cronbach's coefficient alpha

Cronbach's coefficient alphas are presented in Table 16 for the AD and control groups. Again, both raw and standardized data were used in calculating Cronbach's alphas. For the AD

Table 15

Subtest-to-total Pearson's product-moment correlation coefficients for AD and control groups

	AD group		Control group	
Subtest	Raw[a]	Standardized[b]	Raw[a]	Standardized[b]
Left-Right[c]	.58	.61	.16	.14
FGP[d]	.34	..39	.61	.55
Position in Space	.67	.74	.38	.33
Spatial Relations	.54	.53	.47	.44
Porteus Maze	.73	.72	.37	.41
Map-Read	.56	.61	.59	.57
Road-Map	.41	.51	.51	.57
Corsi Block	.52	.57	.39	.38
Stylus Maze	.53	.49	.12	.15
Sketch-map: Fam[e]	.56	.65	.38	.30
Sketch-map: New[f]	.45	.44	.15	.18
FSAQ: Self-rated[g]	.23	.23	.29	.28
FSAQ: Proxy-rated[h]	.34	.39	.20	.22

[a]Raw score correlation with total score

[b]Standardized score correlation with total score

[c]Left-Right Discrimination

[d]Figure-Ground Perception

[e]Sketch-map of familiar environment

[f]Sketch-map of new environment

[g]Functional Spatial Abilities Questionnaire: Self-rated

[h]Functional Spatial Abilities Questionnaire: Proxy-rated

Table 16

Cronbach's alpha coefficients for AD and control groups

	AD group		Control group	
	Raw	Standardized	Raw	Standardized
Overall alpha	.75	.86	.64	.72
Subtests	Alpha[a]	Alpha[b]	Alpha[a]	Alpha[b]
Left-Right[c]	.74	.85	.64	.73
FGP[d]	.74	.86	.53	.68
Position in Space	.71	.84	.63	.71
Spatial Relations	.74	.85	.63	.70
Porteus Maze	.70	.84	.62	.70
Map-Read	.72	.85	.58	.68
Road-Map	.74	.85	.57	.68
Corsi Block	.74	.85	.63	.70
Stylus Maze	.74	.85	.64	.73
Sketch-map: Fam[e]	.75	.85	.64	.71
Sketch-map: New[f]	.75	.86	.64	.73
FSAQ: Self-rated[g]	.75	.87	.62	.72
FSAQ: Proxy-rated[h]	.74	.86	.63	.72

[a]Cronbach's alpha of battery after removing subtest and using raw scores

[b]Cronbach's alpha of battery after removing subtest and using standardized scores

[c]Left-Right Discrimination, [d]Figure-Ground Discrimination

[e]Sketch-map of familiar environment, [f]Sketch-map of new environment

[g]Functional Spatial Abilities Questionnaire: Self-rated, [h]Functional Spatial Abilities Questionnaire: Proxy-rated

group, the overall alpha for the full battery were .75 using raw scores and .86 using standardized scores. For the control group, the overall alphas for the full battery were .64 using raw scores and .72 using standardized scores. Given that little is known about the contribution of each subtest to the measurement of spatial orientation, standardized scores, which give equal weight to each subtest, are more meaningful to interpret. For the AD group, the overall alpha based on standardized scores met the .80 criterion which is indicative of high internal consistency (Carmines & Zeller, 1979, Feinstein, 1987). In contrast, the overall alpha based on standardized scores of the control group did not meet this criterion although it was not much lower.

Table 16 also lists Cronbach's coefficient alphas for the full battery after removal of individual subtests. The highest Cronbach's alpha values (.73) were associated with the Left-Right Discrimination, Stylus Maze and sketch-map of the new environment tests, however, these values were not different from those of the functional questionnaires which had values of .72. For the control group, the highest Cronbach's alpha value (.87) was associated with the self-rated Functional Spatial Abilities Questionnaire.

To summarize, coefficient alphas were generally higher for the AD group. In both groups, coefficient alphas did not fluctuate greatly as each subtest was removed.

3.5.4 Factor analysis using control group data

Two factor analyses were conducted. A factor analysis was initially conducted using only the data from the control group and all 13 subtests of the SOS Battery. Following this, a shortened version of the battery was derived which permitted a second set of factor analyses to be conducted using data from both the control and the AD groups. In the first analysis, five

Table 17

Factor loadings of subtests for control group

Factor	1	2	3	4	5
eigenvalue	3.49	1.76	1.29	1.10	1.04
Proportion of variance explained	.27	.14	.10	.08	.08
Left-Right[a]	.11	-.07	-.14	**.78**	.13
FGP[b]	**.50**	**.49**	.04	.18	.22
Position in Space	.21	.19	.11	**.75**	-.02
Spatial Relations	**.64**	.23	-.01	.13	.22
Porteus Maze	**.82**	-.00	-.03	.04	-.05
Map-Read	.30	**.60**	.32	.20	.30
Road-Map	**.79**	.18	.15	.17	.07
Corsi Block	-.05	.25	.20	.30	**.68**
Stylus Maze	.04	**.61**	-.12	.18	-.22
FSAQ: Self-rated[c]	-.13	-.06	**.90**	.02	.06
FSAQ: Proxy-rated[d]	-.05	.02	**.89**	-.06	.00
Sketch-map: Fam[e]	.15	**.80**	-.04	-.16	.06
Sketch-map: New[f]	.23	-.17	-.09	-.08	**.80**

[a]Left-Right Discrimination

[b]Figure-Ground Perception

[c]Functional Spatial Abilities Questionnaire: Self-rated

[d]Functional Spatial Abilities Questionnaire: Proxy-rated

[e]Sketch-map of familiar environment

[f]Sketch-map of new environment

factors emerged (Table 17). All subtests except for the Figure-Ground Perception Test loaded on only one of the five factors.

The scores on the Figure-Ground Perception, Spatial Relations, Porteus Maze and Road-Map tests loaded on Factor 1. These subtests examine the concepts of mental rotation, planning and left-right orientation. The scores on the Figure-Ground Perception, Map-Reading, Stylus Maze and sketch map of the familiar environment tests loaded on Factor 2. These subtests require mental representation, visuospatial construction ability, spatial learning, mental rotation and long-term memory. The Map-Reading Test and sketch map of the familiar environment both involve a motoric aspect to representation of space. While the Map-Reading Test requires the use of whole body movements, the sketch map of familiar environment involves the subject's interpretation of his or her movement experiences in an environment. Factor 3 consisted of variables from scores on the self-rated and proxy-rated Functional Spatial Abilities Questionnaire. The scores on the Left-Right Discrimination and Position in Space tests loaded on Factor 4. These tests are perceptual in nature and require the ability to recognize the orientation of geometric figures, and to distinguish left from right. Lastly, the two tests loading on Factor 5, the Corsi Block and the sketch map of the new environment, were clearly tests of immediate or short-term spatial memory.

3.5.5 Selection of subtests for a shortened SOS Battery

The shortened version consisted of 6 of the 13 subtests and the selection of these tests was based on the criteria described in the Methods section. The Left-Right Discrimination and Spatial Relations tests were not included in the shortened battery because of their low test-retest and inter-rater reliability. Although the Figure-Ground Perception Test demonstrated good test-retest and inter-rater reliability, it was excluded because it was an extremely difficult test for the AD subjects. All of the raters agreed that it was the subtest requiring the longest amount of time to administer to both the AD and control groups. Although the time taken to administer each subtest was not recorded, raters estimated that the Figure-Ground Perception Test took between 20 to 40 minutes to complete. It also loaded on Factors 1 and 2 suggesting that it involved skills that overlapped with those evaluated by the other spatial tests. The following four tests were not chosen based on the reasons provided: scores on the Porteus Maze Test were associated with level of education in both groups, scores on the Corsi Block Test were correlated with age in the AD group and with both age and education in the control group, scores on the Stylus Maze Test correlated with age in both groups and the self-rated Functional Spatial Abilities Questionnaire correlated with age in the AD group.

The final shortened version of the SOS Battery consisted of the Road-Map Test loading on Factor 1, the Map-Reading Test and sketch map of familiar environment loading on Factor 2, the proxy-rated Functional Spatial Abilities Questionnaire loading on Factor 3, Position in Space loading on Factor 4 and sketch map of new environment loading on Factor 5. Two subtests were selected from Factor 2 because they both appeared clinically relevant. Both tests were quick to administer and were well accepted by the subjects. Although the two subtests loaded on the same factor, they were not significantly correlated with each other for the AD group (Pearson's $r = .25$, $p > .05$) and only slightly correlated with each other for the control group (Pearson's $r = .35$, $p < .05$). Thus, it may be assumed that these two subtests were not redundant. Performance on all subtests of the shortened version was not associated with age or education in either group, except for the sketch map of the new environment which was negatively correlated with age for the control group. The sketch map of the new environment was selected because it was less associated with age and education than the Corsi Block Test and it was the only other subtest representing Factor 5. Although the recommended maximum of variables is 5 for a sample of 25 subjects (Tabachnick & Fidell, 1989), all 6 subtests were kept for the shortened battery because each appeared clinically relevant. Each of the six tests was converted to a score out of 20. The total battery score out of 120 was then converted to a total score out of 100. Table 18 and Figure 13 provide standardized scores on the six subtests and the standardized total score for the shortened SOS Battery.

3.5.6 <u>Factor analyses of the shortened SOS Battery using AD and control group data</u>

The results from the factor analysis of the shortened version of the SOS Battery are presented in Table 19. For the control group, three factors had eigenvalues that were equal to or greater than the criterion of 1 (Tabachnick & Fidell, 1989). A strict application of the minimum eigenvalue criterion would result in selecting only two factors for the AD group. However, in order to facilitate comparison, the third factor was also accepted. It should be noted that the eigenvalue for factor 3 (.93) was close to 1 and substantially greater than the fourth factor which was .64.

For the control group, the results of this factor analysis replicated those of the first factor analysis to the extent that scores on the proxy-rated Functional Spatial Abilities Questionnaire and scores on the sketch map of the new environment again loaded on separate factors. In addition, scores on the Map-Reading and sketch map of the familiar environment loaded on the

Table 18

Standardized subtest and total scores on shortened version of SOS Battery: Control compared to AD group

	AD (GDS 3 & 4)			CONTROL			
	n	range	M (SD)	n	range	M (SD)	t (df)
Position in Space	25	5-20	13.7 (4.5)	97	12-20	18.9 (1.4)	**-5.7**** (25.3)
Road-Map	24	0-20	12.2 (4.3)	95	9-20	17.7 (2.8)	**-6.0**** (28.0)
Map-Read	24	0-20	2.9 (4.7)	96	2.3-20	14.9 (4.8)	**-10.9**** (118.0)
Sketch-map: Fam[a]	24	0-20	5.6 (6.0)	82	8.6-20	18.7 (2.4)	**-5.4*** (22.2)
Sketch-map: New[b]	23	0-20	10.3 (7.4)	95	1.7-20	18.0 (3.6)	**-9.7*** (27.3)
FSAQ: Proxy-rated[c]	23	7.8-18.9	14.2 (3.1)	85	13.9-20	19.6 (1.1)	**-8.2**** (23.4)
Total Score	19	26.5-72.8	48.5 (13.9)	74	63.3-100.0	89.6 (8.0)	**-12.4**** (21.2)

Note. Each of the six subtests was standardized to a maximum of 20 points. The maximum total battery score of 120 was then converted to 100.

[a]Sketch-map of familiar environment

[b]Sketch-map of new environment

[c]Functional Spatial Abilities Questionnaire: Proxy-rated

*$p<.001$, **$p<.0001$, Student's t-test for independent samples.

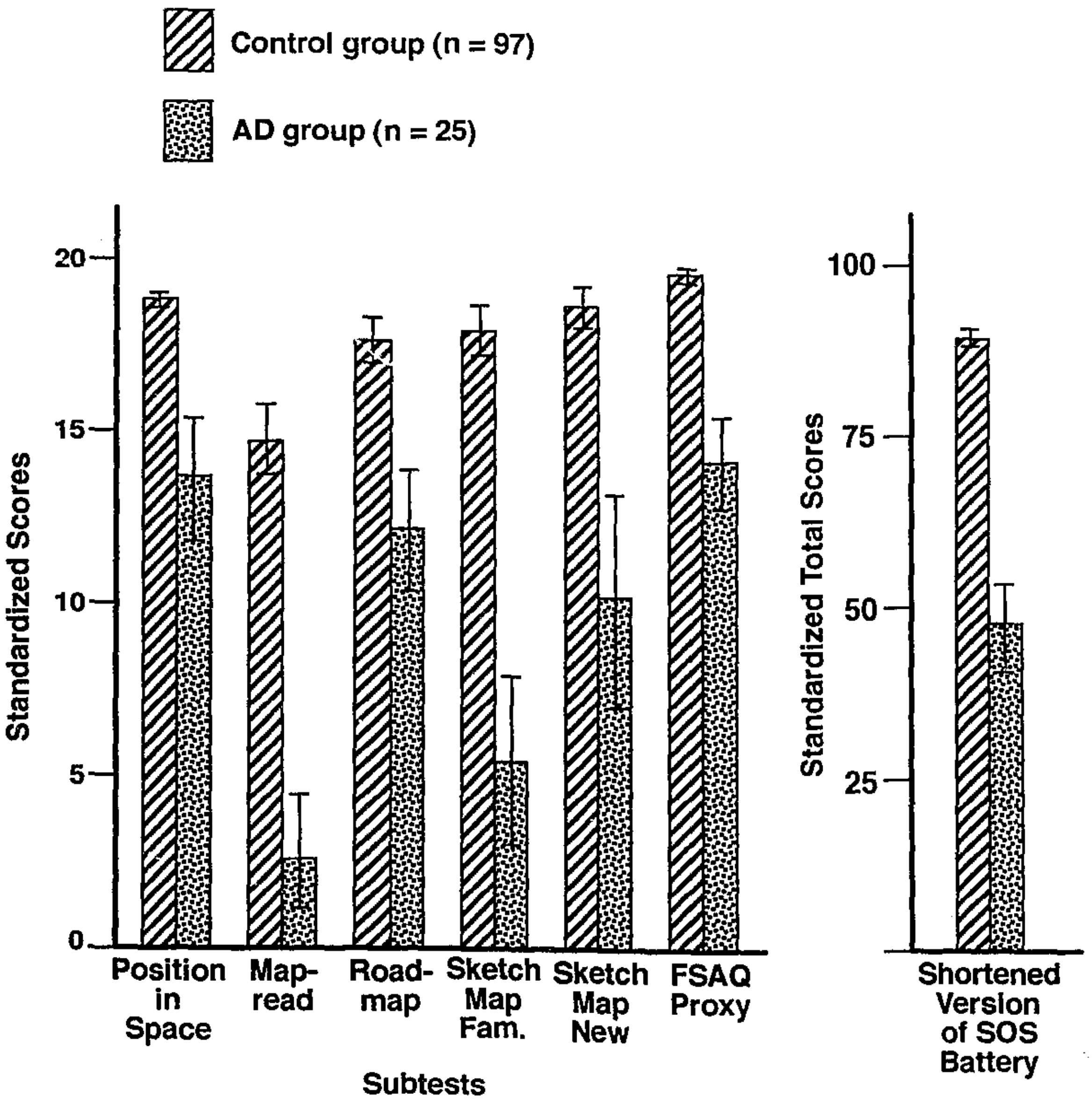


Figure 13. Bar graphs of scores on the shortened version of SOS Battery. The bars indicate 95% confidence intervals.

Table 19

Factor loadings of subtests on shortened version of SOS Battery for AD and control groups

Factor	1		2		3	
Group	AD	C	AD	C	AD	C
Eigenvalue	2.26	1.94	1.52	1.03	0.93	1.00
Proportion of variance explained	.38	.32	.25	.17	.16	.17
Position in Space	**.86**	**.63**	.31	-.28	.07	-.09
Map-Read	-.06	**.78**	**.87**	.17	.09	.10
Road-Map	**.81**	**.66**	-.28	.18	-.09	.20
Sketch-map: Fam[a]	**.77**	**.66**	.13	.01	**.43**	-.07
Sketch-map: New[b]	.09	.03	.07	-.03	**.96**	**.98**
FSAQ: Proxy-rated[c]	.13	.08	**.79**	**.95**	.02	-.04

[a]Sketch-map of familiar environment

[b]Sketch-map of new environment

[c]Functional Spatial Abilities Questionnaire: Proxy-rated

same factor as they did in the first analysis. In contrast to the previous analysis, these results showed that scores of the Position in Space and Road-Map tests, which had represented separate factors, are now loaded on the same factor as two other subtests.

For the AD group, the structure of the factors was more complex than for the control group. Scores on four of the six subtests loaded on the same factors as they did for the control group: Position in Space Test, Road-Map Test, sketch map of the new environment and proxy-rated Functional Spatial Abilities Questionnaire. In contrast, performance on the Map-Reading Test loaded on the factor associated with the proxy-rated Functional Spatial Abilities Questionnaire. Another difference was that the AD group's scores on the sketch map of the familiar environment loaded on two factors, one associated with performance on the Position in Space and Road-Map tests and another associated with the sketch map of the new environment.

3.5.7 Subtest-to-total score correlation coefficients for the shortened battery

Subtest-to-total score Pearson's correlation coefficients for the six subtests of the shortened battery are listed in Table 20 by group. These coefficients were calculated using raw and standardized scores as was done for the full battery. The discrepancies between coefficients based on raw data and those based on standardized data were larger than those observed with the full battery. In general, these coefficients were also lower for both groups in comparison to those seen in the full battery. As with the full battery, standardized coefficients for the shortened battery were higher for the AD group than for the control group. For the AD group, all of the coefficients met the .20 criterion for internal consistency and the lowest coefficient ($r = .32$) was associated with proxy-rated Functional Spatial Abilities Questionnaire scores. For the control

Table 20

Subtest-to-total score correlation coefficients and Cronbach's alphas for AD group: Shortened version of SOS Battery

Subtest	Raw Score Correlation with Total Score	Alpha[a]	Standardized Score Correlation with Total Score	Alpha[a]
Position in Space	.54	.45	.62	.72
Map-Read	.40	.49	.50	.75
Road-Map	.44	.44	.58	.73
Sketch map: Fam[b]	.48	.57	.61	.72
Sketch map: New[c]	.53	.57	.54	.74
FSAQ: Proxy-rated[d]	.31	.52	.32	.79

Note. Overall alpha = .56 for raw scores and .78 for standardized scores.

[a]Cronbach's alpha of battery after removing subtest

[b]Sketch-map of familiar environment

[c]Sketch-map of new environment

[d]Functional Spatial Abilities Questionnaire: Proxy-rated

group, two of the six subtests did not meet this criterion: sketch map of the new environment and proxy-rated Functional Spatial Abilities Questionnaire.

3.5.8 Cronbach's alphas for the shortened battery

Cronbach's coefficient alphas for the shortened battery are listed in Table 21 and were calculated in the same manner as those for the full battery. Overall, Cronbach's alphas were lower for the shortened battery in comparison to those for the full battery. For the AD group, the overall alpha using raw scores was .56 and using standardized scores was .78. For the control group, the overall alpha was .39 using raw scores and .51 using standardized scores. With respect to the alpha values obtained after removing individual subtests, the fluctuations were greater than those obtained for the full battery. For the AD group, alphas based on standardized scores were relatively high for the proxy-rated Functional Spatial Abilities Questionnaire (alpha = .79). For the control group, high alpha values were associated with the sketch map of the new environment (alpha = .56) and with the proxy-rated Functional Spatial Abilities questionnaire (alpha = .53).

3.5.9 Internal consistency: Summary

Correlations between subtests within the perceptual and cognitive categories were of the same magnitude as those measured between these categories. The correlations obtained between scores on the Functional Spatial Abilities Questionnaires and scores on the perceptual and cognitive subtests were low and not statistically significant.

Factor analyses were conducted for the full battery using data from the control group and

Table 21

Subtest-to-total score correlation coefficients and Cronbach's alphas for control group: Shortened version of SOS Battery

Subtest	Raw Score Correlation with Total Score	Alpha[a]	Standardized Score Correlation with Total Score	Alpha[a]
Position in Space	.34	.36	.20	.50
Map-Read	.46	.20	.52	.32
Road-Map	.40	.19	.43	.37
Sketch map: Fam[b]	.38	.40	.29	.45
Sketch map: New[c]	.11	.41	.06	.56
FSAQ: Proxy-rated[d]	.18	.37	.12	.53

Note. Overall alpha = .39 for raw scores and .51 for standardized scores.

