1998

Neuropsychological subgroups of patients with Alzheimer's disease: A look at the first ten years of CERAD data.

Nancy Justina. Fisher

University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/2835

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeib Road, Ann Arbor, MI 48106-1346 USA
800-521-0800

UMI®
NEUROPSYCHOLOGICAL SUBGROUPS OF PATIENTS WITH ALZHEIMER'S DISEASE: A LOOK AT THE FIRST TEN YEARS OF CERAD DATA

by

NANCY JUSTINA FISHER

A Dissertation
Submitted to the Faculty of Graduate Studies and Research
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy at the
University of Windsor

Windsor, Ontario, Canada

©1998 Nancy J. Fisher
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-52421-3
ABSTRACT

Neuropsychological CERAD data from 960 patients diagnosed with probable Alzheimer’s disease (AD) and 465 non-demented controls were examined, to investigate whether previously identified subgroups of patients could be replicated (i.e., externally validated) in this sample, and if so, whether these subgroups were specific to AD. Patient and comparison group data were subjected to separate Q-factor and cluster analytic procedures. Consistent with past research, three qualitatively distinct patient subgroups were reliably identified across classification methods: Subgroup 1 (LAD; n = 312) was characterized by severe naming impairment yet borderline normal figure copying skills; Subgroup 2 (RAD; n = 247) displayed average naming ability within the context of moderately impaired copying performance; Subgroup 3 (GAD; n = 161) evinced profound anomia and constructional dyspraxia. Members of Subgroup 1 were older and less educated than those of the other two subgroups. Subgroup 2 was the highest functioning group in terms of functional levels and overall dementia severity; Subgroup 3 was most impaired in these respects. Analyses of the control data revealed two clusters of individuals that did not resemble the patient subgroups, suggesting that the above 3-subgroup classification is unique to AD. Longitudinal patient data from 7 follow up examinations were also examined. Initial patterns of performance were generally stable across time for the LAD and GAD subgroups. Such data were less consistent for the RAD group, suggesting additional, and as yet unexplained, heterogeneity within this subgroup. Cross-sectional analyses revealed the presence of members from all three subgroups across mild, moderate and severe stages of the disease, regardless of the stratification method employed (i.e., CDR or MMSE).

-iiri-
For the LAD and GAD groups, longitudinal and cross-sectional analyses involving naming/copying discrepancy scores revealed findings in accord with those predicted by an earlier proposed subgroup-specific progression model. Although the results of such analyses for the RAD group did not support the model when the group data were averaged, it should be noted that a subset of RAD members did perform in a manner consistent with the model. The results of this investigation support an integration of the Subgroup and Stage models in conceptualizing this most complex disorder.
ACKNOWLEDGEMENTS

First I would like to thank my committee members for their individual contributions to this dissertation: Byron Rourke, my Mentor, for leading me away from the darkness with one hand, while with the other assuring me the freedom, independence, and space I needed to grow toward the light; Linas Bieliauskas, for his always encouragingly constructive remarks; Doug Shore, for his insights and down-to-earth comments; Dr. Holosko for the thorough, multi-level feedback he offered at all stages of this project.

I am honoured that, Dr. Alex Martin, my all-time idol, agreed to fulfill the role of external examiner for this dissertation. I thank Dr. Martin not only for his landmark inspirational works, but for the gracious manner in which he has always provided me with encouragement and motivation. There could be no other external for this dissertation.

I wish to acknowledge Albert Heyman, M.D., the principal investigator of CERAD, for freely expressing his pleasure with this project, and promoting the presentation and publication of the results reported herein. Special thanks to the CERAD Steering Committee members—Drs. Albert Heyman, Richard Mohs, Gerda Fillenbaum, John Morris, Kathleen Welsh-Bohmer, Suzanne Mirra, and Gerald van Belle, for approving this project back in October of 1995.