[a]Cronbach's alpha of battery after removing subtest

[b]Sketch-map of familiar environment

[c]Sketch-map of new environment

[d]Functional Spatial Abilities Questionnaire: Proxy-rated

for the shortened battery using data from both groups. Results from the first analysis revealed a distinct factor for functional spatial abilities and another for short-term spatial memory. Results from the second factor analysis using the data of the AD group must be interpreted with caution due to the small sample size. The full battery shows high internal consistency when administered to AD patients but not to control subjects, possibly due to scores on the Left-Right Discrimination, Stylus Maze and sketch map of the new environment tests. The overall Cronbach's alpha for the shortened battery almost reached criterion for the AD group but was low for the control group.

3.6 Validity

Evidence for content validity was presented earlier in this section. Results of construct and criterion validity will now be examined. Construct validity was evaluated using multiple logistic regression and by correlating the scores on the shortened SOS Battery with scores on the Reintegration to Normal Living Index.

3.6.1 Construct validity

Construct validity was evaluated first by using multiple logistic regression with the full battery. The hypothesis was that the odds ratio would be greater than one which would indicate that higher scores on a subtest were associated with a decreased likelihood of being an AD patient. Table 22 lists the parameter estimates, odds ratios and the 95% confidence intervals for these ratios. The lower limits of the confidence intervals for each of the subtests were above 1 indicating acceptance of the hypothesis. The disproportionately large odds ratios associated with

Table 22

Logistic regression and odds ratios for subtests of SOS Battery

Subtest	Parameter Estimate*	Standard Error	Odds Ratio	Lower Limit	Upper Limit
Left-Right[a]	2.3	0.5	8.3	3.5	25.0
FGP[b]	0.5	0.1	1.7	1.3	2.2
Spatial Relations	2.0	0.4	7.7	3.3	16.7
Position in Space	0.9	0.2	2.4	1.7	3.6
Porteus Maze	0.9	0.2	2.3	1.6	3.3
Map-Read	0.3	0.1	1.3	1.2	1.5
Road-Map	0.4	0.1	1.5	1.3	1.7
Corsi Block	1.4	0.4	4.2	2.0	8.3
Stylus Maze	0.9	0.2	2.5	1.6	3.9
Sketch Map: Fam[c]	8.4	2.1	4.4×10^3	78.3	2.4×10^5
Sketch Map: New[d]	8.5	1.7	4.7×10^3	1.6×10^2	1.4×10^5
FSAQ: Self-rated[e]	0.5	0.1	1.7	1.4	2.0
FSAQ: Proxy-rated[f]	0.5	0.1	1.7	1.4	2.1

Note. Lower and upper limits define 95% confidence intervals for odds ratios.

[a]Left-Right Discrimination

[b]Figure-Ground Perception

[c]Sketch-map of familiar environment

[d]Sketch-map of new environment

[e]Functional Spatial Abilities Questionnaire: Self-rated

[f]Functional Spatial Abilities Questionnaire: Proxy-rated

*$p < .0001$ for all parameter estimates.

the two sketch map tests reflect the numerical range of the scores, and show that there is almost perfect discrimination between the AD and control groups on these two tests.

Construct validity was also evaluated by calculating Pearson's product-moment correlation coefficients between ratings on the Reintegration to Normal Living Index and the total score of the shortened SOS Battery. The analysis tested the hypothesis that there would be a positive and significant correlation between the scores on the Reintegration to Normal Living Index and those on the shortened SOS Battery. Table 23 and Figure 14 show the AD and control groups' ratings on the Reintegration to Normal Living Index. The control subjects rated themselves similarly to their proxy raters on the Reintegration to Normal Living Index, whereas AD subjects rated themselves significantly higher than their proxy raters. There was a significant between-group difference in both proxy- and self-ratings. For the AD group, the total battery scores were not correlated with either the self-rated Reintegration to Normal Living Index scores ($r = .06$, $p > .05$) or the proxy-rated Reintegration to Normal Living Index scores ($r = .11$, $p > .05$). Similarly, for the control group, the total battery scores were not correlated with the self-rated Reintegration to Normal Living Index scores ($r = .09$, $p > .05$), although they were slightly correlated with the proxy-rated Reintegration to Normal Living Index scores ($r = .24$, $p < .05$). Thus the hypothesis was rejected.

3.6.2 <u>Criterion validity</u>

The demographic and mental status of subjects in the GDS 5 AD, GDS 3 and 4 AD and GDS 1 and 2 groups are listed in Table 24. As mentioned, 10 GDS stage 3 and 4 AD and 10 GDS stage 1 and 2 subjects were matched with the 10 GDS stage 5 AD patients with respect to

Table 23

Comparison of the RNL[a] scores obtained by subjects in the control and AD groups

	AD GDS (3 & 4)			CONTROL			
	n	range	M (SD)	n	range	M (SD)	t (df)
RNL: self-rated	22	67.3-95.3	85.8 (7.4)	97	71.2-100.0	92.3 (4.4)	**-4.0*** (24.5)
RNL: proxy-rated	23	22.6-93.8	69.2 (19.8)	85	73.6-98.0	92.6 (4.8)	**-5.6**** (22.7)

Note. Maximum score on RNL = 100.

[a]Reintegration to Normal Living Index

*p<.001, **p<.0001, Student's t-test for independent samples.

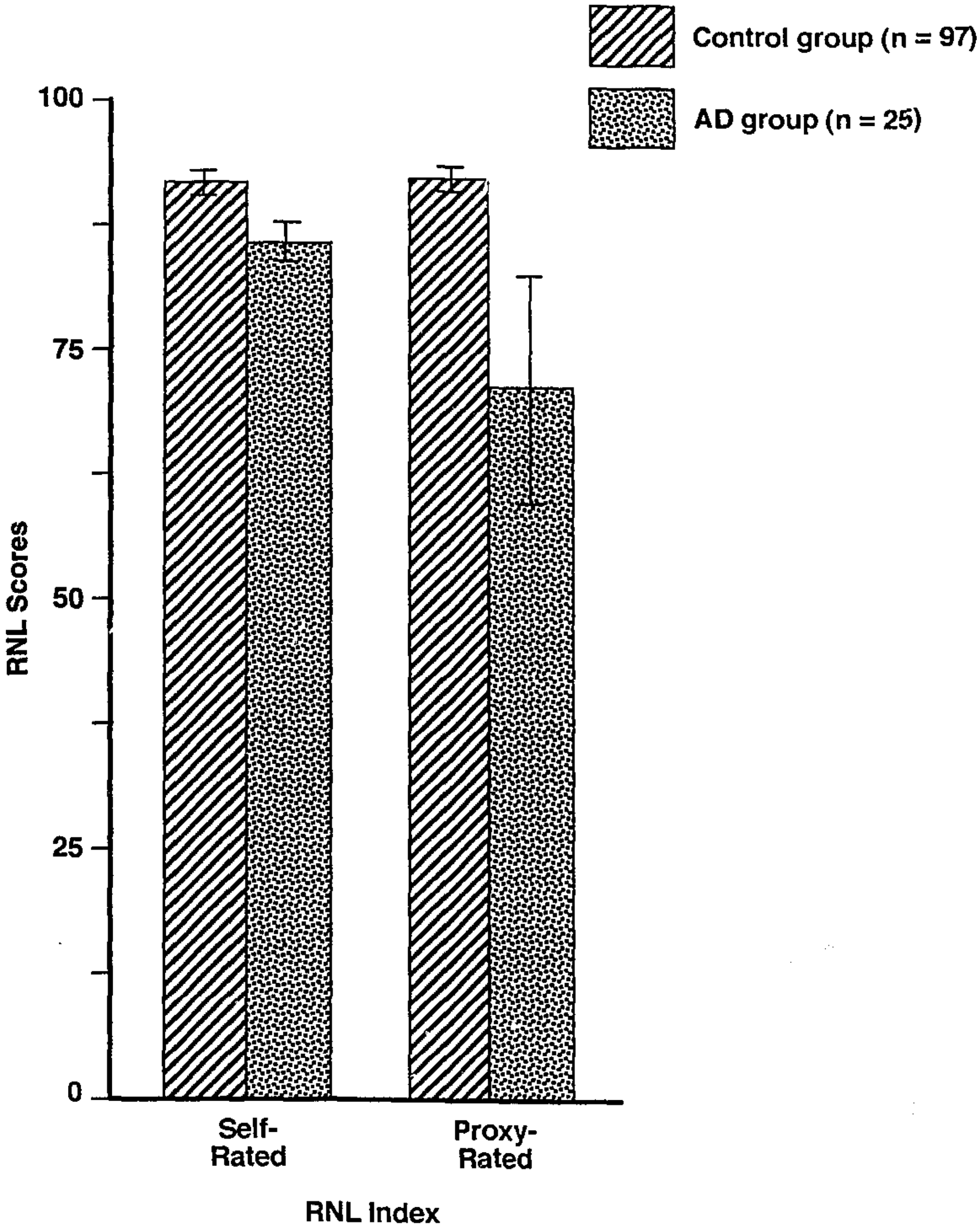


Figure 14. Bar graphs comparing standardized RNL scores. The bars indicate 95% confidence intervals.

Table 24

Comparison of demographic characteristics and mental status of the control, early AD and late AD groups

	GDS 5	GDS 3 & 4	GDS 1 & 2	Chi-square (df)
Male	5	5	5	0.0
Female	5	5	5	(2)
	M (SD)	M (SD)	M (SD)	F (df,df)
Age	68.6	68.6	68.8	0.0
	(7.3)	(6.2)	(7.1)	(2,27)
Education	12.1	10.7	12.8	0.6
	(6.0)	(2.8)	(4.3)	(2,26)
MMSE	11.3[c]	22.8[b]	29.2[a]	**39.9***
	(6.3)	(4.3)	(1.0)	(2,26)
3MS	35.9[c]	72.3[b]	97.5[a]	**59.7***
	(20.0)	(8.8)	(2.4)	(2,26)

Note. Sample size in each group = 10, GDS 5 = late AD, GDS 3 & 4 = early AD, GDS 1 & 2 = control.

*p < .0001, one-way ANOVA.

[a,b,c] Means with the same letter were not significantly different at p<.05, Neuman-Keuls tests.

Table 25

Comparison of performance on the perceptual spatial subtests by control, early AD and late AD groups

	GDS 5	GDS 3 & 4	GDS 1 & 2	
	M (SD)	M (SD)	M (SD)	F (df,df)
Left-Right[1]	7.5[b] (2.5)	8.4[b] (1.2)	9.8[a] (0.4)	**5.3*** (2,27)
FGP[2]	13.4[c] (10.2)	25.0[b] (3.0)	33.3[a] (6.9)	**17.6**** (2,25)
Spatial Relations	6.2[b] (2.6)	7.5[b] (1.4)	9.5[a] (0.7)	**8.8*** (2,27)
Position in Space	9.4[b] (4.5)	11.9[b] (2.7)	14.8[a] (1.8)	**7.0*** (2,26)

Note. Sample size in each group is 10.

GDS 5 = late AD, GDS 3 & 4 = early AD, GDS 1 & 2 = control.

[1]Left-Right Discrimination

[2]Figure-Ground Perception

*p<.05, **p<.0001, one-way ANOVA.

[a,b,c] Means with the same letter were not significantly different at p<.05, Neuman-Keuls tests.

Table 26

Comparison of performance on the cognitive spatial subtests by control, early AD and late AD groups

	GDS 5	GDS 3 & 4	GDS 1 & 2	
	M (SD)	M (SD)	M (SD)	F (df,df)
Corsi Block	2.1[b] (1.6)	2.8[b] (1.6)	4.7[a] (1.3)	**7.5*** (2,26)
Road-Map	8.0[c] (10.2)	21.2[b] (5.4)	28.7[a] (4.8)	**19.8**** (2,24)
Stylus Maze	3.9[b] (1.1)	5.1[b] (1.7)	7.4[a] (1.5)	**8.6*** (2,19)
Porteus Maze	5.2[b] (2.3)	8.0[b] (4.0)	15.5[a] (2.6)	**29.7**** (2,25)
Map-Read	.33[b] (.70)	3.7[b] (6.0)	22.0[a] (9.4)	**30.6**** (2,26)
Sketch map: Fam[1]	.08[c] (.15)	.72[b] (.21)	.94[a] (.09)	**58.1**** (2,23)
Sketch map: New[2]	.02[c] (.06)	.32[b] (.29)	.80[a] (.28)	**21.0**** (2,23)

Note. Sample size in each group is 10.

GDS 5 = late AD, GDS 3 & 4 = early AD, GDS 1 & 2 = control.

[1]Sketch-map of familiar environment

[2]Sketch-map of new environment

*$p<.05$, **$p<.0001$, one-way ANOVA.

[a,b,c] Means with the same letter are not significantly different at $p<.05$, Neuman-Keuls tests.

Table 27

Comparison of performance on the FSAQ[1] by control, early AD and late AD groups

	GDS 5	GDS 3 & 4	GDS 1 & 2	
	M (SD)	M (SD)	M (SD)	F (df,df)
FSAQ: Self-rated	28.3[b] (6.6)	29.5[b] (4.8)	34.4[a] (2.1)	**4.5*** (2,25)
FSAQ: Proxy-rated	19.3[c] (3.7)	25.5[b] (5.6)	36.0[a] (0.0)	**36.0**** (2,24)

Note. Sample size in each group is 10.

GDS 5 = late AD, GDS 3 & 4 = early AD, GDS 1 & 2 = control.

[1]Functional Spatial Abilities Questionnaire

*p<.05, **p<.0001, one-way ANOVA.

[a,b,c] Means with the same letter are not significantly different at p<.05, Neuman-Keuls tests.

sex, age and years of education. The results for the criterion validity study are described in Tables 25, 26 and 27. One-way ANOVAs were performed with group as a factor for each subtest and the results indicated that there were significant group differences on all subtests. Post hoc analyses showed that the following subtests were not only sensitive to the presence of AD but were also sensitive to the severity of the disease: Figure-Ground Perception, the Road-Map Test, sketch maps of the familiar and new environments and the proxy-rated Functional Spatial Abilities Questionnaire. Table 27 also shows that the discrepancies between the self-rated and proxy-rated scores on the Functional Spatial Abilities Questionnaire were greatest for the GDS 5 AD group.

The total battery score was calculated for the GDS 5 AD group and the groups of 10 GDS 3 and 4 AD and GDS 1 and 2 subjects using the shortened version of the SOS Battery. The mean total battery score for the GDS 5 AD group was 23.4 (SD = 11.0), and it was 55.8 (SD = 12.7) and 83.6 (SD = 10.9) for the GDS 3 and 4 and GDS 1 and 2 groups respectively. A one-way ANOVA showed that there were significant group differences between these total scores ($F(2,19) = 47.0$, $p < .0001$). Post hoc analysis showed that the GDS 3 and 4 AD group had significantly lower scores than the GDS 1 and 2 group and that the GDS 5 AD group had lower scores than both the GDS 3 and 4 AD and GDS 1 and 2 group ($p < .05$).

3.6.3 Summary of validity studies

Establishing the validity of a test is an ongoing process and should be the focus of future studies. In this study, content validity was evaluated using a panel of six experts. Each subtest had an odds ratio that indicated a higher risk of AD was associated with a lower score on the test. The Reintegration to Normal Living Index scores did not correlate highly with the SOS Battery scores. Lastly, five subtests of the SOS Battery seem to be valid using the GDS as a criterion, and the shortened SOS Battery demonstrated criterion validity when its total scores were compared across the three GDS categories.

3.7 Results of Other Tests of Cognitive Function

As mentioned, other tests were administered concurrently with the SOS Battery. This section presents results on mental status, attention, language, verbal memory span and clock drawing tests.

3.7.1 Mental status

For the late AD group the MMSE scores did not cluster near the minimum score and the means on both mental status tests were almost perfect for the control group. For the AD group, scores on the MMSE were significantly correlated with performance on seven of the spatial tests and scores on the 3MS were significantly correlated with scores on eight subtests. Thus, the two mental status exams appear to measure the same abilities when used with AD patients. However, for the control group, the MMSE correlated with only one subtest while the 3MS correlated with four, suggesting that for normal subjects a greater range of scores is obtained with the 3MS.

The test-retest and inter-rater reliabilities of the mental status examinations were also examined. All reliability coefficients were low for the control group. This can be attributed to the small range of scores obtained by this group. For the control group, test-retest ICCs were $r = .05$ ($p > .05$) for the MMSE and $r = .23$ ($p > .05$) for the 3MS, while the inter-rater ICCs were $r = .02$ ($p > .05$) for the MMSE and $r = .43$ ($p < .05$) for the 3MS. For the AD group, test-retest ICCs were $r = .86$ ($p < .0001$) for the MMSE and $r = .85$ ($p < .0001$) for the 3MS, and the inter-rater ICCs were $r = .82$ ($p <.0001$) for the MMSE and $r = .85$ ($p < .0001$) for the 3MS.

There was no difference between the test-retest and inter-rater reliability of the MMSE and the 3MS. Both mental status examination demonstrated low test-retest and inter-rater reliability when used with the control group, and showed high test-retest and inter-rater coefficients when used with the early AD group.

3.7.2 Attention

For the AD group, the median, mode and skewness for Albert's Test were 40, 40 and -2.0 respectively and for the Bells Test they were 31, 30 and -0.8 respectively. For the control group, the median, mode and skewness for Albert's Test were 40, 40, -6.9 and for the Bells Test, they were 34, 35 and -2.1.

The AD patients demonstrated a slight impairment in attention as assessed by these two cancellation tests. The scores for the AD group on the cancellation tests (Albert: $M = 39.6$, $SD = 0.8$; Bells: $M = 31.0$, $SD = 3.2$) were slightly lower than those for the control group (Albert: $M = 40.0$, $SD = 0.1$; Bells: $M = 33.4$, $SD = 2.1$). These differences were statistically significant (Albert: $t(22) = -2.4$, $p < .05$; Bells: $t(27) = -3.3$, $p < .01$). For both groups, scores on the cancellation tests were not correlated with age, years of education or mental status scores.

3.7.3 Language

The AD group was impaired on the Boston Naming Test ($M = 14.5$, $SD = 7.0$) in comparison to the control group (M = 26.0, SD = 3.0), ($t(24) = -7.7$, $p < .0001$). The distribution of scores on the Boston Naming Test was normal in the AD group (median = 15, mode = 20, skewness = 0.0) and negatively skewed in the control group (median = 27, mode = 28, skewness = -1.0). For the AD group, scores on the Boston Naming Test were correlated with performance on both the MMSE ($r = .65$, $p < .001$) and 3MS ($r = .79$, $p < .0001$), but not with age or education. For the control group, scores on the Boston Naming Test were positively correlated with scores on the 3MS ($r = .40$, $p < .0001$), negatively correlated with age ($r = -.22$, $p < .05$), but not with education or the MMSE score.

3.7.4 Verbal memory span (Hebb verbal span)

The AD subjects were capable of repeating a mean verbal span of 5.4 (SD = 1.4) digits which was significantly less than the mean span of the control subjects of 6.6 (SD = 1.4) digits, ($t(120) = -3.8$, $p < .001$). The data were normally distributed for the AD group (median = 5, mode = 5 and skewness = 0.0) but slightly skewed for the control group (median = 7, mode = 8 and skewness = -0.4). Span size was correlated with mental status scores for the AD group (MMSE: $r = .55$, $p < .05$; 3MS: $r = .41$, $p < .04$), but not with age or education. In the control group, verbal span was negatively correlated with age ($r = -.24$, $p < .01$), but not correlated with education or mental status.

3.7.5 Clock drawing test

The AD subjects were significantly more impaired (M = 9.7, SD = 5.6) than the control subjects (M = 19.4, SD = 1.4) on the clock drawing test ($t(24) = -8.4$, $p < .0001$). The distribution of scores for the AD group was skewed to the lower scores (median = 9.3, mode = 5.5 and skewness = 0.27) and there was a ceiling effect in the control group (median = 20, mode = 20 and skewness = -3.6). For the AD group, clock drawing scores were correlated with performance on the MMSE ($r = .42$, $p = .03$) but not on the 3MS and not with age or education. The converse was true for the control group whose scores were correlated with age ($r = -31$, $p < .01$), education ($r = .25$, $p < .01$) and the 3MS scores ($r = .22$, $p < .05$), but not with the MMSE scores. Test-retest coefficients (ICC) for the AD group were $r = .37$ ($p < .05$, $n = 25$) and for the control group $r = .16$ ($p > .05$, $n = 33$). Inter-rater reliability coefficients for the AD group were $r = .59$ ($p < .01$, $n = 18$) and for the control group $r = .68$ ($p < .001$, $n = 27$).

3.7.6 Summary

Performance of the AD patients on tests of attention, word finding ability, verbal memory span and visuospatial construction was impaired as was found for the subtests of the SOS Battery. For the AD group, mental status scores, especially on the MMSE, were related to scores on the Boston Naming Test, verbal memory span and the clock drawing test. For the control group it was the 3MS that correlated best with scores on the Boston Naming Test and clock drawing test. These data suggest that the Boston Naming Test, verbal memory span and clock drawing tests reflect general cognitive functioning. The Bells Test was more sensitive than Albert's Test in detecting attentional deficits in AD patients. Also, the MMSE appeared to be a more sensitive test for the AD group, but the 3MS seems to be more sensitive for normal healthy subjects.

3.8 Summary of Results

The early AD group consisted of 12 individuals in GDS 3 and 13 individuals in GDS 4. Patients in stages 3 and 4 differed with respect to their MMSE and 3MS scores and on only two of the 13 subtests of the original SOS Battery. In all subsequent analyses, subjects in GDS 3 and 4 were grouped together as is commonly done in dementia research.

The SOS Battery differentiated between AD and control subjects. In general, the battery showed good internal consistency. For the AD group, Cronbach's coefficient alphas for the standardized subtest scores were all above the .80 criterion for the original battery, yet did not exceed .90, which would have suggested that the subtests were very similar and evaluated the same set of skills (Nunnally, 1978). For the control group, the alphas were all close to .70.

A preliminary shortened SOS Battery was obtained after examining the correlations between subtest scores and age and education, test-retest and inter-rater reliabilities, and the data obtained from an initial factor analysis based on the full battery and on the control group's scores. Age and education may have been confounding variables if they were significantly correlated with performance on any of the subtests. The small range of scores on some of the subtests was taken into consideration when interpreting the results.

Factor analysis of the shortened version of the SOS Battery produced three factors. One factor represented functional spatial orientation as a distinct concept. Another factor was associated with spatial memory function and the third factor represented spatial orientation skills, which were both perceptual and cognitive in nature. These data clearly show that spatial orientation skills are distinct from spatial memory skills.

As expected, Cronbach's alphas for the shortened version of the SOS Battery were slightly lower than those obtained with the full battery. The lower alphas could result from the combined effects of the smaller number of subtests in the shortened battery and the requirement that they represent each of the five original factors. The content of the original battery was validated using the expert opinions of professionals. The construct validity of the battery was supported by the results from a the logistic regression analysis. However, correlations between the total scores on the shortened SOS Battery and the scores on the self-rated and proxy-rated Reintegration to Normal Living Index were low. Criterion validity was established using the GDS as a criterion. Subjects in GDS 1 and 2, 3 and 4, and 5 were matched with respect to age, education and sex. On 5 of the 13 subtests of the full battery and on the total scores of the shortened battery there was a significant difference in the performance of these three groups.

4.0 DISCUSSION

The findings are now discussed with respect to the existing data, their clinical relevance and theoretical implications. Development of the shortened SOS Battery was based on limited data and so the results pertaining to it should be interpreted with caution. The limitations of the study are also addressed in this section. Finally, the specific original contributions of the study are discussed.

4.1 Psychometric properties of the SOS Battery

The content of the SOS Battery was based on the results from an earlier study (Liu et al., 1991a) which examined tests commonly used in the clinic and in research to measure the concept of spatial orientation. Most of these tests had not been administered to AD patients who are recognized by clinicians, caregivers and researchers as individuals at risk for spatial disorientation sometime during the course of the disease. Thus, the preliminary study examined which tests could distinguish between early AD and control subjects. Most of the tests studied could distinguish between the two types of subjects, however, psychometric data were not available for many of these tests. The present study was able to provide normative data on the tests as well as information on their reliability, internal consistency and validity. The analyses focused on individual subtests rather than on a total battery score because it was premature to recommend the entire battery for clinical or research purposes. These findings are now discussed with respect to the entire battery.

4.1.1 Normative Data

The mean scores obtained by the control group reached near maximum values for 8 of the 13 subtests thereby limiting the usefulness of these subtests for normal control subjects. The scores on these eight subtests also had large values for skewness. The ranges were larger for scores on the Figure-Ground Perception, Map-Reading, Road-Map, Corsi Block and Stylus Maze tests. Thus, these five tests are useful for identifying individual differences in spatial skills related to normal aging. Most of the subtests lack normative data for adults aged 50 to 80 years. The modifications made to the subtests or scoring methods further limit the comparability of the results obtained in this study with those already published. For the Figure-Ground Perception Test, the normal subjects in this study performed similarly to a group of 124 female subjects (mean: 38.2 as compared to 35.6 in the present study) studied by Petersen et al. (1985), even though those subjects were considerably younger (mean age = 34.9 years, range = 18 to 81 years), and a time limit of one minute was implemented for each plate.

Normative data also exist for the performance of adults on the Position in Space Test. Mahoney and Siev (in Siev et al, 1986) reported that a group of normal adults aged 20 to 49 years scored perfectly on the first two parts of this test which were also used in this study. The control subjects in the present study also obtained near perfect scores, thus, it may be assumed that healthy individuals aged 50 to 80 years can perform this test without making any errors. Scores on the Left-Right Discrimination and Spatial Relations tests cannot be compared to other data because the tests were modified from their original formats.

For the Porteus Maze Test, norms exist for children aged 5 to 15 years, but they do not exist for adults. However, the test ages range from 5 to 17 and the range of performance was almost within that specified by Porteus (1959) who recommended that adults start at test age 11 and progressively increase to the maximum test age of 17. The control subjects performed between test ages 10 and 17 years. The mean test score was close to the maximum and was likely an over-estimate because qualitative errors were not recorded. Nevertheless, the scores were similar to those obtained in the preliminary study (Liu et al., 1991a).

Data for the Map-Reading Test cannot be compared to other data because the scoring system used in this study, is not comparable to those used in other studies. When compared to the control group in the present study, the normal control group of the preliminary study performed similarly (mean = 34 as compared to 26 in this study) because the mean score obtained in the previous study is within one standard deviation of the mean obtained in this study. Normative data do not exist for the Road-Map or Corsi Block Tapping tests, but control groups of both the current and preliminary studies performed similarly. Scores on the Stylus Maze Test and the two sketch-map tests were transformed and thus, cannot be compared to those obtained in the preliminary study.

Based on the results from the correlational analyses, four subtests were found to be sensitive to normal aging: Figure-Ground Perception, Corsi Block, Stylus Maze and sketch map of the new environment. Correlational analyses also showed that there was an association between the 3MS and four subtests compared to an association with only one subtest for the MMSE for the control group. This could be attributed to a wider range of scores for the 3MS compared to the MMSE.

In summary, mean scores of normal subjects on 5 of the 13 subtests have wide ranges and are not close to the maximum scores. These subtests are useful for normal subjects aged 50 to

80 years. On the majority of subtests normal subjects have scores close to the maximum and there is little variability in their performance which makes them less useful for normal control subjects, although they are appropriate for differentiating control from AD subjects.

4.1.2 Test-retest and inter-rater reliability

In general, the subtests demonstrated better reliability for the AD group than for the control group. The small range of scores obtained by the control subjects on many of the tests no doubt contributed to the low ICC values. For the AD group, subtests for which statistically significant ICC values were obtained and that met the .70 criterion for both test-retest and inter-rater reliability were the Position in Space, Porteus Maze, and Map-Reading tests, as well as the sketch-map of familiar environment and both self- and proxy-rated Functional Spatial Abilities Questionnaires. For the control group, three subtests met these criteria for both test-retest and inter-rater reliability: the Figure-Ground Perception Test, the Map-Reading Test and the self-rated Functional Spatial Abilities Questionnaire. The high ICC values obtained for the Figure-Ground Perception test in control subjects may be attributed in part to the large number of items on this particular test.

4.1.3 Internal consistency of the full battery

Examination of the subtest-to-subtest correlation coefficients among perceptual and cognitive subtests, revealed that correlations reaching statistical significance for the AD group did not for the control group and vice versa. These differences probably reflect the differences in distributions of the scores between the two groups. As presented in the Results section, for both AD and control groups approximately half of the correlations between perceptual and cognitive subtests were significant and there was a lack of high correlations between subtests within each of the two categories. These results suggest that categorizing the subtests as perceptual or cognitive may not be appropriate. Subtests were categorized as perceptual based on the intent of the original author to use the tests to measure perceptual skills (Ayres, 1972). Subtests in the cognitive category seem to require more complex cognitive abilities than the perceptual subtests. Correlational analysis did show that the performance of AD and control subjects on the Functional Spatial Abilities Questionnaires was clearly not related to scores on the other subtests. Two reasons may account for the difference between the questionnaires' scores and those on the other subtests. First, the format of the questionnaires may sufficiently differ from that of the other subtests and the lack of correlation may reflect the way the subtests were administered. Alternatively, the questionnaires may truly measure functional spatial ability which is not measured by the other subtests in which case it would appropriate to categorize functional spatial abilities as a distinct spatial ability.

With respect to subtest-to-total score Pearson's correlation coefficients, these results need to interpreted with caution. Because the Pearson's correlation coefficient represents a linear relationship that is dependent on the ranges of the two scores being correlated, small ranges in scores can produce coefficients that were not statistically significant or small. For the control group, three subtests did not meet the criterion for internal consistency ($r = .20$) of the full battery. This could be attributed to the small ranges in the scores of these tests.

Cronbach's coefficient alphas were used as another measure of internal consistency. For both the full and shortened versions of the battery, the alpha values were lower for the control group in comparison to the AD group. Again, this could be attributed to the near perfect scores and extremely low variability of the control subjects' scores because Cronbach's alpha is partly dependent on the correlations between subtests. The magnitude of Cronbach's alpha is also partly dependent on the number of subtests entered into the calculation (Carmines & Zeller, 1979; Golden et al., 1984). Therefore, the high overall alpha obtained for the AD group for the full battery may be partly due to the large number of subtests.

Factor analysis of the full battery was used in this study to describe further the internal consistency of the battery. Results of this analysis using data from the control group indicated that the 13 subtests can be subdivided into 5 categories of spatial skills. Scores on each of the subtests, except for the Figure-Ground Perception Test, loaded on only one of the five factors. When a subtest loads on more than one factor, this may be an indication that the subtest is not measuring what the battery is intended to measure and thus, should be altered or discarded (Streiner & Norman, 1989). Although the scores on the Figure-Ground Perception Test correlated with scores on only one other subtest, the Road-Map Test, scores on the Figure-Ground Perception Test were highly correlated with the total score and Cronbach's alpha was low. These results suggest that for the control group, although the Figure-Ground Perception Test is associated with two of the five factors of spatial abilities, it is still internally consistent with the full battery.

It is proposed that the five factors describe the following types of spatial abilities. Factor 1 describes spatial representation on a difficult level requiring mental rotation. Factor 2 describes spatial learning and problem solving. Factor 3 describes functional spatial abilities. Factor 4 is associated with simple concepts of left/right and up/down. Factor 5 is related to immediate spatial memory.

4.1.4 Construct and criterion validity of the full battery

Construct validity was evaluated using multiple logistic regression analysis. At the very least, scores on the subtests of the SOS Battery should be able to be used to discriminate between AD and control subjects after controlling for gender, age and years of education. High scores on all subtests were associated with a decreased likelihood of having AD. Criterion validity was studied with the realization that there is currently no gold standard for measuring spatial disorientation in AD patients. The GDS (Reisberg et al., 1982) served as a criterion based on the premise that a patient's spatial orientation skills deteriorate with the progression of AD, and that the GDS accurately measures the disease progression. According to the GDS, spatial disorientation begins in novel environments for patients in stages 3 and 4 and occurs in familiar environments for patients in stages 5 and 6. In stage 7, wandering behaviour is observed. Of the 13 subtests, 5 were sensitive to the severity of AD. These results provided evidence for criterion validity for the Figure-Ground Perception, Road-Map, sketch map of the new and familiar environments and the proxy-rated Functional Spatial Abilities Questionnaire. This analysis was limited in that the sample sizes (10) were small given the large variability of the GDS 5 group's scores. The variability in scores increased with the severity of AD and could account for the lack of statistically significant differences between the GDS 3 and 4 and GDS 5 groups for 8 of the 13 subtests despite the substantially lower means for the GDS 5 group. A longitudinal study, preferably with larger samples, would be required to confirm whether or not the SOS Battery is sensitive to the progression of AD.

4.2 Characteristics of the Shortened SOS Battery

The preliminary shortened SOS Battery was developed because the full battery was very time-consuming for clinical use and would be burdensome to clients. As presented previously, the selection of subtests for the shortened battery was based on several criteria. For the control group two subtests did not meet the criterion for adequate subtest-to-total score correlation in the shortened battery. Again, this could be attributed to the small ranges in the scores of these tests. Although the overall Cronbach's alpha value of the shortened battery is near the criterion (alpha = .78), its decline from that of the full battery probably reflects the smaller number of subtests in the shortened battery. The overall alpha indicates that the shortened battery is not internally consistent when administered to control subjects.

It could be argued that the shortened battery contained one variable more than permitted for a factor analysis to be conducted using data from only 25 AD subjects, or fewer given that there were missing data. All six subtests were used because the intention was to explore how a representative sample of subtests would behave in a factor analysis using data from the AD group. The proxy-rated Functional Spatial Abilities Questionnaire was the only subtest that permitted an examination of the subject's performance in daily activities that are related to spatial orientation. The proxy-rated version was selected because, for AD subjects it was more internally consistent with the full battery than the self-rated version. This selection was also based on the clinical consideration that in the late stages of AD, patients are unable to complete such a questionnaire thereby making the self-rated version less useful. In retrospect, the selection of the proxy-rated version could also be supported by its criterion validity which was established using the GDS. In contrast, the self-rated Functional Spatial Abilities Questionnaire was not sensitive to the severity of AD. Finally, missing data were not extensive. One subtest had data from 25 subjects, three subtests had data from 24 subjects and two had data from 23 subjects. Thus, the number of missing data did not exceed two.

With respect to the AD group, the results can only be compared to those of the control group because it was not possible to conduct a factor analysis using the full battery. Further, these results are interpreted with the realization that the sample size was still inadequate to perform a factor analysis with confidence even when the number of subtests was reduced. Two major differences emerged from these analyses. For the AD subjects, scores on the Map-Reading Test were associated with scores on the Functional Spatial Abilities Questionnaire and scores on the sketch map of the familiar environment were associated with two factors instead of one. This implies that an impairment in functional spatial skills can be associated with performance on the Map-Reading Test which is administered in the laboratory. Also, for AD patients, the ability to sketch a map of one's familiar environment was associated with sketching a new map from memory. In contrast, scores on these two sketch maps were not related for the control group. Results from the control group's data indicate that sketch map of the familiar environment may not be related to the construct being tested by the battery because it loaded on two factors and should therefore be discarded. But results from the AD group's data show that the sketch map of the familiar environment correlated highly with the total score (r = .61) and it was associated with a relatively low Cronbach's alpha (alpha = .72) thereby demonstrating consistency with the shortened battery. Scores on the sketch map of the familiar environment load with scores on the sketch map of the new environment probably because AD results in a deficit in visuospatial constructional ability which is required to perform both tests.

Construct validity of the shortened battery has been partially addressed in the discussion of the full battery because the six subtests were included in the multiple logistic regression analyses. On all six subtests, higher scores were associated with a lower likelihood of being an AD subject.

Further analysis correlating the Reintegration to Normal Living Index scores with the shortened battery total scores did not provide evidence for construct validity. The use of the Reintegration to Normal Living Index was based on the rationale that a person who is spatially disoriented should experience a poorer ability to integrate into normal living. Also, it was suggested that performance on the SOS Battery could reflect an individual's quality of life. The results suggest that the subtests of the shortened battery measured cognitive abilities which are quite different from the construct of quality of life which was what the Reintegration to Normal Living Index measured. Further work to address the issue of construct validity could correlate scores on the SOS Battery with scores on other cognitive tests. One battery of cognitive tests that is of interest is the Hierarchic Dementia Scale (Cole & Dastoor, 1987) which consists of 20 subscales that examine cognitive functions which include orientation, concentration, drawing and recent memory. In addition, the Functional Spatial Abilities Questionnaire could be correlated with other functional scales in order to examine construct validity.

Criterion validity of the shortened battery was also addressed in the context of the full battery. Five of the six subtests were sensitive to the severity of AD. Scores on the Map-Reading Test did not differentiate between the GDS 3 and 4 and GDS 5 group, possibly due to the large variability of the scores from the GDS 3 and 4 group. Despite this, differences between the three groups of subjects in the standardized total scores for the shortened battery were statistically significant indicating that the shortened battery is valid when using the GDS as a criterion. In summary, the subtests of the shortened battery are valid to the extent that they can discriminate between early (GDS 3 and 4) AD and control (GDS 1 and 2) subjects and they are sensitive to the severity of AD.

4.3 Variables that may be related to spatial skills

In addition to cognitive abilities, other variables may be related to spatial orientation. These variables include gender, handedness, psychomotor speed, attention, language ability and education. Where possible, an attempt was made to control for the effects of these variables in the design of the study.

The relationship of gender and handedness to spatial abilities has been extensively studied (Buffery & Gray, 1972; De Renzi, 1982; Maccoby & Jacklin, 1974). The effect of gender on spatial skills was not an issue in this study because the early AD and control groups were equivalent with respect to the proportion of men and women. Gender was controlled for in the logistic regression and also in the study of criterion validity by matching the GDS 1 and 2, GDS 3 and 4 AD and GDS 5 AD subjects on this variable. The effect of handedness was controlled for by including only subjects who were right-handed.

Psychomotor speed, attention, language comprehension and years of education were other factors that were considered. As described in the Results section, the AD patients took on average 16 minutes longer than the control subjects to complete the entire battery. This number underestimates the actual time taken because many AD patients were unable to complete certain subtests, as reflected in the smaller sample sizes on some test scores. Prior to commencing the study, it was considered a possibility that AD patients would require more time to complete the battery. It was therefore considered inappropriate to set a time limit on each subtest as this would have introduced a speed factor, and may not have permitted the AD patients to perform at their optimum. A review of the medical charts revealed that none of the AD patients had physical impairments or motor disorders that could have accounted for the longer completion time. AD does not usually involve the primary sensory and motor areas of the brain (Van Hoesen & Damasio, 1987). However, as mentioned in the Introduction, AD patients have been shown to have impaired visual selective attention on tasks requiring effortful processing as indicated by a slower speed of performance with an increasing number of visual distractors (Foster et al., 1992). Hence slower performance can be attributed to a slowing of cognitive processes, rather than to a basic sensory or motor deficit.

Attentional deficits would definitely affect performance on the SOS Battery. Attention was measured using Albert's Test and the Bells Test. Vanier et al. (1990) have suggested a criterion of four or more errors on the Bells Test and a criterion of two or more on the Albert's Test for diagnosing hemi-neglect. There was no evidence of unilateral neglect in the AD group as shown by the location of errors on both tests. Nevertheless, the average on the Bells Test for the AD group did meet this criterion while the average score on Albert's Test did not. These data suggest that the Bells Test was capable of detecting a deficit in attention in AD patients although this attentional deficit was not due to hemi-neglect. The Bells Test was more sensitive to AD possibly because the test contained distractors whereas Albert's Test did not and therefore was easier. These results are consistent with those of Foldi et al. (1992) who used nine cancellation tests that contained line drawings of geometric shapes instead of objects. Although attention was not tested using a working memory model, that is, subjects were not given combined tasks to perform, the attentional deficit detected on the Bells Test could be attributed to an impaired central executive system (CES) as proposed by Morris and Baddeley (1988). As described in the Introduction, the CES has a limited capacity and is responsible for the maintenance, rehearsal and storage of information. In this study, all of these abilities were required for understanding and remembering instructions as well as for the performance of the subtests. Thus, impaired performance of the AD group on the SOS Battery could be partly related to impaired performance on the Bells Test.