Lastly, I would like to thank my dear friend John DeLuca, who lived through much of the dissertation process with me on a day-to-day basis. I thank him for providing me access to computer hardware/software resources and offering his statistical/methodological opinions. Moreover, the multiple occasions on which he offered theoretical insight and practical advice on surviving the dissertation process were much appreciated.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS AND ACRONYMS</td>
<td>xiii</td>
</tr>
</tbody>
</table>

## CHAPTER

### I. INTRODUCTION

Clinical Heterogeneity of Alzheimer’s Disease: Subgroup versus Stage Models

Unravelling the Heterogeneity: Neuropsychological Profile Analysis

Scope of Literature Review

### II. REVIEW OF THE LITERATURE

Neuropsychological Subgrouping Studies

--Landmark Contributions of Martin et al. (1986)

--Other Neuropsychological Right-Left Subgrouping Studies

--Neuropsychological Subgrouping Studies of Normal Elderly

--Integrative Summary
Toward a Neuropsychological Model of AD

--Goldberg-Costa Model

--White Matter/NLD Model

--Minimal Progression Model

--An Integration of Theoretical Applications: The Subgroup-Specific Progression Model

Where do we go from here?

On CERAD

CERAD Neuropsychological Studies

Purpose of the Present Study

Hypotheses

III. METHODOLOGY

Subjects

The CERAD Assessment Battery

--Clinical Assessment

--Neuropsychological Measures

--Psychometric Properties of the CERAD Neuropsychological Battery

On Cluster Analysis
Clustering Methods Utilized in Data Analysis

--Q-Type Factor Analysis

--Hierarchical Agglomerative Methods

--K-Means Iterative Partitioning

Overview of Data Analysis Process

IV. RESULTS

Preliminary Examination of the Data

R-Factor Analyses

Q-Factor Analyses

Cluster Analyses

Evaluation of Clustering Results

Inter-Subgroup Comparisons

Longitudinal Analyses of Patient Data

Discrepancy Score Changes as a Function of Dementia Severity/Stage

V. DISCUSSION

Summary of Results

--Subtyping Results

--Inter-Subgroup Comparisons

--Longitudinal Results

--CDR Stage Findings

-viii-
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation of Hypotheses</td>
<td>110</td>
</tr>
<tr>
<td>Integration with the Literature</td>
<td>117</td>
</tr>
<tr>
<td>Limitations of Present Study</td>
<td>127</td>
</tr>
<tr>
<td>Directions for Future Research</td>
<td>128</td>
</tr>
<tr>
<td>Concluding Comments: Significance and Implications of the Current Findings</td>
<td>130</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>133</td>
</tr>
<tr>
<td>APPENDIX A -- Participating CERAD Sites and Principal Investigators</td>
<td>150</td>
</tr>
<tr>
<td>APPENDIX B -- Description of Tests Comprising the CERAD Clinical and Neuropsychological Battery</td>
<td>152</td>
</tr>
<tr>
<td>APPENDIX C -- Glossary of Terms</td>
<td>158</td>
</tr>
<tr>
<td>VITA AUCTORIS</td>
<td>160</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Descriptive Statistics for Demographic and Neuropsychological Data
2. Principal Component R-Type Factor Loadings for the Patients
3. Principal Component R-Type Factor Loadings for the Controls
4. Results of Q-Type Factor Analysis—Patients
5. Results of Q-Type Factor Analysis—Controls
6. Misclassification Analysis: (a) Patients (b) Controls
7. Multiprofile-Multimethod Matrix: Patients
8. Multiprofile-Multimethod Matrix: Controls
9. Means and Standard Deviations for the Q-Factor Patient Subgroups on Measures not Employed in their Derivation
10. Mean BNT/PRAX Discrepancy T Scores across Assessments
11. Comparison of Incidence Rates of the Three Subgroups across Studies
LIST OF FIGURES