In the Introduction, evidence was presented to suggest that although AD patients can present with language and/or visuospatial deficits, these deficits are probably independent of each other. The results from the Boston Naming Test are consistent with other findings that report deterioration of confrontation naming in AD patients. This impairment may have affected performance on the sketch maps which required that subjects name each of the rooms. Some of the AD patients misspelled the names but they were not penalized in the scoring process. Errors made on the identification or location of rooms may have been a result of an impairment in naming. For example, if a subject could not name a room, he or she may have chosen not to include it in the sketch map or he or she may have chosen to assign another name to the room.

Although language comprehension was not directly measured, it was assumed that AD patients would experience difficulty comprehending and remembering instructions, especially if the instructions were long. Verbal comprehension has been shown to be affected by AD (Murdock et al., 1987). This study attempted to control for this in three ways. First, only subjects who were fluent in written and spoken English were recruited, although English was not necessarily a subject's mother tongue; second, other investigators' test instructions were made shorter in order to facilitate remembering them; and third, all instructions were followed by a confirmation from the patient to indicate that the instructions were understood. Each examiner was also informed to repeat any portion of the instructions as frequently as needed throughout the testing session. This also served to control for the effects of poor memory or poor concentration in following instructions. As reading comprehension has also been shown to deteriorate with the progression of AD, performance on the self-rated Functional Spatial Abilities Questionnaire could have been affected by impaired reading comprehension.

Education and training may have an effect on spatial abilities and should be considered when assessing patients with dementia (Kittner et al., 1986). Appleyard (1970) reported that individuals with higher levels of education tend to represent their environment in a more spatially coherent manner and infer more from limited experience. In this study, the AD and control groups were comparable with respect to the mean number of years of education. Education was controlled for in the logistic regression. For the criterion validity study, the GDS 3 and 4 AD and GDS 1 and 2 subjects were matched to the late AD group with respect to years of education. The lack of a significant correlation between education and scores on the subtests of the shortened SOS Battery for both the AD and control groups is further evidence that spatial performance was not influenced by education in this study.

Given that the above variables were considered in the inclusion criteria for the subjects and in the analyses, it was not likely that gender, handedness, psychomotor speed, language or years of education affected the performance of subjects on the SOS Battery. Other variables that should be taken into consideration in future studies include lifestyle differences, active versus passive travel modes, and cultural values (Moore, 1979). These factors would be especially important in studies of functional spatial orientation. For example, Appleyard (1970) has reported that individuals who drive automobiles (active mode) are better oriented than those who travel only by public transportation (passive mode).

4.4 Theoretical Implications

The results of this study have implications for the interpretation of other findings on spatial abilities, as well as for the development of a framework for studying spatial orientation in AD.

4.4.1 Spatial skills in Alzheimer's disease

In the Introduction, spatial abilities in AD were reviewed in the context of figure-ground discrimination, personal orientation, extrapersonal orientation, visuospatial constructional ability, mental representation and functional spatial orientation pertaining to the capacity to drive and way-finding ability. These topics will be discussed with respect to the results of this study.

4.4.1.1 Figure-ground discrimination

The AD group in this study was significantly impaired in discriminating figure from ground as tested on the Figure-Ground Perception Test. The mean scores obtained by the AD and control groups were similar to those obtained in a previous study using similar, but smaller, groups of subjects (Liu et al., 1991a). Impairment in figure-ground discrimination in AD patients has also been reported by Mendez et al. (1990) who used three items from the Figure-Ground

Perception Test, and by Loring and Largen (1985) who used the Embedded Figures Test which has been validated with the Figure-Ground Perception Test (Petersen & Wikoff, 1983).

Scores on the Figure-Ground Perception Test were normally distributed for the control group, had a wide range and the mean score was not near the maximum. For the control group, the subtest also met the criterion for good test-retest and inter-rater reliability. Thus, it was an appropriate test to use with normal control subjects. In addition, it was the only perceptual subtest that was sensitive to the severity of AD which was consistent with results from other studies (Gauthier et al., 1990, Mendez et al., 1990). However, despite its high test-retest reliability, it had low inter-rater reliability when administered to the AD group. Whether or not an AD subject completed the test seemed to depend on the ability of the examiner to encourage the patient to perform. Although 22 of the 25 AD subjects were able to complete the test, many of the errors were the result of not providing a response and not a result of providing the wrong response. Thus, this test could be made more appropriate for AD subjects if the examiners were trained to be persistent in obtaining a response for each item.

Results from the first factor analysis indicated that the ability to discriminate figure from ground was related to performance on the other cognitive subtests. It was the only subtest whose scores loaded equally on two factors that were described earlier as associated primarily with spatial representation requiring mental rotation and spatial recognition (Factor 1) and spatial learning and problem solving (Factor 2). It was not associated with two of the other three perceptual subtests (Left-Right Discrimination and Position in Space). Correlational analyses showed that Figure-Ground Perception Test scores of the control group were significantly and adequately correlated ($r \geq .20$) with five of the seven cognitive subtests. Similarly, scores of the

AD group were significantly and adequately correlated with two of the cognitive subtests. These results suggest that either figure-ground perception involves more than just visual perceptual skills as proposed by Cohn et al. (1991) and Ayres (1972) or the other cognitive tests also require similar perceptual skills. It is more likely that the Figure-Ground Perception Test requires cognitive spatial skills because it involves effort on the part of the subject to suppress two other targets in addition to the distractors in order to search and identify a third target. A memory component is also involved in remembering which target has been identified so that time is not spent on perseverating over one target at the expense of finding the other two targets. In addition, the subject is also required to plan in order to systematically eliminate choices that are not appropriate. Attention, memory and planning are skills previously associated with performance on cognitive tests and could explain why the scores of the Figure-Ground Perception Test were not highly correlated with the other perceptual subtests.

4.4.1.2 <u>Personal orientation</u>

With respect to personal orientation, the data confirmed Cogan's (1985) observation that patients with early AD experience deficits in left-right discrimination. In a previous study (Liu et al., 1991a), an impairment on the Left-Right Discrimination Test was not seen in AD patients on an initial evaluation, although these same patients did demonstrate a significant deterioration when retested one year later (Gauthier et al., 1990). The impairment observed in the present study could be attributed to an increased sensitivity of the test which was modified to include more items pertaining to identification of the examiner's body as opposed to the subject's body. This would be consistent with the results of Fischer et al. (1990) who compared the performance of mild (MMSE: 16 to 23), moderate (MMSE: 6 to 15) and severe AD (MMSE: < 6) patients on items requiring the identification of the left and right side of the subjects' own body as well as of a confronting doll. These investigators found that AD patients at every stage were impaired on items concerning the doll, but only severe AD patients were impaired on items pertaining to their own body. Although scores from the Left-Right Discrimination Test were not analyzed with respect to whether the items involved the subject's or the examiner's body, based on the findings of Fischer et al. (1990), it is likely that the impairment observed in this study was due to difficulty with items pertaining to the examiner's body. The low test-retest and inter-rater reliability of this test with the AD group suggest that it requires further examination and revision before it can be used with confidence with the AD population.

Another test commonly used to assess personal orientation is the Road-Map Test. Results of the factor analyses with the full and shortened batteries showed that the scores on the Left-Right Discrimination Test were not associated with scores on the Road-Map Test. Although both tests require personal orientation, that is, left-right orientation with respect to one's own body, the Road-Map Test is more difficult. Mental rotation is involved in the Left-Right Discrimination Test when the subject has to identify which of the examiner's hands is left or right. But the mental rotation involved in the Road-Map Test is more difficult since the subject has to imagine him- or herself rotating the entire body to face several directions. In addition, the Road-Map Test involves more than three times the number of items of the Left-Right Discrimination Test. This may have introduced a fatigue component in addition to the factor of increased complexity.

4.4.1.3 Extrapersonal orientation

The mean Map-Reading Test scores obtained by the AD and control groups in this study were considerably lower than those obtained by comparable groups of subjects in a previous study (Liu et al., 1991a). This difference may be due to the conditions under which the test was performed. Bylsma et al. (1992) reported that patients with Huntington's disease were impaired on the Map-Reading Test when using a TURN condition, which required subjects to turn and face the direction in which they were walking, but they were not impaired in a NO-TURN condition, which required that the subjects face a predetermined north wall throughout the entire test. These investigators argued that the NO-TURN condition measures extrapersonal orientation, while the TURN condition requires mental rotation in addition to extrapersonal orientation, and is more difficult to perform (Bylsma et al., 1992). In the study of Liu et al. (1991a), it was not specified to the subjects which condition they were to use, and this could have resulted in higher scores if many subjects used the easier NO-TURN condition. In this study, the subjects were instructed to use only the TURN condition. Thus, the lower scores may reflect the more difficult nature of the task that combined mental rotation with extrapersonal orientation.

Results from the first factor analysis, which used data from the control group, indicated that scores on the Map-Reading Test were related to Figure-Ground Discrimination, learning as assessed on the Stylus Maze Test, and mental representation as assessed on the sketch map of the familiar environment. In the second factor analysis, for the control group, the Map-Reading Test was related to three of the other five subtests, and for the AD group, it was related to scores on the Functional Spatial Abilities Questionnaire. These results indicate that the Map-Reading Test and the proxy-rated Functional Spatial Abilities Questionnaire involve similar abilities which are affected by AD. The Map-Reading Test is the only subtest that requires movement of the whole body. This dynamic component is an integral part of functional spatial orientation which the Functional Spatial Abilities Questionnaire was assumed to measure.

It has been proposed that successful performance on the Map-Reading Test depends on an intact working memory which allows the concurrent use of several cognitive functions (Aubrey & Dobbs, 1989). In view of the many cognitive skills required to perform the test, the impairment observed in the AD group on this test may be attributed to a deficiency in working memory. It would be of interest for a future study to examine extrapersonal orientation separately from mental rotation by comparing the TURN and NO-TURN conditions, as was done with Huntington's disease patients by Bylsma et al. (1992). If AD patients are not impaired in the NO-TURN condition, then it can be argued that the impairments observed in this study were due to mental rotation deficits and not an impairment of extrapersonal orientation. At present, these results can only suggest that the difficulties AD patients have with the Map-Reading Test can be attributed to deficits in extrapersonal orientation, mental rotation and working memory.

4.4.1.4 Visuospatial constructional ability

The AD subjects were impaired in visuospatial constructional skills as evaluated on the clock drawing test. Impairments on this task have been reported by several investigators studying AD subjects (Doyon et al., 1991; Mohr et al., 1991; Rouleau et al., 1992; Sunderland et al., 1989; Wolf-Klein et al., 1989). Clock drawing involves the ability to organize spatially the numbers (Rouleau et al., 1992; Shulman et al., 1986). This spatial organization was also required on the mental status exams, which involved copying two intersecting pentagons, and on the sketch map tests. Thus, the impairments seen on the copying task and on the sketch maps may be attributed partly to the spatial organization aspect of visuospatial constructional ability. Further examination using qualitative analyses of the clock drawing results could provide more insight into the specific types of errors made. Although the analysis used in this study incorporated some qualitative aspects (ie. spacing of numbers) this was not evident in the final scores.

The method of administering or scoring the test is also subject to further investigation given the low test-retest and inter-rater reliability coefficients for both the AD and control groups. Since the conclusion of this study, a new method for administering and scoring a clock drawing task has been reported (Mendez, Ala & Underwood, 1992). This method requires subjects to draw a clock in contrast to the method used in the current study which provided the subjects with the outline of a clock. Another difference was that Mendez et al. (1992) presented written as well as oral instructions. While Mendez et al. (1992) also used a maximum score of 20, the points were allocated differently. Using this method with 42 AD patients, Mendez et al. (1992) reported high test-retest reliability coefficients measured over three different periods of time (r = at least .70, $p < .001$). Inter-rater reliability coefficients were also reported to be high ($r = .94$, $P < .001$). It would be interest to compare the two methods of scoring in another study using the same patients and statistical analysis.

4.4.1.5 Mental Representation

Both sketch map tasks, chosen to assess mental representation, involved visuospatial constructional ability and produced floor plans as end products. Despite this, results of the factor analyses indicated that the processes involved were clearly different. For the AD group, factor analysis showed that performance on the sketch map of the familiar environment was related to that on the Position in Space and Road Map tests (Factor 1) and also to sketch map of the new environment, whereas for the same factor analysis involving control group data, performance on this task was related to that on the Position in Space and Road-Map tests in addition to that on the Map-Reading Test. The Position in Space Test requires the ability to distinguish left from right and up from down. These abilities could be related to the ability to place the rooms in correct relationships to each other. The Road-Map Test requires mental rotation of one's own body which may be needed in order to place the rooms in the correct relationship with respect to the point of reference the subject has chosen. For example, if a subject has chosen to place the front entrance of the floor plan at the bottom of the page, all of the other rooms would have to be placed with respect to that reference point. The results of the two groups differed in that scores on the sketch map of the familiar environment loaded with scores on the Map-Reading Test for the control group, whereas scores on the sketch map of the familiar environment loaded with scores on the sketch map of the new environment. These results suggest that for the normal control subjects, the ability to sketch a floor plan of one's home is associated with extra-personal orientation skills. It is proposed that this association does not exist in the AD group in which the ability to sketch a floor plan of one's home is related to visuospatial constructional abilities which were also required for the sketch map of the new environment, and which have been shown to be impaired in AD.

With respect to results of the factor analysis involving the 13 subtests (control group data), performance on the sketch map of the familiar environment was related to that on the Figure-Ground Perception, Map-Reading and Stylus Maze tests, whereas performance on the sketch map of the new environment was related to that on the Corsi Block Test. These results were explained earlier by associating sketch map of the familiar environment with requiring mental rotation, planning and problem solving and associating sketch map of the new environment with immediate spatial memory.

In the preliminary study, AD patients were impaired on sketch map tasks (Liu et al., 1991a). However, a different method of scoring was used in that study which involved analyzing two scores, one for frequency and another for accuracy. In this study one score was calculated based on the mean of these measures. The use of a mean value meant that the score reflected both the frequency and the accuracy components of the task. Another difference was that instead of conducting a home visit in order to score the sketch maps, they were scored by comparing them to sketch maps drawn by caregivers. Given that the two scoring methods were different, it was not possible to compare the scores between the two studies for the normal control subjects. Thus, it is unknown whether using a caregiver's sketch map with which to score a subject's sketch map was comparable to conducting a home visit.

The sketch map of a familiar environment met the criteria for high test-retest and inter-rater reliability in AD patients. This further supports the use of this task to study mental representation in AD patients as it relates to spatial orientation. More sensitive measures of visuospatial constructional performance and mental representation could be applied to the maps drawn by subjects. For example, they could be analyzed using qualitative measures based on the overall layout of elements on the maps (Moeser, 1988). The task could also be expanded to involve mental representation of a macrospatial context, that is, sketch a map of one's neighbourhood or city (Rovine & Weisman, 1989). It would also be of interest to relate sketch maps drawn by AD patients to their way-finding abilities in a new environment such as a hospital, nursing home or a new city, as is commonly done with normal subjects (Moeser, 1988; Rovine & Weisman, 1989).

4.4.1.6 Functional spatial orientation skills

As expected, AD patients were impaired in functional spatial orientation as assessed by the Functional Spatial Abilities Questionnaire. The content of the Functional Spatial Abilities Questionnaire was based on responses from caregivers of AD patients to an open-ended questionnaire. This questionnaire asked caregivers to describe behaviours or actions that they had seen in AD subjects who had difficulty with orientation to place. The 12 items on the Functional Spatial Abilities Questionnaire pertained to one's function in both new and familiar environments. The items addressed the perception of one's own sense of direction and activities of daily living (ADL) including driving, way-finding and following maps.

Results from the item analysis indicate that 11 of the 12 items were sensitive to differences between AD and control subjects. The only exception was item 9 which states, "I require supervision travelling in the neighbourhood". This item was sensitive to group differences when it was rated by the proxies but not when it was rated by the subjects themselves. A discrepancy was also observed in the correlational analyses which showed that self- and proxy-ratings were highly correlated ($r = .68$) for the control group but not for the AD group. Thus, although the two questionnaires are highly associated with each other as shown in the factor analysis results (using the full battery), users of the questionnaires are alerted to the discrepancy between AD subjects' responses and the responses of their caregivers.

Clinicians and caregivers are often faced with deciding when an AD patient should no longer be permitted to drive. Two surveys suggest that there is a higher incidence of driving accidents associated with AD (Friedland et al., 1988; Lucas-Blaustein et al., 1988). A necessary component of driving is adequate way-finding ability. Way-finding has been described as a complex ability involving not only intact cognitive abilities but also an ability to interact with and integrate environmental information (Downs & Stea, 1981; Passini, 1984a; Rovine & Weisman, 1989).

Functional spatial orientation is a construct separate from spatial orientation skills that are evaluated in the laboratory. Results from the first factor analysis using the control group data showed that both self- and proxy-rated Functional Spatial Abilities Questionnaires loaded on a single factor. The Functional Spatial Abilities Questionnaire appeared to be valid for evaluating functional spatial orientation because it detected a deficit in AD patients when compared to the control subjects. The self-rated Functional Spatial Abilities Questionnaire was excluded from the shortened SOS Battery in order to decrease the number of subtests and because it was assumed that the proxy-ratings were more accurate. It may be argued that self-ratings were more accurate than proxy-ratings and that the proxies under-estimated the functional performance of the AD patients. Although the AD patients rated themselves significantly higher than their proxies, both self- and proxy-rated scores were lower than those for the control group. In this study, AD patients also rated themselves higher than their proxies on the Reintegration to Normal Living Index. McGlynn and Kaszniak (1991) reported similar findings when they compared AD patients' reports with those of their caregivers on another questionnaire. They found that AD patients tended to under-estimate their difficulties with cognitive tasks in everyday life and over-estimate their memory abilities.

In another study, Magaziner, Simonsick, Kashner and Hebel (1988) compared elderly hip fracture patients and their proxies' responses to questions pertaining to the patients' pre-fracture health and functional status. These investigators reported that the best agreement occurred where the patient had no cognitive impairments and was not depressed. Patients who were cognitively impaired, as measured on the Mental Status Questionnaire of Kahn et al. (1960), rated their health and functional status better than the proxies. However, Rocca et al. (1986) found that when questions pertaining to family or medical history were used, the responses of AD patients agreed with those of their proxies. Based on these studies, it may be hypothesized that AD patients over-estimate their cognitive and functional abilities because of a denial of the disease or in reaction to social pressures to appear competent. Questions pertaining to topics such as family history pose less threat to one's self-esteem, hence there was an agreement between the self and proxy-ratings in the study of Rocca et al. (1986). However, it is also likely that AD patients are more accurate in responding to questions pertaining to facts in contrast to subjective appraisals of recent changes in their cognitive or functional abilities. Although for the AD group both self- and proxy-rated Functional Spatial Abilities Questionnaire responses demonstrated high test-retest and inter-rater reliability, it appears that when there is a discrepancy, proxy-rated scores are lower.

4.4.1.7 Summary

In general, the spatial deficits observed in the AD patients in this study are congruent with reports by other investigators. However, this study differed from all other studies in that several spatial tests ranging from those that assess perception to those that evaluate cognitive skills were administered concurrently. In addition, the Functional Spatial Abilities Questionnaire was introduced as a measure of functional spatial orientation. This permitted analyses of different aspects of spatial orientation in order to understand better the processes involved in spatial disorientation.

The results from this study can help with the interpretation of findings from other studies reported in the literature. One could deduce from the results of the Figure-Ground Perception Test that training of the examiner is crucial to its reliability between raters when used with AD patients. It has also been seen that performance on the Figure-Ground Perception Test is related to performance on many cognitive spatial tests. This study confirms the results from other studies which indicate that the tests used for assessing personal and extrapersonal orientation evaluate other spatial skills as well. Both the Road-Map and the Map-Reading Tests, used to assess personal and extrapersonal orientation, respectively, involve mental rotation.