1 Comparison of current sample to previous samples on Visual Construction and Semantic Memory composites

2 Comparison of current and past samples on demographic and symptom severity indicators

3 R-Factor plot of patient $T$ score variables in rotated factor space

4 R-Factor plot of control $T$ score variables in rotated factor space

5 Mean performance profiles of patient subgroups derived by Q-Factor analysis

6 Mean performance profiles of normal control subgroups derived by Q-Factor analysis

7 Mean performance profiles of normal control subgroups (all included [Q1, Q2]; reduced sample with those having MMSE scores $< 25$ excluded [RQ1; RQ2]) derived by Q-Factor analysis

8 Comparative figures depicting mean patient subgroup profiles generated by each clustering method

9 Comparative figures depicting mean control subgroup profiles generated by each clustering method

10 Histograms depicting BNT/PRAX $T$ score discrepancies across time for 5 randomly selected Q1 patients

11 Histograms depicting BNT/PRAX $T$ score discrepancies across time for 5 randomly selected Q2 patients

12 Histograms depicting BNT/PRAX $T$ score discrepancies across time for 5 randomly selected Q3 patients

13 BNT–PRAX discrepancy $T$ scores for the patients across time
14 Mean BNT and PRAX T scores across testings for the Q1 group
15 Mean BNT and PRAX T scores across testings for the Q3 group
16 Mean BNT and PRAX T scores across testings for the Q2 group
17 Plots of initial BNT/PRAX mean discrepancy scores by CDR stage for each subgroup
ABBREVIATIONS & ACRONYMS

AD ................................................................. Alzheimer’s Disease
ADRDA ......................................................... Alzheimer’s Disease and Related Disorders Association
ANOVA .......................................................... Analysis of Variance
BNT ................................................................. Boston Naming Test
CERAD ......................................................... Consortium for the Establishment of a Registry for Alzheimer’s Disease
DVBM .......................................................... Delayed Verbal Memory
GAD ............................................................... Global Alzheimer’s Disease
IVBM ............................................................. Immediate Verbal Memory
LAD ............................................................... Left Hemisphere Alzheimer’s Disease
MMSE ............................................................ Mini Mental State Exam
NINCDS ......................................................... National Institute of Neurological and Communicative Disorders and Stroke
NLD ............................................................. Nonverbal Learning Disability
PRAX ............................................................ Constructional Praxis (Figure Copying Task)
RAD .............................................................. Right Hemisphere Alzheimer’s Disease
RECOG .......................................................... Recognition Memory
VF ................................................................. Verbal Fluency for Animal Names
WMS ............................................................. Wechsler Memory Scale
CHAPTER I

Introduction

The research project outlined herein involved the utilization of factor and cluster analytic techniques in a large scale examination of the neuropsychological profiles of patients with Alzheimer's disease (AD). This investigation was conducted with the goal of replicating the results of an earlier study, in which three neuropsychological subgroups of AD patients were reliably identified across clustering methods (Fisher et al., 1996). The current study employed methodological improvements over the previous one, by inclusion of a demographically similar healthy (i.e., non-demented) comparison group, utilization of a larger, multicenter sample, and incorporation of a longitudinal design. The longitudinal design allowed for evaluation and refinement of an earlier developed subgroup-specific neuropsychological model of AD (Fisher, Rourke, & Bielialuskas, 1997). This model incorporates both the Subgroup and Stage Models of the disease, by outlining progression patterns for each of three neuropsychological subgroups.

Clinical Heterogeneity of Alzheimer's Disease: Subgroup versus Stage Models

Over the past decade or so, researchers from several fields (i.e., neurology, neuroimaging, neuroepidemiology, neuropathology, neuropsychology) have become increasingly aware of the heterogeneous nature of AD (for review, see Fisher, Rourke, & Bielialuskas, 1997). As the present state of affairs remains one in which the antemortem diagnosis of AD is based upon clinical history and presentation (McKann et al., 1984), it follows that attempts at investigating the parameters of this heterogeneity begin on clinical grounds.
There is agreement in the clinical literature that patients with AD vary greatly in regard to their clinical presentation (Becker, Huff, Nebes, Holland, & Boller, 1988; Jagust, Davies, Tiller-Borcich, & Reed, 1990; Joanette, Ska, Poissant, & Beland, 1992; Martin, 1990; Martin, Cox, Brouwers, & Fedio, 1985; Neary et al., 1986; Price et al., 1993; Schwartz, 1987; Shuttleworth, 1984; Swash et al., 1991; Wishart & Henderson, 1998). In this manner, although memory deficits can almost always be shown to be present if adequate neuropsychological testing is undertaken, great inter-individual variability is apparent with respect to other areas of cognitive impairment, in addition to psychiatric and strictly neurological symptomatology.

Two main models which attempt to explain the heterogeneity of symptom presentation among AD patients are the Subgroup and Stage Models. Proponents of the Subgroup Model of AD contend that patients vary with regard to their evolutionary pattern of neuropsychological/psychiatric/neurological degeneration. As such, individuals at comparable phases of the illness exhibit distinct patterns of spared and affected cognitive and functional abilities, and also, varying degrees of impairment of these capacities. Furthermore, not only do the neuropsychological functions affected differ among individuals in the early and middle stages of AD, but the patterns by which the subcomponents of these functions breakdown also show inter-individual variability (Joanette et al., 1992; Jorm, 1985; Martin et al., 1986). The goal of research conducted by those adhering to this conceptual framework of AD is to identify subgroups of patients with similar patterns of degeneration. Such identification would allow for study of relatively similar cohorts, which may have etiologic, prognostic, and/or therapeutic importance (Joanette et al., 1992; Martin, 1990). As well, the study of various
patterns of neuropsychological degeneration is important in advancing an understanding of
the organization of functional neural systems (e.g., the organization of semantic knowledge).
Identification of double dissociations of function, and subsequent study of the ways in which
these independent systems breakdown, allows for increased awareness of the functional
organization of the brain (Martin, 1990).

In contrast to the view of individuals at comparable phases of AD as heterogeneous
in symptom presentation, opponents argue that deficits incurred by AD victims are
homogeneous and, relatively equal across cognitive domains (Constantinidis, 1978; Hom,
1992; Reisberg, Ferris, & Crook, 1982). In this way, proponents of the classical Stage Model
contend that there is a single global pattern of cognitive impairment characterizing AD
patients (Hom, 1992), and that any differences observed between patients merely reflect the
distinct stages of the disease and hence, the severity and/or duration of the disorder
(Constantinidis, 1978; Reisberg, Ferris, & Crook, 1982). Various Stage Models have been
presented in the literature (e.g., Cummings & Benson, 1992), each assuming a more or less
global, homogeneous pattern of cognitive deterioration, which increases as a function of
disease progression. Qualitatively distinct symptoms among AD victims are recognized by
proponents of these models, but are attributed to specific stages of disease progression (e.g.,
see Reisberg et al., 1982), and as such, are thought to adhere to a timetable.

While contributory in terms of understanding the progressive nature of AD, there are
limitations of the strict Stage Model conceptualization of the disease. First of all, the fact that
several different Stage Models are noted in the literature, some of which clash markedly in
terms of the temporal sequence in which certain symptoms are scheduled to first appear (e.g.,
personality disturbance; Martin et al., 1986), raises concerns regarding the validity of this approach (Liston, 1979; Schwartz, 1987) as a lone explanatory model. Furthermore, there does not appear to be a consensus as to the appropriate number of stages of AD, or the approximate duration of each of the different stages (c.f. Schwartz, 1987).

Another limitation is the failure of this model to account for autopsy confirmed case examples in the literature, which contradict the assumption of homogeneous dissolution of memorial, visual-spatial, and linguistic functions. For example, AD may initially present as an isolated memory impairment, in the face of otherwise normal neuropsychological functioning (Haxby et al., 1988; Neary et al., 1986; Price et al., 1993). Furthermore, individuals have been identified who exhibit severe impairment on neuropsychological measures of visual-construction, while maintaining relatively normal levels of performance on tests of word generation and naming, and vice versa (Martin, 1990). Other contradictory autopsy confirmed examples include AD initially presenting most prominently as a slowly progressive attentional deficit (Price et al., 1993), fluent aphasia (Pogacar & Williams, 1984), and parietal lobe syndrome (Crystal, Horoupian, Katzman, & Jotkowitz, 1982). Further, not only do certain AD cases initially present with "focal" impairment, but generally speaking, the most prominent initial area of deficit remains salient during the course of the disease; preserved areas remain relatively less affected, until the terminal stages are reached, at which time all areas of neuropsychological functioning become disrupted profoundly (Price et al., 1993).