Visuospatial constructional ability was also impaired in the AD group as evaluated on the clock drawing test. The impairments seen on the sketch map tasks may therefore be partially attributed to difficulties with visuospatial construction. However, the sketch map of the new environment was primarily a memory test while the sketch map of the familiar environment involved spatial orientation skills. No other study has examined the performance of AD patients on a sketch map task for a familiar environment as was done in the present study. But studies have confirmed a deficit in immediate spatial memory in people with AD (Corkin, 1982).

Deficits in functional spatial orientation were detected using a questionnaire that was based on information provided by caregivers. As yet there has been no report in the literature of an evaluation for examining functional spatial orientation in AD patients. Skills assessed on the Functional Spatial Abilities Questionnaire overlapped with those evaluated on the Map-Reading Test. But functional spatial orientation was associated with a factor that was distinct from the factors representing the other spatial skills evaluated in the laboratory.

Other spatial tests used in the study that have not been discussed are the Spatial Relations, Porteus Maze and the Stylus Maze tests. It is not known why the Spatial Relations Test was unreliable for the AD group. The test was modified so that only 10 of the 60 original test items were used. A previous study (Liu et al., 1991a) used 15 items and found a statistically significant difference between the scores of the AD and control groups. It was felt that a test of 15 items was still too long for AD patients and so it was shortened to 10 items. An assumption was made that they were representative of the range of difficulty of items on the original test. However, no analysis was conducted to see if this subtest was equivalent to the original test. Nevertheless, it does not seem that this alone would have resulted in the low reliability. The Porteus Maze Test had high reliability coefficients, but was influenced by education in both groups. Thus, use of this test to assess problem solving and spatial planning in AD patients should take into account the effects of education.

The reliability coefficients for the Modified Stylus Maze Test were just below the criterion of .70. Although this test was not included in the shortened SOS Battery, it should still be considered as a useful tool for studying spatial learning. Other versions of this test have been found to detect spatial deficits in AD patients in comparison to normal control subjects (Brouwers et al., 1984; Liu et al., 1991a) and in comparison to Huntington's Disease patients (Brouwers et al., 1984). For the control group, performance on the Stylus Maze Test was found to be related to scores on the Map-Reading Test, sketch map of the familiar environment and the Figure-Ground Perception Test. It is not known if this test would have loaded in a similar way if data from a larger group of AD patients had been available.

In short, the number and range of tests used in this study were more comprehensive than those used in other studies examining spatial deficits in AD patients. This study included spatial tests that have been used previously with AD patients as well as other tests that have not been employed with this particular clinical population. The concurrent use of a wide range of spatial tests permitted factor analyses to be used. Results from these analyses can contribute to the development of a framework for studying spatial orientation in AD patients.

4.4.2 Framework for studying spatial orientation in AD patients

This study began with an hierarchical framework of spatial orientation based on a review of the literature. Patients with AD had been shown previously to experience difficulty on a number of different spatial tests. The deficits were reported on tests of figure-ground discrimination (Loring & Largen, 1985; Mendez et al., 1990), personal orientation (Bylsma et al., 1992; Fischer et al., 1990; Flicker et al., 1988), extrapersonal orientation (Liu et al., 1991a; Liu et al., 1991b), visuospatial constructional ability (Henderson et al., 1989; Mohr et al., 1991; Teng & Chui, 1987), and mental representation (Liu et al., 1991a). Impaired functional spatial orientation has been of concern to caregivers and health professionals with respect to an AD patient's capacity to drive and to travel independently (Drachman, 1988; Friedland et al., 1988; Lucas-Blaustein et al., 1988).

The findings reported in the literature and the results from this study suggest that AD patients are impaired on all aspects of spatial orientation. This is consistent with the neuropathology of the disease. As discussed in the Introduction, the pathological changes associated with AD affect all areas of the cortex (Van Hoesen & Damasio, 1987) and subcortical structures (Morris & Kopelman, 1986). Some of the specific deficits detected were with visual perceptual skills (Left-Right Discrimination and Position in Space), planning (Porteus Maze) and immediate memory (Corsi Block). These deficits are consistent with reports of damage in the visual, frontal and temporal (hippocampal regions) cortices (Morris & Kopelman, 1986; Van Hoesen & Damasio, 1987). Computed Tomography (CT) results were available for 18 of the 25 GDS 3 and 4 AD patients and 5 of the 10 GDS 5 AD patients who participated in this study. In these scans, there was no evidence of localized lesions and the areas of the brain affected by atrophy varied. Some of these CT results were normal, others showed that different cortical or subcortical areas were affected. The generalized deterioration in cognitive function that would be associated with this atrophy is seen in the results of the mental status examinations and nonspatial tests such as the Hebb Verbal Memory test and the Boston Naming Test.

This study examined the contribution of various skills to spatial orientation by using a comprehensive battery of spatial tests that evaluated skills believed to be required for adequate spatial orientation. A framework for studying spatial orientation in AD must consider the multi-dimensional and hierarchical aspects of the construct of spatial orientation. It is multi-dimensional in the sense that different spatial skills are involved. The hierarchy refers to the concept that certain spatial skills are prerequisites for other more difficult spatial skills. Concurrent deficits on many types of spatial skills would be consistent with the global cognitive deterioration that is characteristic of AD.

The hierarchical and multi-dimensional aspects of these spatial skills need to be further explored using a larger sample of AD patients which would permit the use of more variables in a factor analysis. With respect to functional spatial orientation, this framework must consider that there are other variables that have an impact on functional performance. These variables may be environmental, physical or behavioural. Studies are needed to identify other factors and their relationships to each other as well as their influence on one's functional spatial ability.

4.5 Clinical Implications

The results of this study have implications for clinical assessments as well as for the clinical management and treatment of spatial disorientation in AD patients.

4.5.1 Clinical assessment of spatial orientation in AD patients

On the basis of this study, four subtests of the shortened SOS Battery can be recommended for the clinical assessment of spatial orientation in AD patients. These are the Position in Space Test, the Map-Reading Test, the Road-Map Test and the sketch map of the familiar environment. The Functional Spatial Abilities Questionnaire has potential for clinical use, but requires additional research to establish its content validity. This study demonstrated that all of these tests have high test-retest and inter-rater reliabilities. After using a screening measure to assess global cognitive function, a clinician could select one of these tests and do further evaluation of spatial orientation. The normative data, which are based on a large sample of control subjects, can be used by a clinician for the purpose of comparison.

To the author's knowledge, the Position in Space Test is used by some clinicians to evaluate perceptual spatial skills and recognition of shapes in patients with dementia. However, there are no documented reports of its use with AD patients. This study used the first two parts of the original test but not the third part which assesses spatial memory. The results showed that the first two parts are reliable and valid for the assessment of perceptual spatial abilities in AD patients.

To date, the Map-Reading and Road-Map tests have been primarily used for research purposes. These tests can be easily administered in the clinic as they are relatively short and inexpensive. A clinician using these tests must consider that mental rotation is involved in both tasks. Further research is needed to determine if AD patients are also impaired on the Map-Reading Test when administered in the NO-TURN condition. For the AD patients, performance on the Map-Reading Test was related to that on the Functional Spatial Abilities Questionnaire. If one assumes that the Functional Spatial Abilities Questionnaire is an accurate and good measure of functional spatial orientation, then it may be argued that the Map-Reading Test is currently the only standardized clinical assessment that is related to functional spatial orientation in AD patients. If this is the case, it may be because the Map-Reading Test requires body movements in space that are related to navigation, unlike most clinical tests that are performed by a patient seated at a table.

The sketch map of the familiar environment task demonstrated excellent psychometric properties. This task has been used by clinicians with other populations to obtain a qualitative impression of a patient's mental representation (McFie et al., 1980; Ross, 1980). There is only one report of the use of a sketch map with an AD patient (Cogan, 1985), and one other study that has quantified AD patients' performance on this task (Liu et al., 1991a). The present study provides data to support its use based on its stable psychometric properties. The sketch map task as applied to the familiar environment can also be applied to unfamiliar environments as was done in the preliminary study.

There is one potential limitation involved with using the sketch map as a clinical tool. In this study, a floor plan drawn by a relative was used as the standard for scoring, which replaced the time-consuming home visits in the previous study (Liu et al., 1991a). The caregiver was also asked to produce a second sketch map at home in order to confirm the accuracy and consistency of the first sketch map. In the majority of cases, the two sketch maps drawn were consistent. In three cases, they were not and so the caregivers were requested to provide a third sketch map which was consistent with either the first or second one. Although this method of scoring the patient's sketch map proved to be reliable, a clinician wishing to use this task must take into consideration the compliance and ability of a caregiver to perform the task. Thus, a caregiver's sketch map could not be used if the person is unable to provide two sketch maps that are similar. Even if the two maps are consistent, a clinician must assume that these maps are accurate.

The method of scoring the sketch map can be further refined to provide more information. In this study, the score was a mean of two ratios, one being frequency (i.e. number of rooms correctly identified) and the other being accuracy (i.e. number of rooms correctly placed in relation to the others). The use of a mean of the two ratios served the purpose of decreasing the number of variables used in the factor analysis. Prior to using the mean, it had been established that the two ratios were highly correlated, suggesting that they were related. For clinical assessments, both ratios could be of interest. The study of Liu et al. (1991a) reported that AD patients had more difficulty with the accuracy than with the frequency aspects of the sketch map task. In a study using university students, Rovine and Weisman (1989) compared the ability of several pencil-paper measures to predict way-finding performance. These investigators used measures of visualization, sense-of-direction as well as sketch maps. They reported that the sketch map measures, in particular the accuracy score, were the best predictors of way-finding performance.

As mentioned earlier, the Functional Spatial Abilities Questionnaire requires further work before it can be considered for use as a clinical measure. Although the content was based on reports of caregivers, the Functional Spatial Abilities Questionnaire was not sent back to these caregivers for further feedback. The self- and proxy-rated Functional Spatial Abilities Questionnaires demonstrated high test-retest and inter-rater reliability and were able to detect differences between the AD and control groups in functional performance. These data suggest that it was an adequate measure of functional spatial orientation. There was one problem associated with its use. The scale for scoring each item was categorical but the final score was treated as interval data. A "not applicable" category was included for situations that individuals did not have to face. For example, a person with AD may rate the fourth item "I have difficulty following a map" as not applicable because this task has always been managed by the caregiver. This category may have encouraged neutral responses and may have accounted for the lack of a significant correlation between the spatial tests and the Functional Spatial Abilities Questionnaire. It may be argued that the Functional Spatial Abilities Questionnaire data would have behaved differently if the scales used ordinal data that was treated as continuous data. A study is needed to validate further the content of the Functional Spatial Abilities Questionnaire and to test a more sensitive scaling method such as the visual analogue scale.

With respect to functional assessment of spatial orientation, other more direct methods should be explored. For example, way-finding in AD patients can be evaluated using Easy Street Environments (Dix, 1991). These customized environmental simulations are being designed to be used for functional assessment and training of the physically disabled. To date, there has been no research reported on the use of Easy Street Environments for assessing way-finding ability in AD patients. These environmental simulations provide a safe, objective alternative to the assessment of way-finding in outdoor environments. This method would also be realistic and meaningful to patients. However, the major drawback is its high cost and, therefore, limited availability.

4.5.2 Clinical management of spatial disorientation in AD patients

Some of the findings can be related to the management of spatial disorientation in AD patients. The results from this study, specifically from the factor analyses, suggest that, at least in early AD, the ability to be functionally orientated may depend on factors other than cognitive spatial skills. It stands to reason that on tasks such as driving or way-finding, the environment plays an integral role in a patient's spatial orientation. The contribution of the environment to a patient's functional ability may be what is neglected in the laboratory spatial tests used. This may also explain why the scores on the Functional Spatial Abilities Questionnaire did not correlate with performance on the spatial tests. The proxy-rated Functional Spatial Abilities Questionnaire scores did not correlate with mental status scores, further supporting an environmental influence on functional ability.

Clinically, this finding suggests that environmental factors should be studied for their effects on spatial orientation. Several investigators have shown that various strategies can improve the legibility of a physical environment, or the ability of an individual to understand the spatial nature of the environment (Weisman, 1981). These strategies include arrangement of furniture, use of recognizable landmarks, use of signs and colours, and promoting the visibility of destinations (Copper, Mohide & Gilbert, 1989; Hyde, 1989; Moore, 1979; Shroyer, Hutton & Anderson, 1987; J. Weisman, 1981; G.D. Weisman, 1987). Although AD patients may have poor spatial skills, as reflected on their scores on perceptual and cognitive tests, these strategies may facilitate their functional performance.

Wandering is usually observed in patients who are in the later stages of AD and in patients who are institutionalized. Martino-Saltzman et al. (1991) have suggested that environmental factors may contribute to the types of wandering behaviour they have observed, namely, lapping, pacing and random travel. The factors they have identified are the layout of the facility, placement of furniture, lighting, noise levels and staffing (Martino-Saltzman et al., 1991). Based on these observations, it is likely that the physical aspects of an environment, as well as the capabilities of the caregivers, have an impact on the functional spatial orientation of AD patients. The impact of environmental features on functional spatial orientation will certainly be the focus of future studies, with the increasing need for special facilities for individuals with dementia (McGrowder-Lin & Bhatt, 1988; Munson et al., 1991; Ohta & Ohta, 1988; Sands & Suzuki, 1983; Weisman, 1987).

Finally, wandering behaviour is usually associated with the behavioural changes that are characteristic of late AD (American Psychiatric Association, 1987; Austrom & Hendrie, 1990;

Foldi et al., 1987). Personality changes, inappropriate behaviour, disinhibition and physical restlessness can contribute to wandering. Clinical intervention must take these factors into consideration. In addition to providing an environment that facilitates way-finding, other intervention strategies to decrease wandering include the use of structured activities and physical exercise in order to help decrease restlessness (Zgola, 1987).

4.6 Limitations of the Study

It may be argued that one limitation of this study was that many of the tests, which had standardized instructions and methods of administration, were modified. Given the global cognitive deterioration associated with AD, it was of primary concern that these patients be able to perform the tests. In order to study their spatial abilities, it was necessary to simplify the instructions, eliminate time constraints, and remind the patients of the instructions whenever necessary. As discussed, training the examiners to repeat the instructions was not adequate in ensuring good inter-rater reliability of the Figure-Ground Perception Test. In retrospect, the examiners should have also been instructed to try to elicit a response for each item when testing the AD patients. The procedures used to modify the tests have been clearly described in the Methods section so that any part of this study can be replicated.

For the control group, a small range of scores was obtained on many tests which probably accounts for the moderately low subtest-to-subtest, subtest-to-total, Cronbach's alpha and intra-class correlation coefficients. However, if the tests were made more difficult in order to increase the variability, the scores from the AD group may have clustered at the lower end of the range, resulting in the opposite effect.

The small sample size of the AD group did not permit a factor analysis to be conducted using the full battery. Therefore, interpretation of the first factor analysis cannot, at this time, be generalized to the AD group. Further study involving more AD subjects is required. With regard to sample size, the decision to use a sample of 25 subjects was based on the feasibility of the study, the clinical population under study and time constraints. This sample size is comparable to that used in most dementia studies. In addition, stringent criteria for handedness, language and age were implemented, and the requirement that patients were community residing with a caregiver who would participate in the study all ensured that the AD group was as homogeneous as possible. Finally, statistically significant findings were obtained, suggesting the sample size was adequate for the purpose of comparison with control subjects.

Ideally, separate groups of patients should be used for the reliability studies. It may be argued that by using the same patients for both the test-retest and inter-rater reliability studies, a systematic bias was introduced, particularly in the estimates of inter-rater reliability which were conducted four weeks after the initial visit. Thus, the inter-rater reliability coefficients could have been under-estimated because they were contaminated with the passage of time.

Future studies may consider other methods of estimating reliability that would not require as many visits from the AD patients. There may, however, be difficulties associated with these methods. Prior to this study, the use of another rater attending the test-retest session, or a video camera had been considered. These options were decided against because it was firmly believed that the performance and attention of the AD subjects would be negatively influenced by the presence of an observer or camera. Not all tasks (e.g. the Map-Reading Test) were administered with the subject seated which meant that a camera would need to be moved and readjusted during the session.

All of the subjects in both studies were from the vicinity of Montreal. Thus, the data from the control subjects reflect this urban, multicultural population and so the findings may not be generalizable to other populations. Also, this study did not include AD patients in a stage earlier than GDS 3, and patients in GDS 6 and 7. Patients in these stages were not studied due to the difficulty in accurately diagnosing AD in a person earlier than GDS 3, and in testing AD patients in stages 6 and 7.

4.7 Other contributions of this study

In addition to contributing to the understanding of spatial orientation, this study has made other important contributions. Based on the large size of the control group, normative data are available for many spatial tests in the format presented. Prior to this study, normative data for adults aged 50 to 80 were not available for any of the perceptual and cognitive tests used in this study. Although normative data have been reported for the Figure-Ground Perception Test (Petersen & Wikoff, 1983; Petersen et al., 1985), the Position in Space Test (Siev et al., 1986) and for the Spatial Relations Test (Taylor, 1968), all of these come from younger adults. The normative data derived from this study can contribute to other studies that examine spatial performance in AD patients or in the normal aged population.

Another contribution of this study is the information obtained on the psychometric properties of these spatial tests. Despite their long-standing use with various clinical populations, the reliability of most of these tests, in particular the Porteus Maze, Road-Map and Map-Reading tests, has never been evaluated. This information is now available not only for AD patients but also for normal elderly subjects. Although normative and psychometric data are now available for these tests, generalizability is limited because of the adaptations made to some of the tests used.

With respect to mental status, the reason for administering two versions of the mental status examination was to determine whether their psychometric properties differed. The 3MS was designed to meet the need for a screening test that is more sensitive and that samples a wider range of cognitive abilities than the MMSE (Teng & Chui, 1987). The findings indicate that both the MMSE and the 3MS are equally able to detect differences between the AD and control groups. The test-retest and inter-rater reliability of the MMSE has already been shown to be quite high (Anthony et al., 1982; Fillenbaum et al., 1987; Thal et al., 1986; Uhlmann et al., 1987), and these results were confirmed in this study. However, until now, there has been no report on the reliability of the 3MS. This study established that the 3MS does have high test-retest and inter-rater reliability when used with AD patients. Both mental status examinations were sensitive to the change in cognitive function associated with the progression of AD.

This study also demonstrated that the Bells Test was able to detect a difference between AD and control groups while Albert's Test was not. Despite this, the AD group's performance on Albert's Test was significantly lower than that of the control group. The administration of these tests to AD patients has not been reported prior to this study.

With respect to the clock drawing test, the test-retest and inter-rater reliability coefficients did not meet the .70 criterion. The low coefficients in the control group could be attributed partly to the fact that mean score was close to the maximum and the range in scores was small. However, this was not the case for the AD group which had a lower mean score with a larger range. Possibly, the scoring instructions were not as clear as was assumed. Although Doyon et al. (1991) reported a high reliability coefficient for this task, the type of coefficient used and whether the reliability pertained to test-retest or inter-rater was not specified. Using a different scoring system, Sunderland et al. (1989) established high inter-rater reliability, but there have been no reports of test-retest reliability. Most studies have reported high sensitivity and specificity for the clock drawing test (Doyon et al., 1991; Shulman et al., 1986; Sunderland et al., 1989; Wolf-Klein et al., 1989). As the inter-rater ICCs were higher than the test-retest ICCs for both groups, it is proposed that variability within subjects may also account for the overall low reliability coefficients. The findings of this study point to the need for further work to establish the reliability of the clock drawing test.

The construct validity study suggests that spatial orientation is a construct that is not related to reintegration to normal living. An explanation for the lack of correlation for the self-rated scores of the AD group, may be that the nature of the visual analogue scale was difficult for the AD patients to understand (Streiner & Norman, 1989; Wood-Dauphinee & Williams, 1987). The Reintegration to Normal Living Index includes quality of life as a major component (Wood-Dauphinee & Williams, 1987) which may explain the lack of correlation with the SOS Battery. Other measures that could have been used to examine construct validity of the SOS Battery would be cognitive and functional measures. It has been seen that many of the subtests evaluate cognitive abilities and some had components of functional abilities. In future studies, the perceptual and cognitive subtests could be validated with a test using the working memory model which deals with the visuospatial domain. The Functional Spatial Abilities Questionnaires could be validated with functional measures that involve instrumental activities of daily living such as going to the bank and travelling by bus.

Although it was not a focus of the study, the lower scores of the AD patients on the self- and proxy-rated Reintegration to Normal Living Index suggest that they experience a poorer quality of life in comparison to the control group. Finally, the normative data for the Reintegration to Normal Living Index contributes to future studies examining this construct.

5.0 CONCLUSION

This study has made an original contribution by providing data on the relationships between various spatial skills in order to understand better spatial orientation in AD. This was achieved by selecting a wide range of spatial tests that required perceptual and cognitive skills to complete. A new measure was developed in order to study the relationship between functional spatial orientation and performance on spatial tests. The conclusions drawn from this study are: 1) For healthy individuals, spatial orientation skills as measured on the full SOS Battery can be categorized into five types. These categories probably do not apply to AD patients and further research is needed to determine how scores of a larger group of AD patients would load in a similar analysis. 2) Seven subtests of the full battery demonstrate high ICC values for both test-retest and inter-rater reliability when used with AD patients. These are the Position in Space, Porteus Maze, Map-Reading, Road-Map, sketch map of the familiar environment and the self-rated and proxy-rated Functional Spatial Abilities Questionnaires. These results may guide the clinician in selecting spatial tests to use with AD patients. 3) All of the subtests are sensitive to differences between early AD and control subjects. But only five subtests are sensitive to the severity of AD. These subtests are the Figure-Ground Perception, Road-Map, sketch map of the familiar environment, sketch map of the new environment and proxy-rated Functional Spatial Abilities Questionnaire.

This study has also provided data on the psychometric properties of well-known spatial tests. Results that can be specifically applied to the AD population are needed in this area of research. While the establishment of the validity of a test is an ongoing process, the test must first prove to be stable or consistent over time and between raters. The Map-Reading and Road-

Map tests have been used frequently to study spatial orientation in AD, but reliability data for these measures were lacking. The high reliability coefficients and good validity data obtained in this study support the use of these tests. The Position in Space Test appears to be a good clinical tool for assessing perceptual spatial skills. The sketch map task is not new to environmental psychologists who use it to study mental representation in the normal population. Based on this study, the sketch map task is reliable and valid for use with AD patients. These tests should not be limited to research studies, however, and should be used by clinicians to evaluate AD patients. A new measure, the Functional Spatial Abilities Questionnaire, has good psychometric properties but requires more research to establish its content validity and to develop a more sensitive scoring system.

Other unique contributions of this study are the normative data obtained on a wide range of spatial tests. While some normative data exist for the perceptual tests, these pertain to younger adult groups. The data from this study will be useful in future research on AD as well as on normal aging. With respect to the other tests used, which are not part of the SOS Battery, data now exist pertaining to their psychometric properties. For example, psychometric data are now available for the 3MS and the clock drawing task. Normative data are also available for the Reintegration to Normal Living Index.

Some immediate research questions to be addressed are: 1) If factor analysis is conducted using a larger group of AD patients and all 13 subtests, would the results replicate those obtained in the first factor analysis of this study which used data from the control subjects; 2) How would AD patients perform if some of these spatial tests were reexamined with respect to the specific skills required? For example, would AD patients also be impaired in the NO-TURN condition of the Map-Reading Test which examines only extrapersonal orientation and eliminates the need for mental rotation? 3) Would the Functional Spatial Abilities Questionnaire be a better measure if a more sensitive scale was applied? 4) What are the environmental factors that influence functional spatial orientation? Other variables to consider would be a person's premorbid personality and spatial ability, caregiver availability and interventions or strategies used by caregivers and patients to deal with functional spatial problems.

In conclusion, as the general population ages, the number of individuals with AD in Canada is increasing. Spatial disorientation is a serious problem in the management of these patients. This study has made a unique contribution to knowledge with respect to the assessment and understanding of spatial disorientation in AD patients. With further research and increased understanding of this deficit, it is anticipated that a positive impact can be made on the clinical intervention and daily management of AD patients.

Bibliography

Albert, M. L. (1973). A simple test of visual neglect. Neurology, 23, 658-664.

Aldelstein, T. B., Kesner, R. P., & Strassberg, D. S. (1992). Spatial recognition and spatial order memory in patients with dementia of the Alzheimer's type. Neuropsychologia, 30(1), 59-67.

Alexander, D., Walker, H. T., & Money, J. (1964). Studies in direction sense, I. Turner's Syndrome. Archives of General Psychiatry, 10, 337-339.

Alexander, M. P., & Geschwind, N. (1984). Dementia in the elderly. In M.L. Albert (Ed.), Clinical Neurology of Aging (pp. 254-276). New York: Oxford University Press.

Allen, G. L., Siegel, A. W., & Rosinski, R. R. (1978). The role of perceptual context in structuring spatial knowledge. Journal of Experimental Psychology, Human Learning and Memory, 4(6), 617-630.

American Psychiatric Association (1987). Diagnostic and Statistical Manual of Mental Disorders (3rd ed.). Washington: American Psychiatric Association.

Anthony, J. C., LeResche, L., Niaz, U., Von Korff, M. R., & Folstein, M. F. (1982). Limits of the "Mini-Mental State" as a screening test for dementia and delirium among hospital patients. Psychological Medicine, 12, 397-408.

Anastasi, A. (1982). Psychological testing (5th ed.), (Chapters 4-8). New York: MacMillan.

Appleyard, D. (1970). Styles and methods of structuring a city. Environment and Behavior, 2, 100-117.

Aubrey, J. B., & Dobbs, A. R. (1989). Age differences in extrapersonal orientation as measured by performance on the locomotor maze. Canadian Journal on Aging, 8(4), 333-342.

Austrom, M. G., Hendrie, H. C. (1990). Death of the personality: The grief response of the Alzheimer's disease family caregiver. The American Journal of Alzheimer's Care and Related Disorders & Research, 5(2), 16-27.

Ayres, A. J. (1972). Southern California Sensory Integration Tests. Los Angeles: Western Psychological Services.

Baddeley, A. (1992). Working Memory. Science, 255, 556-559.

Baddeley, A. D., Bressi, S., Della Sala, S., Logie, R., & Spinnler, H. (1991). The decline of working memory in Alzheimer's disease: A longitudinal study. Brain, 114, 2521-2542.

Baddeley, A., Logie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and working memory. Quarterly Journal of Experimental Psychology, 38A, 603-618.

Baddeley, A., Sunderland, A., & Harris, J. (1982). How well do laboratory-based psychological tests predict patients' performance outside the laboratory? In S. Corkin, K. L. Davis, J. H. Growdon, E. Usdin, & R. J. Wurtman (Eds.). Alzheimer's disease: A Report of Progress in Research, Aging, 19, 141-148. New York: Raven Press.

Barker, R. G. (1931). The stepping-stone maze: A directly visible space problem apparatus. Journal of General Psychology, 5, 280-285.

Barr, W. B., Bilder, R. M., & Kaplan, E. (1990). Pathophysiologic mechanisms underlying spatial disorientation in patients with Alzheimer's disease. Archives of Neurology, 47, 618.

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. Psychological Reports, 19, 3-11.

Baum, C. M., Edwards, D. F., Leavitt, K., Grant, E., & Deuel, R. M. (1988). Performance components in senile dementia of the Alzheimer type: Motor planning, language, and memory. The Occupational Therapy Journal of Research, 8(6), 356-368.

Bayles, K. A. & Kaszniak, A. W., & Tomeada, C. K. (1987). Perception and attention in normal aging and dementia. Communication and cognition in normal aging and dementia (pp. 249-263). Boston: Little and Brown.

Becker, J. T., Huff, F. J., Nebec, R. D., Holland, A., & Boller, F (1986). Neuropsychological function in Alzheimer's disease: Pattern of impairment and rates of progression. Archives of Neurology, 45, 263-268.

Benton, A. L. (1959). Right-Left Discrimination and Finger Localization. New York: Hoeber.

Bleathman, D. (1987). The practical management of the Alzheimer's disease patient in the hospital setting. Journal of Advanced Nursing, 12, 531-534.

Bleecker, M. L., Bolla-Wilson, K., Kawas, C., & Agnew, J. (1988). Age-specific norms for the Mini-Mental State Exam. Neurology, 38, 1565-1568.

Bohrnstedt, G. W. (1983). Measurement. In P. H. Rossi, J. D. Wright, & A. B. Anderson (Eds.), Handbook of Survey Research, (pp. 69-121). New York: Academic Press.

Bowles, N. L., Obler, L. K., & Albert, M. L. (1987). Naming errors in healthy aging and dementia of the Alzheimer type. Cortex, 23, 519-524.

Botwinick, J. (1984). Aging and Behavior (3rd ed.). New York: Springer.

Branconnier, R. J., & DeVitt, D. R. (1983). Early detection of incipient Alzheimer's disease: Some methodological considerations on computerized diagnosis. In B. Reisberg (Ed.), Alzheimer's disease: The standard reference (pp. 214-227). New York: The Free Press.

Brantjes, M., & Bouma, A. (1991). Qualitative analysis of the drawings of Alzheimer's patients. The Clinical Neuropsychologist, 5(1), 41-52.

Breitner, J., Foldi, N. S., Rabins, P. V., Sunderland, T., & Butler, R. N. (1987). Practical considerations in managing Alzheimer's disease: I. Geriatrics, 42, 78-98.

Broe, G. A., Akhtar, A. J., Andrews, G. R., Caird, F. I., Gilmore, A. J. J., & McLennan, W. J. (1976). Neurological disorders in the elderly at home. Journal of Neurology, Neurosurgery, and Psychiatry, 39, 362-366.

Brouwers, P., Cox, C., Martin, A., Chase, T., & Fedio, P. (1984). Differential perceptual-spatial impairment in Huntington's and Alzheimer's dementias. Archives of Neurology, 41, 1073-1076.

Bryden, M. P. (1977). Measuring handedness with questionnaires. Neuropsychologia, 15, 617-624.

Buffery, A. W. H., & Gray, J. A. (1972). Sex differences in the development of spatial and linguistic skills. In C. Ounsted, & D. C. Taylor (Eds.). Gender differences: Their ontogeny and significance. London: Churchill-Livingston.

Butters, N., Barton, M., & Brody, B. A. (1972). Role of the right parietal lobe in the mediation of cross-modal associations and reversible operations in space. Cortex, 6, 174-190.

Butters, N., & Soeldner, C. (1972). Comparison of parietal and frontal lobe spatial deficits in man: Extrapersonal versus personal (egocentric) space. Perceptual and Motor Skills, 34, 27-34.

Bylsma, F. W., Brandt, J., & Straus, M. (1992). Personal and extra-personal orientation in Huntington's disease patients and those at risk. Cortex, 28, 113-122.

Carmines, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. Beverly Hills: Sage Publications.

Carpman, J. R., Grant, M. A., & Simmons, D. A. (1985). Hospital design and wayfinding: A video simulation study. Environment and Behavior, 17(3), 296-314.

Chenoweth, B., & Spencer, B. (1986). Dementia: the experience of family caregivers. The Gerontologist, 26(3), 267-272.

Cogan, D. G. (1985). Visual disturbances with focal progressive dementing disease. American Journal of Ophthalmology, 100, 68-72.

Cogan, D. G. (1987). Alzheimer syndromes. American Journal of Opthalmology, August, 183-184.

Cohen, G. D. (1983). Historical views and evolution of concepts. In B. Reisberg (Ed.), Alzheimer's disease: The standard reference (pp. 29-33). New York: Free Press.

Cohn, J. B., Wilcox, C. S., & Lerer, B. E. (1991). Development of an "early" detection battery for dementia of the Alzheimer type. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 15, 433-479.

Cole, M. G., & Dastoor, D. P. (1987). A new hierarchic approach to the measurement of dementia. Psychosomatics, 28, 298-304.

Consensus Conference (1987). Differential diagnosis of dementing diseases. Journal of the American Medical Association, 258, 3411-3416.

Cooper, B., Mohide, A., & Gilbert, S. (1989). Testing the use of color in a long-term care setting. Dimensions in Health Service, 66(6), 22-26.

Corkin, S. (1982). Some relationships between global amnesias and the memory impairments in Alzheimer's disease. In S. Corkin, K. L. Davis, J. H. Growdon, E. Usdin and R. J. Wurtman (Eds.), Alzheimer's disease: A report of research in progress. New York: Raven Press.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334.

Crook, T., Ferris, S., & McCarthy, M. (1979). The misplaced-objects task: A brief test for memory dysfunction in the aged. Journal of the American Geriatrics Society, 27(6), 284-287.

Culver, C. M. (1969). Test of right-left discrimination. Perceptual and Motor Skills, 29, 863-867.

Cummings, J. L., & Benson, D. F. (1986). Dementia of the Alzheimer type: an inventory of diagnostic clinical features. Journal of the American Geriatrics Society, 34, 12-19.

Cummings, J. L., Houlihan, J. P., & Hill, M. A. (1986). The pattern of reading deterioration in dementia of the Alzheimer type: Observations and implications. Brain and Language, 29(2), 315-323.

Del Greco, L., Walop, W., & McCarthy, R. H. (1987). Questionnaire development: 2. Validity and reliability. Canadian Medical Association Journal, 136, 699-700.

De Renzi, E. (1982). Disorders of space exploration and cognition. New York: John Wiley & Sons.

Dix, A. (1991). Easy does it. Hospital Development, Nov., 23-25.

Downs, R. M., & Stea, B. D. (1981). Essai sur la cartographie mentale des cartes plein la tete. St. Hyacinthe, Quebec: Edisem Inc.

Doyon, J., Bouchard, R., Morin, G., Bourgeois, C., & Côté, D. (1991). Nouveau système de cotation quantitative de l'horloge: Sensibilité et spécificité dans la démence de type Alzheimer. L'Union Médicale au Canada Français, 120, 119.

Drachman, D. (1988). Who may drive? Who may not? Who shall decide? Annals of Neurology, 24(6), 787-788.

Dubinsky, R. M., Williamson, A., Gray, C. S., & Glatt, S. L. (1992). Driving in Alzheimer's disease. Journal of the American Geriatrics Society, 40(11), 1112-1116.

Eichelman, B., Phelps, C. (1992). Management of behavior disturbances in Alzheimer's disease - current knowledge and future directions. Alzheimer's disease and Associated Disorders, 6(2), 77-88.

Eustache, F. (1992). Les troubles de la mémoire et du language dans les démences de type Alzheimer. Revue Neurologigue (Paris), 148(3), 18-192.

Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. Quarterly Journal of Experimental Psychology, 38A, 675-688.

Feinstein, A. R. (1987). Clinimetrics. (Chapter 11). New Haven: Yale University Press.

Ferris, S. H., Crook, T., Clark, E., McCarthy, M., & Rae, D. (1980). Facial recognition memory deficits in normal aging and senile dementia. Journal of Gerontology, 35(5), 707-714.

Fine, E. J., Mellstrom, M., Mani, S. S., & Timmins, J. (1980). Spatial disorientation and the Dyke-Davidoff-Masson Syndrome. Cortex, 16, 493-499.

Fischer, P., Marterer, A., & Danielczyk, W. (1990). Right-left disorientation in dementia of the Alzheimer type. Neurology, 40, 1619-1620.

Fillenbaum, G. G., Hayman, A., Wilkinson, W. E., & Hanes, C. S. (1987). Comparison of two screening tests in Alzheimer's disease. The correlation of the reliability of the Mini-Mental State Examination and the Modified Blessed Test. Archives of Neurology, 44, 924-927.

Flicker, C., Bartus, R. T., Crook, T. H., & Ferris, S. H. (1984). Effects of aging and dementia upon recent visuospatial memory. Neurobiology of Aging, 5, 275-283.

Flicker, C., Ferris, S. H., Crook, T., & Bartus, R. T. (1987). Implications of memory and language dysfunction in the naming deficit of senile dementia. Brain and Language, 31, 187-200.

Flicker, C., Ferris, S. H., Crook, T., Bartus, R. T., & Reisberg, B. (1986). Cognitive decline in advanced age: Future directions for the psychometric differentiation of normal and pathological age changes in cognitive function. Developmental Neuropsychology, 2(4), 309-322.

Flicker, C., Ferris, S. H., Crook, T., Reisberg, B., & Bartus, R. T. (1988). Equivalent spatial-rotation deficits in normal aging and Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology, 10(4), 387-399.

Foldi, N. S., Jutagir, R., Davidoff, D., & Gould, T. (1992). Selective attention skills in Alzheimer's disease: Performance on graded cancellation tests varying in density and complexity. Journal of Gerontology: Psychological Sciences, 47(3), 146-153.

Folstein, M., Anthony, J. C., Parhad, T., Duffy, B., & Gruenberg, E. M. (1985). The meaning of cognitive impairment in the elderly. Journal of the American Geriatric Society, 33, 228-235.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental" A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12, 189-198.

Foster, J. K., Behrmann, M., & Stuss, D. T. (1992). Attentional dysfunction in Alzheimer's disease. Transcript of Poster presented at TENNET III. Montreal: TENNET.

Fozard, J. L., Wolf, E., Bell, B., McFarland, R. A., & Podolsky, S. (1977). Visual perception and communication. In J.E. Birren & K.W. Schaie (Eds.), Handbook of the Psychology of Aging (pp. 497-534). New York: Van Nostrand Reinhold.

Friedland, R. P., Koss, E., Haxby, J. V., Grady, C. L., Luxenberg, J., Schapiro, M. B., & Kaye, J. (1988). Alzheimer's disease: Clinical and biological heterogeneity. Annals of Internal Medicine, 109, 298-311.

Furrie, A. (1987). A national database on disabled persons making disability data available to users. Ottawa: Statistics Canada.

Gauthier, L., Dehaut, F., & Joanette, Y. (1989). The Bells Test: A quantitative and qualitative test for visual neglect. International Journal of Clinical Neuropsychology, 11, 49-54.

Gauthier, L., & Joanette, Y. (1992). Elaboration of an assessment of hemispatial neglect: The

Bells Test. Abstracts Attention: Theoretical and clincial perspectives, p. 14. Toronto: Rotman Research Institute.

Gauthier, L., Liu, L., & Gauthier, S. (1990). La désorientation spatiale chez les personnes ayant la maladie d'Alzheimer. L'Union Médicale au Canada Français, 120, 167.

Golden, C. J., Sawicki, R. F., & Franzen, M. D. (1984). Test construction. In G. Goldstein & M. Hersen (Eds.), Handbook of psychological assessment (pp. 19-37). New York: Pergamon.

Gottfries, C. G. (1985). Definition of normal aging, senile dementia and Alzheimer's disease. In C. G. Gottfries (Ed.), Normal aging, Alzheimer's disease, and senile dementia: Aspects on etiology, pathogenesis, diagnosis, and treatment. Proceedings of two symposia held at the Collegium Internale Neuropharmologicum 14th Congress (pp. 11-17). Florence, Italy.

Gottlieb, G. L., Gur, R. E., & Gur, R. C. (1988). Reliability of psychiatric scales in patients with dementia of the Alzheimer type. American Journal of Psychiatry, 145(7), 857-860.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. Journal of Experimental Psychology: Learning, Memory and Cognition, 11, 501-518.

Grosse, D. A., Wilson, R. S., & Fox, J. H. (1991). Maze learning in Alzheimer's disease. Brain and Cognition, 15, 1-9.

Gustafason, L., & Milsson, L. (1982). Differential diagnosis of presenile dementia on clinical grounds. Acta Psychiatrica Scandanavica, 65, 194-209.

Guyatt, G., Walter, S., Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. Journal of Chronic Diseases, 40(2), 171-178.

Habib, M., & Sirigu, A. (1987). Pure topographical disorientation: A definition and anatomical basis. Cortex, 23, 73-85.

Hachinski, V. C., Iliff, L. D., Zilhka, E., Du Boulay, G. H., McAllister, V. L., Marshall, J., Russell, R. W. R., & Symon, L. (1975). Cerebral blood flow in dementia. Archives of Neurology, 32, 622-637.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In A. Fessard, R. W. Gerard, J. Konorski, & J. F. Delafresnaye (Eds.), Brain mechanisms and learning, p. 37-51. Oxford: Blackwell Scientific Publications.

Henderson, V. W., Mack, W., & Williams, B. W. (1989). Spatial disorientation in Alzheimer's disease. Archives of Neurology, 46, 391-394.

Huff, F. J., Becker, J. T., Belle, S. H., Nebes, R. D., Holland, A. L., & Boller, F. (1987). Cognitive deficits and clinical diagnosis of AD. Neurology, 37, 1119-1124.

Hyde, J. (1989). The physical environment and the care of Alzheimer's patients: An experiential survey of Massachusetts' Alzheimer Units. The American Journal of Alzheimer's Care and Related Disorders & Research, 4(3), 36-44.

Jorm, A. F. (1991). Cross-national comparisons of teh occurrence of Alzheimer's and vascular dementias. European Archives of Psychiatry and Clinical Neuroscience, 240: 218-222.

Kahn, R. L., Goldfarb, A. I., Pollack, M., Peck, A. (1960). Brief objective measures for the determination of mental status in the aged. American Journal of Psychiatry, 117, 326-328.

Kaplan, E., Goodglass, H., & Weintraub, S. (1976). The Boston Naming Test (Experimental ed.). Boston.

Kaplan, H. I., & Sadock, B. J. (1981). Modern synopsis of comprehensive textbook of psychiatry (3rd ed.). Baltimore: Williams and Wilkins.

Katzman, R. (1986). Alzheimer's disease. New England Journal of Medicine, 314(15), 964-973.

Kaneko, Z. (1875). Care in Japan. In Howells, J. G. (ed). Modern Perspectives in the Psychiatry of Old Age. New York: Brunner/Mazel, pp. 519-530.

Kay, D. W. K., Henderson, A. S., Scott, R., Wilson, J., Ridkwood, D., & Grayson, D. A. (1985). Dementia and depression among the elderly living in the Hobart Community: The effect of the diagnostic criteria on the prevalence rates. Psychological Medicine, 15, 771-788.

Kertesz, A. (1980). Western Aphasia Battery. London, Canada: University of Western Ontario.

Kesner, R. P., Adelstein, T. B., & Crutcher, K. A. (1989). Equivalent spatial location memory deficits in rats with medial septum or hippocampal formation lesions and patients with dementia of the Alzheimer's type. Brain and Cognition, 9, 289-300.

Kirk, A., & Kertesz, A. (1991). On drawing impairment in Alzheimer's disease. Archives of Neurology, 48, 73-77.

Kittner, S. J., White, L. R., Farmer, M. E., Wolz, M., Kaplan, E., Moes, E., Brody, J. A., & Feinleib, M. (1986). Methodological issues in screening for dementia: The problem of education adjustment. Journal of Chronic Diseases, 39(3), 163-170.

Kline, P. (1979). Psychometrics and psychology. New York: Academic Press.

Kline, P. (1986). A Handbook of test construction. London: Methuen.

Knesevich, J. M., LaBarge, E., & Edwards, D. (1986). Predictive value of the Boston Naming Test in mild senile dementia of the Alzheimer type. Psychiatric Research, 19, 155-161.

Kokmen, E. (1989). Occurrence of motor vehicle crashes among patients with Alzheimer's disease. Annals of Neurology, 26(3), 415-416.

Kopelman, M. D. (1985). Rates of forgetting in Alzheimer-type dementia and Korsakoff's syndrome. Neuropsychologia, 23(5), 623-638.

Kozlowski, L. T., & Bryant, K. J. (1977). Sense of direction, spatial orientation, and cognitive maps. Journal of Experimental Psychology, 3(4), 590-598.

Kukull, W. A., Larson, E. B., Reifler, B. V., Lampe, T. H., Yerby, M. S., & Hughes, J. P. (1990). The validity of 3 clinical diagnostic criteria for Alzheimer's disease. Neurology, 40, 1364-1369.

Liu, L., Gauthier, L., & Gauthier, S. (1991a). Deficits in spatial orientation skills in persons with early Alzheimer's disease. American Journal of Occupational Therapy, 45(1), 67-74.

Liu, L., Gauthier, L., & Gauthier, S. (1991b). Assessment of extrapersonal orientation in individuals with Alzheimer's disease using the locomotor maze test. Canadian Journal of Occupational Therapy, 58, Supp.: 26.

Loring, D. W., & Largen, J. W. (1985). Neuropsychological patterns of presenile and senile dementia of the Alzhiemer type. Neuropsychologia, 23(3), 351-357.

Lucas-Blaustein, M. J., Filipp, L., Dungan, C., & Tune, L. (1988). Driving in patients with dementia. Journal of the American Geriatrics Society, 36, 1087-1091.

Maccoby, E. E., & Jacklin, C. N. (1974). The psychology of sex differences. Stanford: Stanford University Press.

Mack, W. J., Freed, D. M., White Williams, B., & Henderson, V. W. (1992). Boston Naming Test: Shortened versions for use in Alzheimer's disease. Journal of Gerontology: Psychological Sciences, 47(3), P154-158.

Magaziner, J., Simonsick, E. M., Kashner, M., & Hebel, J. R. (1988). Patient-proxy response comparability on measures of patient health and functional status. Journal of Clinical Epidemiology, 43(11), 1065-1074.

Martin, A. (1987). Representation of semantic and spatial knowledge in Alzheimer's patients: Implications for models of preserved learning in amnesia. Journal of Clinical and Experimental Neuropsychology, 9(2), 191-224.

Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. Brain and Language, 19, 124-141.

Martin, A., Brouwers, P., Lalonde, F., Cox, C., Teleska, P., & Fedio, P. (1986). Towards a behavioral typology of Alzheimer's patients. Journal of Clinical and Experimental Neuropsychology, 9(2), 191-224.

Martin, G. M., & Bryant, E. M. (1988). Genetics of aging and of disease models. In R. D. Tery (Ed.), Aging: Aging and the Brain (Vol. 32, pp. 1-16). New York: Raven Press.

Martino-Saltzman, D., Blasch, B. B., Morris, R. D., & McNeal, L. W. (1991). Travel behavior of nursing home residents perceived as wanderers and nonwanderers. The Gerontologist, 31(5), 666-672.

Mas, J. L., Alperovitch, A., & Derouesne, C. (1987). Epidémiologie de la démence de type

Alzheimer. Review Neurologue (Paris), 143(3), 161-171.

Mayeux, R., Stern, Y., & Spanton, S. (1985). Heterogeneity in dementia of the Alzheimer type: Evidence of subgroups. Neurology, 35, 453-461.

McEwan, K. L., Donnelly, M., Robertson, D., & Hertzman, C. (1991). Mental health problems among Canada's seniors: Demographic and epidemiologic considerations. Ottawa: National Health and Welfare.

McFie, J., Peircy, M. F., & Zangwill, O. L. (1950). Visual-spatial agnosia associated with lesions of the right cerebral hemisphere. Brain, 73, pp. 167-190.

McGlynn, S. M., & Kaszniak, A. W. (1991). When metacognition fails: Impaired awareness of deficit in Alzheimer's disease. Journal of Cognitive Neuroscience, 3(2), 183-189.

McGrowder-Lin, R., & Bhatt, A. (1988). A wanderer's lounge program for nursing home residents with Alzheimer's disease. The Gerontologist, 28(5), 607-609.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (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, 939-944.

Mendez, M. F., Ala, T., & Underwood, K. L. (1992). Development of scoring criteria for the clock drawing task in Alzheimer's disease. Journal of the American Geriatrics Society, 40(11), 1095-1099.

Mendez, M. F., Mendez, M. A., Martin, R., Smyth, K. A., & Whitehouse, P. J. (1990). Complex visual disturbances in Alzheimer's disease. Neurology, 40, 439-443.

Milner, B. (1965). Visually-guided maze learning in man: Effects of bilateral hippocampal, bilateral frontal, and unilateral cerebral lesions. Neuropsychologia, 3, 317-338.

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. British Medical Bulletin, 27(3), 272-277.

Milner, B. (1980). Complementary functional specializations of the human cerebral hemispheres. In R. Levi-Montalcini (Ed.), Nerve Cells, Transmitters and Behavior (pp. 601-616). Vatican City: Pontificia Academia Scientiarum.

Moeser, S. D. (1988). Cognitive mapping in a complex building. Environment and Behavior, 20(1), 21-49.

Mohr, E., Claus, J. J., Mann, U. M., & Chase, T. N. (1991). Dementia disorders. In E. Mohr & P. Brouwers (Eds.), Handbook of clinical trials: The neurobehavioral approach. Amsterdam: Swets & Zeitlinger.

Molsa, P. K., Paljarri, L., Rinne, J. O., Rinne, U. K., & Sako, E. (1985). Validity of clinical diagnosis in dementia: A prospective clinicopathological study. Journal of Neurology, Neurosurgery, and Psychiatry, 48, 1085-1090.

Money, J. A. (1976). Standardized Road Map Test of Directional Sense. Manual. Sam Rafael, CA: Academic Therapy Publications.

Moore, G. T. (1979). Knowing about environmental knowing: The current state of theory and research on environmental cognition. Environment and Behavior, 11(1), 33-70.

Morris, R. G. (1984). Dementia and the functioning of the Articulatory Loop System. Cognitive Neuropsychology, 1, 143-157.

Morris, R. G., & Baddeley, A. D. (1988). Primary and working memory functioning in Alzheimer-type dementia. Journal of Clinical and Experimental Neuropsychology, 10(2), 279-296.

Morris, R. G., & Kopelman, M. D. (1986). The memory deficits in Alzheimer-type dementia: a review. Quarterly Journal of Experimental Psychology, 38A, 575-602.

Morris, J. C., McKeel, D. W., Fulling, K., Torack, R. M., & Berg, L. (1988). Validation of clinical diagnostic criteria for Alzheimer's disease. Annals of Neurology, 24, 17-22.

Morris, R. G., Wheatley, J., & Britton, P. G. (1983). Retrieval from long-term memory in senile dementia: Cued and recall revisited. British Journal of Clinical Psychology, 22, 141-142.

Moscovitch, M. (1982). A neuropsychological approach to perception and memory in normal and pathological aging. In F. I. M. Craik, & S. Trehub (Eds.), Aging and cognitive processes. New York: Plenum Press.

Munson, P., Awde, R., Badiou, D., Bourke, E., Jones, M. J. D., Lips, T. J., & Schiff, M. (1991). Designing facilities for people with dementia, Ottawa: Health and Welfare Canada.

Murdoch, B. E., Chenery, H. J., Wilks, V., & Boyle, R. S. (1987). Language disorders in dementia of the Alzheimer type. Brain and Language,31, 122-137.

Newcombe, F., & Russell, R. (1969). Dissociated visual perceptual and spatial deficits in focal lesions of the right hemisphere. Journal of Neurology, Neurosurgery, and Psychiatry, 32, 73-81.

Nissen, M. J., Corkin, S., Buonanno, F. S., Growdon, J. H., Wray, S., & Bauer, J. (1986). Spatial vision in Alzheimer's disease: General findings and a case report. Archives of Neurology, 42, 667-671.

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw Hill.

Obler, L. K., & Albert, M. L. (1979). Action Naming Test (Experimental ed.).

Ohta, R. J., & Ohta, B. M. (1988). Special units for Alzheimer's disease patients: A critical look. The Gerontologist, 28(6), 803-808.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia, 9, 97-113.

Passini, R. (1984a). Wayfinding in Architecture. New York: Van Nostrand Reinhold.

Passini, R. (1984b). Spatial representations, a wayfinding perspective. Journal of Environmental Psychology, 4, 153-164.

Pappas, B.A., Sunderland, T., Weingartner, H.M., Vitiello, B., Martinson, H., & Putnam, K. (1992). Alzheimer's disease and feeling-of-knowing for knowledge and episodic memory. Journal of Gerontology: Psychological Sciences, 47(3), 159-164.

Petersen, P., Goar, D., & Van Deusen, J. (1985). Performance of female adults on the Southern California figure-ground visual perception test. American Journal of Occupational Therapy, 39(8), 525-530.

Petersen, P., & Wikoff, R. L. (1983). The performance of adult males on the Southern California figure-ground visual perception test. American Journal of Occupational Therapy, 37(8), 554-560.

Plude, D. J., Milberg, W. P., & Cerella, J. (1986). Age differences in depicting and perceiving tridimensionality in simple line drawings. Experimental Aging Research, 12(4), 221-225.

Porteus, S. D. (1959). The Maze Test and Clinical Psychology. Palo Alto: Pacific Books.

Pothig, V. D., Pogelt, B., & Roth, N. (1985). Zur Anwendung des Labyrinthtestes in der gerontologischen Funktionsdiagnostik (Use of labyrinth tests in geriatric functional testing). Z. Alternsforsch, 40(1), 23-28.

Raczkowski, D., & Kalat, J. W. (1974). Reliability and validity of some handedness questionnaire items. Neuropsychologia, 12, 43-47.

Ratcliff, G. (1982). Disturbances of spatial orientation associated with cerebral lesions. In M. Portegal (Ed.), Spatial Abilities (pp. 301-331). New York: Academic Press.

Ratcliff, G., & Newcombe, F. (1973). Spatial orientation in man: Effects of left, right, and bilateral posterior cerebral lesions. Journal of Neurology, Neurosurgery and Psychiatry, 36, 448-454.

Reisberg, B., Ferris, S. H., De Leon, M. J., & Crook, T. (1982). The global deterioration scale for assessment of primary degenerative dementia. American Journal of Psychiatry, 139(9), 1136-1139.

Reisberg, B., Schneck, M. K., Ferris, S. H., Schwartz, G. E., & De Leon, M. J. (1983). The brief cognitive rating scale (BCRS): Findings in primary degenerative dementia (PDD). Psychopharmacology Bulletin, 19(1), 47-50.

Rissenberg, M., & Glanzer, M. (1987). Free recall and word finding ability in normal aging and senile dementia of the Alzheimer's type: The effect of item concreteness. Journal of Gerontology, 42(3), 318-322.

Rocca, W. A., Amaducci, L. A., & Schoenberg, B. S. (1986). Epidemiology of clinically diagnosed Alzheimer's disease. Annals of Neurology, 19(5), 415-424.

Rocca, W. A., Fratiglioni, L., Bracco, L., Pedone, D., Groppi, C., & Schoenberg, B. S. (1986). The use of surrogate respondents to obtain questionnaire data in case-control studies of neurologic diseases. Journal of Chronic Diseases, 39(1), 907-912.

Ross, E. D. (1980). Sensory-specific and fractional disorders of recent memory in man. Archives of Neurology, 37, 193-200.

Rouleau, I., Salmon, D. P., Butters, N., Kennedy, C., & McGuire, K. (1992). Quantitative and qualitative analyses of clock drawings in Alzheimer's and Huntington's disease. Brain and Cognition, 18, 70-87.

Rovine, M. J., & Weisman, G. D. (1989). Sketch-map variables as predictors of way-finding performance. Journal of Environmental Psychology, 9, 217-232.

Rudel, R. G., Teuber, H. L., & Twitchell, T. E. (1974). Levels of impairment of sensory-motor functions in children with early brain damage. Neuropsychologia, 12, 95-108.

Sahakian, B. J., Morris, R., Evenden, J. L., Heald, A., Levy, R., Philpot, M., & Robbins, T. W. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. Brain, 111, 695-718.

Sands, D., & Suzuki, T. (1983). Adult day care for Alzheimer's patients and their families. The Gerontologist, 23(1), 21-23.

SAS Institute Inc. (1990). SAS/STAT User's Guide (version 6, 4th ed.), volumes 1 and 2. Cary, NC: SAS Institute Inc.

Satterthwaite, F. W. (1946). An approximate distribution of estimates of variance components. Biometrics Bulletin, 2, 110-114.

Scarborough, D., Cortese, C., & Scarborough, H. (1977). Frequency and repetition effects in lexical memory. Journal of Experimental Psychology: Human Perception and Performance, 3, 1-17.

Schoenberg, B. S., Anderson, D. W., & Haerer, A. F. (1985). Severe dementia. Prevalence and clinical features in a biracial U.S. Population. Archives of Neurology, 42, 740-743.

Semmes, J., Weinstein, S., Ghent, L.. & Teuber, H.-L. (1963). Correlates of impaired orientation in personal and extrapersonal space. Brain, 86, 747-772.

Semmes, J., Weinstein, S., Ghent, L., & Teuber, H.-L. (1955). Spatial orientation in man after cerebral injury: I. Analysis by locus of lesions. The Journal of Psychology, 39, 227-243.

Shroyer, J. L., Jutten, J. T.. & Anderson, G. M. (1987). The Alzheimer patient: Interior design considerations. Texas Medicine, 83, 654-657.

Shulman, K., Shedletsky, R., & Silver, I. L. (1986). The challenge of time: Clock-drawing and cognitive function in the elderly. Internal Journal of Geriatric Psychiatry, 1, 135-140.

Siev, E., Freishtat, B., & Zoltan, B. (1986). Perceptual and cognitive dysfunction in the adult stroke patient (rev ed.). New Jersey: Slack.

Sluss, T. K., Gurenberg, E. M., & Kramer, M. (1981). The use of longitudinal studies in the investigation of risk factors for dementia-Alzheimer type. In J. A. Mortimer & L. M. Schuman (Eds.), The Epidemiology of Dementia. Oxford: Oxford University Press.

Smyth, M. M., Morris, P. E., Levy, P., & Ellis, A. W. (1987). Cognition in Action (pp. 239-259). London: Lawrence Erlbaum Associates.

Spitzer, W. O., Dobson, A. J., Hall, J., Chesterman, E., Levi, J., Shepherd, R., Battista, R. N.,

& Catchlove, B. R. (1981). Measuring the quality of life of cancer patients. Journal of Chronic Diseases, 34, 585-597.

Spreen, O., & Benton, A. L. (1977). Neurosensory Centre Comprehensive Examination for Aphasia. Victoria: University of Victoria.

Spreen, O., & Strauss, E. (1991). A compendium of neuropsychological tests: Administration, norms, and commentary. New York: Oxford University Press.

Squire, L. R. (1987). Memory and brain. New York: Oxford University Press.

Steel, R. G. D., & Torrie, J. H. (1980). Principles and procedures of statistics, (2nd ed.). New York: McGraw-Hill.

Stiles-Davis, J., Janowsky, J., Engel, M., & Nass, R. (1988). Drawing ability in four young children with congenital unilateral brain lesions. Neuropsychologia, 26(3), 359-371.

Streiner, D. L., & Norman, G. R. (1989). Health measurement scales. New York: Oxford University Press.

Strub, R. L., & Black, F. W. (1985). The mental status examination in neurology (2nd ed.). Philadelphia: F. A. Davis Co.

Sulkava, R., Wikstrom, J., Aromaa, A., Raitasalo, R., Lehtinen, V., Lahtela, K., & Palo, J. (1985). Prevalence of severe dementia in Finland. Neurology, 35, 1025-1029.

Sunderland, T., Hill, J. L., Mellow, A. M., Lawlor, B. A., Gundersheimer, J., Newhouse, P. A., & Grafman, J. H. (1989). Clock drawing in Alzheimer's disease. Journal of the American Geriatrics Society, 37, 727-729.

Tabachnick, B. G., & Fidell, L. S. (1989). Using Multivariate Statistics (2nd ed.). New York: Harper and Row.

Tariot, P. N., Sunderland, T., Weingartner, H., Murphy, D. L., Cohen, M. R., & Cohen, R. M. (1986). Naloxone and Alzheimer's disease. Archives of General Psychiatry, 43, 727-732.

Taylor, M. M. (1968). Analysis of dysfunction in the left hemiplegia following stroke. American Journal of Occupational Therapy, 22, 512-520.

Teng, E. L., & Chui, H. C. (1987). The Modified Mini-Mental State (3MS) Examination. Journal of Clinical Psychiatry, 48(8), 314-318.

Thal, L. J., Grundman, M., & Golden, R. (1986). Alzheimer's disease: A correlational anlaysis of the Blessed Information-Memory-Concentration Test and the Mini-Mental State Exam. Neurology, 36, 262-264.

Thorn, D. W., & Deitz, J. C. (1990). A content validity study of the Test of Orientation for rehabilitation patients. The Occupational Therapy Journal of Research, 10(1), 27-40.

Tolman, E. (1948). Cognitive maps in rats and men. Psychological Review, 55, 189-208.

Uhlmann, R. F., Larson, E. B., & Buchner, D. M. (1987). Correlations of Mini-Mental State and modified Dementia Rating Scale to measures of transitional health status in dementia. Journal of Gerontology, 42, 33-36.

Van Hoesen, G. W., & Damasio, A. R. (1987). Neural correlates of cognitive impairment in Alzheimer's disease. In F. Plum (Ed.), Handbook of physiology: Higher functions of the brain (Vol. 5, pp. 871-898). Washington: DC: American Physiological Society.

Vanier, M., Gauthier, L., Lambert, J., Pepin, E. P., Robillard, A., Dubouloz, C. J., Gagnon, R.,

& Joanette, Y. (1990). Evaluation of left visuospatial neglect: Norms and discrimination power of two tests. Neuropsychology, 4, 87-96.

Waber, D. P., & Holmes, J. M. (1986). Assessing children's memory productions of the Rey-Osterrieth Complex Figure. Journal of Clinical and Experimental Neuropsychology, 8, 565-580.

Wagner, O., Oesterreich, K., & Hoyer, S. (1985). Validity of the ischemic score in degenerative and vascular dementia and depression in old age. Archives of Gerontology and Geriatrics, 4, 333-345.

Watkins, P., Gouvier, D., Callon, E., & Barkemeyer, C. (1989). Equivalence of items on the Mini-Mental State. Archives of Clinical Neuropsychology, 4, 381-384.

Weintraub, S., & Kapust, L. R. (1989). Non-amnestic dementias: Special considerations for clinicians and caregivers. The American Journal of Alzheimer's Care and Related Disorders & Research, 4(3), 6-11.

Weisman, G. D. (1987). Improving way-finding and architectural legibility in housing for the elderly. In V. Regnier & J. Pynoos (Eds.). Housing the aged: Design directives and policy considerations. New York: Elsevier.

Weisman, J. (1981). Evaluating architectural legibility: Way-finding in the built environment. Environment and Behavior, 13(2), 189-204.

Whitkin, H. A. (1950). Individual differences in ease of perception of embedded figures. Journal of Personality, 19, 1-15.

Whitehouse, P. S., & Keller, K. J. (1987). Nicotinic and muscarinic cholinergic receptors in Alzheimer's disease and related disorders. In R. J. Wurtman, S. Corkin, & J. H. Growdon

(Eds.), Advances in basic research and therapies. Proceedings of the Fourth Meeting of the International Study Group on the Pharmacology of Memory Disorders Associated with Aging, (pp. 169-177). Cambridge, MA: Center for Brain Sciences and Metabolism Charitable Trust.

Williams, B. W., Mack, W., & Henderson, V. W. (1989). Boston Naming Test in Alzheimer's disease. Neuropsychologia, 27, 1073-1079.

Wilkins, A. (1986). Remembering to do things in the laboratory and everyday life. Acta Neurologica Scandanvica, 74 (suppl. 109), 109-112.

Wolf-Klein, G. P., Silverstone, F. A., Levy, A. P., Brod, M. S., & Breuer, J. Screening for Alzheimer's disease by clock drawing. Journal of the American Geriatrics Society, 37, 730-734.

Wood-Dauphinee, S. L., Opzoomer, A., Williams, J. I., Marchand, B., & Spitzer, W. (1988). Assessment of global function: The Reintegration to Normal Living Index. Archives of Physical Medicine and Rehabilitation, 69, 583-590.

Wood-Dauphinee, S. L., & Williams, J. I. (1987). Reintegration to normal living as proxy to quality of life. Journal of Chronic Diseases, 40, 491-499.

Zgola, J. M. (1987). Doing things. Baltimore: Johns Hopkins University Press.

Appendix A

List of Abbreviations

1. Activities of Daily Living	ADL
2. Alzheimer's disease	AD
3. Analysis of Variance	ANOVA
4. Central Executive System	CES
5. Computed Tomography	CT
6. Functional Spatial Abilities Questionnaire	FSAQ
7. Global Deterioration Scale	GDS
8. Intra-class Correlation Coefficient	ICC
9. Mini-Mental State Examination	MMSE
10. Modified Mini-Mental State Examination	3MS
11. National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association	NINCDS-ADRDA
12. Occupational Therapist	OT
13. Reintegration to Normal Living Index	RNL Index
14. Revised Diagnostic and Statistical Manual of Mental Disorders	DSM III-R
15. Spatial Orientation Skills	SOS
16. Visuospatial Scratch Pad	VSSP
17. Wechsler Adult Intelligence Scale	WAIS

Appendix B

General Information

Subject No.: ____________________ Date:______________________________

Last Name: ____________________ First Name:_________________________

Caregiver/Spouse:____

Address:____

____

Duration of Residence at this address:_____________________________________

Telephone No.:____

Sex: F ___ M ___

Marital Status: Single___ Married___ Common-law___ Divorced___ Widow___ Widower___

Date of Birth: D ___ M___ Y ___ Age: ______

Educational level:____

Other training:____

Past Medical History: ____

____

Current Medications:____

____

Global Deterioration Scale Level: ____

Hachinski Ischemic Scale Score: ____

MMSE: ___/30 3MS: ___/100 RNL: ___/100

Family History of AD: Y ___ N ___

Edinburgh Inventory administered: Y ___ N ___ Handedness: R ___ L ___

Language(s) spoken/written: English ___ French ___ Other ___

Corrective Lenses: Y ___ N ___

Referring Neurologist:____

Appendix C

Consent Form (for AD subjects)

McGill University
School of Physical and Occupational Therapy

I am informed that this is a research study undertaken by Lili Liu under the supervision of Louise Gauthier. I am also informed that the purpose of this study is to develop a battery for the assessment of spatial orientation skills in older persons.

I am aware that I will undergo a series of assessments during a period of one and a half hours, and that these assessments will take place at the School of Physical and Occupational Therapy, McGill University. I realize that a 10 minute rest period will be provided half way through the session. I have been told that no foreseeable risks are involved in the assessments.

I agree to return for two more sessions with a two week interval between each session.

I understand that my primary caregiver (spouse, next of kin, friend) will complete a questionnaire concerning my functional abilities.

I realize that although the results from the study will be published, my identity will be held in confidence. I am aware that my participation in this study is on a volunteer basis and that I will not be paid, however my transportation will be reimbursed if applicable.

I understand that I may withdraw my consent as well as withdraw from the study at any time without prejudice to my treatment.

________________________   __________________

Signature of Subject   Date

________________________   __________________

Signature of Caregiver   Date

I have explained to ____________________ the procedures of the study and I have informed him/her that he/she may withdraw from the study at any time.

________________________   __________________

Signature of Investigator   Date
(Tel: 398-4500)

Appendix C (Continued)

Consent Form (for Control subjects)

McGill University
School of Physical and Occupational Therapy

I am informed that this is a research study undertaken by Lili Liu under the supervision of Louise Gauthier. I am also informed that the purpose of this study is to develop a battery for the assessment of spatial orientation skills in older persons.

I am aware that I will undergo a series of assessments during a period of one and a half hours, and that these assessments will take place at the School of Physical and Occupational Therapy, McGill University. I realize that a 10 minute rest period will be provided half way through the session. I have been told that no foreseeable risks are involved in the assessments.

I realize that I may be selected to return for a second visit with a two week interval between the sessions.

I understand that my spouse, next of kin or friend will complete a questionnaire concerning my functional abilities.

I realize that although the results from the study will be published, my identity will be held in confidence. I am aware that my participation in this study is on a volunteer basis and that I will not be paid, however my transportation will be reimbursed if applicable.

I understand that I may withdraw my consent as well as withdraw from the study at any time without prejudice.

________________________ ____________________
Signature of Subject Date

________________________ ____________________
Signature of Witness Date

I have explained to ______________________ the procedures of the study and I have informed him/her that he/she may withdraw from the study at any time.

________________________ ____________________
Signature of Investigator Date
(Tel: 398-4500)

Appendix D

Functional Spatial Abilities Questionnaire (self-rated)

Please rate yourself on each item according to the following scale:

YES (Y)	NOT APPLICABLE (N/A)	NO (N)
(1)	(2)	(3)

	Y	N/A	N
1. I get lost in new or nonfamiliar environments when walking or driving.	1	2	3
2. I require supervision when travelling to a new environment.	1	2	3
3. I have difficulty following a map (ex. subway map, city map).	1	2	3
4. I am uncomfortable when travelling alone.	1	2	3
5. I have difficulty remembering the destination when I travel.	1	2	3
6. I have difficulty returning home after an outing (ex. take longer than is required, get off at wrong bus/subway make a wrong turn).	1	2	3
7. My sense of direction has changed over time.	1	2	3
8. I get lost in previously familiar environments (homes of relatives/friends, shopping center).	1	2	3
9. I require supervision when I travel in the neighbourhood.	1	2	3
10. I get lost in the home.	1	2	3
11. I am uncomfortable when I am alone at home.	1	2	3
12. I place objects in inappropriate locations in the home (ex. put kitchen item in bathroom).	1	2	3

Total: _____ (max = 36)

Appendix D (Continued)

Functional Spatial Abilities Questionnaire (proxy-rated)

You have been chosen to complete this questionnaire regarding: ____________________.
Please fill form without consultation from anyone and return it to the examiner in the envelope provided.
Please rate each item according to the following scale:

YES (Y)	NOT APPLICABLE (N/A)	NO (N)
(1)	(2)	(3)

	Y	N/A	N
1. This person gets lost in new or nonfamiliar environments when walking or driving.	1	2	3
2. This person requires supervision when travelling to a new environment.	1	2	3
3. This person has difficulty following a map (ex. subway map, city map).	1	2	3
4. This person is uncomfortable when travelling alone.	1	2	3
5. This person has difficulty remembering the destination when travelling.	1	2	3
6. This person has difficulty returning home after an outing (ex. takes much longer than is required, gets off at wrong bus/subway station, makes a wrong turn).	1	2	3
7. This person's sense of direction has changed over time.	1	2	3
8. This person gets lost in previously familiar environments (homes of relatives/friends, shopping center).	1	2	3
9. This person requires supervision when travelling in the neighbourhood.	1	2	3
10. This person gets lost in the home.	1	2	3
11. This person expresses discomfort when left alone at home.	1	2	3
12. This person places objects in inappropriate locations in the home (ex. put kitchen item in bathroom).	1	2	3

Total: _____ (max = 36)

Appendix E

Questionnaire sent to Caregivers

Dear Madam/Sir:

We are conducting a research project sponsored by the Alzheimer Society of Canada. We are interested in knowing the changes in people's orientation to place (knowing where one is and reaching a destination): for example, not being able to go to the store alone). Please describe ybehaviours or actions you have seen in persons who have difficulty with orientation to place. Thank you for your cooperation.

Sincerely yours,

Please send this questionnaire to:

Louise Gauthier
School of Physical and Occupational Therapy
3654 Drummond Street
Montréal (Québec) H3G 1Y5

Appendix E (Continued)

Questionnaire sent to Caregivers (French Version)

Chère Madame, Cher Monsieur,

Nous effectuons présentement un projet de recherche subventionné par la Société Alzheimer du Canada. Nous sommes intéressés à mieux connaître les changements qui se produisent chez une personne dans son orientation dans l'espace (savoir où on est et pouvoir atteindre une destination; par exemple, ne plus se rendre au dépanneur seul). Pourriez-vous décrire les comportements ou les actions que vous avez observés chez une personne qui a de la difficulté à s'orienter dans l'espace. Merci de votre collaboration.

Bien à vous,

____

____

____

____

____

____

____

____

____

____

S'il-vous-plaît, faire parvenir votre réponse à:

Louise Gauthier
School of Physical and Occupational Therapy
3654 Drummond Street
Montréal (Québec) H3G 1Y5

Appendix F

Table 28

Comparison of characteristics and mental status performance by the AD patients in stages 3 and 4 of the GDS

	AD (GDS 3)		AD (GDS 4)		
Variable	range	M (SD)	range	M (SD)	t* (df)
Age	51-78	64.5 (9.4)	56-76	68.5 (6.4)	-1.3 (23)
Education	5-16	11.1 (3.5)	6-35	11.9 (7.6)	-0.3 (17)
MMSE	15-29	23.1 (4.4)	6-25	17.9 (5.8)	**2.5*** (23)
3MS	61-93	74.3 (10.1)	28-72	56.6 (15.7)	**3.3**** (23)

Note. n = 10-12 for GDS 3 group and n = 9-13 for GDS 4 group.

*p<.05, **p<.01, two-tailed Student's t-test for independent samples.

Appendix G

Summary of Results

SUBTEST	SCORE	MAXIMUM SCORE
PERCEPTUAL SPATIAL SKILLS SUBTESTS		
1. Left-Right Discrimination	________	10
2. Figure-Ground Perception	________	48
3. Position in Space	________	16
4. Spatial Relations	________	15
COGNITIVE SPATIAL SKILLS SUBTESTS		
5. Porteus Maze Test	________	17 yrs
6. Map-Reading Test	________	35
7. Road-Map Test	________	32
8. Corsi Block Tapping Test	________	9
9. Modified Stylus Maze Test	________	10
10. Mental Representation (Sketch Map): Familiar Environment	________	1
11. Sketch Map: New Environment	________	1
FUNCTIONAL SPATIAL ABILITIES		
12. FSAQ: Self-Rated	________	36
13. FSAQ: Proxy-Rated	________	36

Appendix H

Content Validation Form for SOS Battery

1. Is the total group of skills tested (perceptual, cognitive and functional skills) appropriate to spatial orientation? Yes ___ No ____

2. Considered individually, is each skill included in the battery important to a person's spatial orientation? Yes ____ No ____

If not, which of the skills would you omit?

____

____

3. Are there sufficient skills included in the battery to provide a realistic assessment of a person's spatial orientation? Yes ____ No ____

If not, which additional skills should be included?

____

____

4. Would you agree that this battery will show differences in spatial skills between a group of control subjects and another with Alzheimer's disease?

Strongly Agree:____ Agree:____ Disagree:____ Strongly Disagree:____

5. Would you agree that this battery will show differences in spatial orientation between a group of control subjects and another with Alzheimer's disease?

Strongly Agree:____ Agree:____ Disagree:____ Strongly Disagree:____

6. Are there any other comments you would like to make? (Please use back of sheet)

____ ____________________________

Signature & title(s) Date

Appendix I

Table 29

Comparison of performance on the SOS Battery by the AD patients in stages 3 and 4 of the GDS

	AD (GDS 3)		AD (GDS 4)		
Variable	range	M (SD)	range	M (SD)	t* (df)
Left-Right	6-10	8.8 (1.4)	4-10	7.9 (1.8)	1.3 (23)
FGP	21-31	24.4 (3.3)	5-31	23.9 (7.5)	0.2 (12)
Spatial Relations	5-10	7.6 (1.8)	7-9	7.6 (0.8)	-0.1 (15)
Position in Space	7-16	12.4 (0.8)	4-15	9.6 (3.7)	**2.1*** (23)
Porteus Maze	5-17	8.2 (3.3)	5-17	7.8 (4.3)	-0.1 (22)
Map-Read	0-35	6.1 (10.2)	0-20	4.2 (5.9)	0.6 (22)
Road-Map	0-32	22.7 (6.0)	0-32	16.4 (6.4)	**2.5*** (22)
Corsi Block	0-5	3.1 (2.0)	0-4	2.5 (1.5)	0.8 (23)
Stylus Maze	3.3-10.0	5.7 (1.8)	3.0-7.1	4.9 (1.6)	1.2 (20)

(continued)

Appendix I (Continued)

Table 29

Comparison of performance on the SOS Battery by AD patients in stages 3 and 4 of the GDS

	AD (GDS 3)		AD (GDS 4)		
Variable	range	M (SD)	range	M (SD)	t* (df)
Sketch map: Familiar	0-1	.58 (.31)	0-1.0	.45 (.42)	0.9 (21)
Sketch map: New	0-1	.35 (.28)	0-.92	.22 (.31)	1.1 (22)
FSAQ: Self-rated	22-36	28.5 (4.9)	23-36	31.1 (3.5)	-1.5 (21)
FSAQ: Proxy-rated	22-34	26.9 (4.7)	14-34	24.6 (6.2)	1.0 (21)
Total Battery Score	35.4-61.8	51.2 (8.7)	26.9-73.3	46.2 (18.3)	0.7 (11)

Note. n = 10-12 for GDS 3 group and n = 9-13 for GDS 4 group.

*$p<.05$, **$p<.01$, two-tailed Student's t-test for independent samples.

Appendix J

Table 30

Frequencies of self-rated responses on items of FSAQ

	AD (GDS 3 & 4) n = 25	CONTROL n = 97	
Question	n(%)	n(%)	/Z/ value
(1) Y	11(48)	24(25)	
N/A	0(0)	0(0)	**2.2***
N	12(52)	73(75)	
(2) Y	11(48)	6(6)	
N/A	0(0)	0(0)	**5.1*****
N	12(52)	91(94)	
(3) Y	8(35)	6(6)	
N/A	3(13)	0(0)	**5.0*****
N	12(52)	91(94)	
(4) Y	7(30)	5(5)	
N/A	3(13)	0(0)	**4.9*****
N	13(57)	92(95)	
(5) Y	3(13)	0(0)	
N/A	0(0)	0(0)	**3.6*****
N	20(87)	97(100)	
(6) Y	2(9)	1(1)	
N/A	2(9)	1(1)	**3.0****
N	19(82)	95(97)	

(continued)

Appendix J (Continued)

Table 30

Frequencies of self-rated responses on items of FSAQ

	AD (GDS 3 & 4) n = 25	CONTROL n = 97	
Question	n(%)	n(%)	/Z/ value
(7) Y	11(48)	4(4)	
N/A	1(4)	0(0)	**6.0*****
N	11(48)	93(96)	
(8) Y	3(13)	1(1)	
N/A	1(4)	0(0)	**3.5****
N	19(83)	96(99)	
(9) Y	0(0)	0(0)	
N/A	0(0)	0(0)	0.0
N	23(100)	97(100)	
(10) Y	2(9)	0(0)	
N/A	0(0)	0(0)	**2.9****
N	21(91)	97(100)	
(11) Y	3(13)	0(0)	**2.9****
N/A	0(0)	1(1)	
N	20(87)	96(99)	
(12) Y	6(26)	1(1)	
N/A	0(0)	0(0)	**4.6*****
N	17(74)	96(99)	

Note. Percentage was calculated based on number of subjects who responded.

*p<.05, **p<.001, ***p<.0001, Wilcoxon-rank sum test.

Appendix K

Table 31

Frequencies of proxy-rated responses on items of FSAQ

	AD (GDS 3 & 4) n = 25	CONTROL n = 97	
Question	n(%)	n(%)	/Z/ value
(1) Y	16(67)	9(11)	
N/A	3(12)	2(2)	**6.4****
N	5(21)	75(87)	
(2) Y	16(67)	1(1)	
N/A	2(8)	1(1)	**8.2*****
N	6(25)	84(98)	
(3) Y	15(63)	6(7)	
N/A	3(12)	0(0)	**6.9*****
N	6(25)	79(93)	
(4) Y	8(35)	4(5)	
N/A	7(30)	1(1)	**6.4*****
N	8(35)	81(94)	
(5) Y	9(38)	1(1)	
N/A	3(12)	0(0)	**6.5*****
N	12(5)	85(99)	
(6) Y	9(38)	1(1)	
N/A	4(17)	1(1)	**6.5*****
N	11(46)	84(98)	

(continued)

Appendix K (Continued)

Table 31

Frequencies of proxy-rated responses on items of FSAQ

	AD (GDS 3 & 4) n = 25	CONTROL n = 97	
Question	n(%)	n(%)	/Z/ value
(7) Y	13(54)	2(2)	
N/A	3(13)	1(1)	**7.2*****
N	8(33)	83(97)	
(8) Y	6(25)	1(1)	
N/A	0(0)	0(0)	**4.2*****
N	18(75)	85(99)	
(9) Y	3(13)	0(0)	
N/A	1(4)	0(0)	**3.8*****
N	20(83)	86(100)	
(10) Y	2(8)	0(0)	
N/A	0(0)	0(0)	**2.7****
N	22(92)	85(100)	
(11) Y	3(13)	1(1)	
N/A	0(0)	1(1)	**2.1***
N	21(87)	84(98)	
(12) Y	10(42)	0(0)	
N/A	1(4)	1(1)	**6.2*****
N	13(54	85(99)	

Note. Percentage was calculated based on number of subjects who responded.
*p<.05, **p<.001, ***p<.0001, Wilcoxon-rank sum test.