A final limitation of the Stage Model is its adherence to a prescribed time framework by which the appearance of certain qualitatively distinct symptoms are supposed to appear, as several autopsy confirmed cases of AD do not follow such a linear sequence. For example,
motor deficits may appear early in the course of AD (Funkenstein et al., 1993; Jagust et al., 1990), but most Stage Models do not ascribe such impairment until the final stages of the disease. Similarly, personality, and/or affective disturbance may present as an initial symptom in some patients, while in others it does not occur until the later stages of AD (Adams & Victor, 1993).

As indicated above, the assertion that the deficits of AD progress in a parallel, and/or single predictable sequence for all so afflicted is tenuous on several grounds; contradictory evidence is readily available in the literature, and has been for at least the past decade. There appears to be a tendency for proponents of the stage approach to dismiss as “atypical” the cases mentioned above which contradict their assumptions. However, since these cases were autopsy confirmed as AD and only AD, such a dismissal is difficult to justify. Incorporation of a Subgroup Model in conceptualizing AD would appear necessary in accounting for the case history literature to date. In this manner, it would seem that a theoretical meshing of the Stage and Subgroup Models, in which the latter refines the former, would lead to a more advanced understanding of AD.

Unravelling the Heterogeneity: Neuropsychological Profile Analysis

As previously mentioned, because the diagnosis of AD remains one largely based upon clinical symptomatology and history, it would appear to make sense that in beginning the search for consistencies within the overlying heterogeneity, one look to measurable differences in clinical presentation. Neuropsychological measures would seem appropriate instruments to employ in this regard, given their objective nature. To date, research originating from this field has been both reliable and promising. Such studies have consistently
reported three qualitatively distinct clinical subgroups of AD patients based on patterns of performance on neuropsychological measures (Fisher et al., 1996; Martin et al., 1986; Strite, Massman, Cooke, & Doody, 1997). This research has suggested that while the majority of AD patients appear to undergo global decline in cognitive functioning, in a manner consistent with the Stage Model approach, small yet significant numbers of patients demonstrate relatively preserved functioning in one neuropsychological domain (either visual-constructional functioning or semantic accessibility) during the early and middle periods of the disease. Thus, in addition to the case history literature reviewed above, the results of neuropsychological subgrouping studies also suggest a blending of the Stage and Subgroup Models in conceptualizing the complex nature of AD.

Scope of Literature Review

Before proceeding to outline the rationale, hypotheses, and methodology of the current investigation, a literature review is in order, and will be presented in the chapter which follows. In this review, prior studies investigating neuropsychological profiles of AD patients and the application of pre-existing explanatory models arising from these studies will be presented, in essence providing the background and rationale for the current investigation. The chapter concludes with a summary and critique of the neuropsychological subgrouping literature to date, followed by a discussion of the specific research questions and hypotheses suggested by the review and examined in this dissertation.
CHAPTER II

Review of the Literature

Neuropsychological Subgrouping Studies

In this section, empirical investigations suggesting a right-left factor contributing to the neuropsychological heterogeneity of AD are reviewed. Neuropsychological subgrouping studies focusing on heterogeneous breakdown of the component processes within particular domains of functioning (e.g., memory, spatial processing, attention, etc.) are not included. Although such studies are important, they are beyond the scope of this review and have been covered sufficiently elsewhere (Fisher, 1995; Fisher, Rourke, & Bielaiskas, 1997). Following the review of AD subgrouping studies, subgrouping studies of normal elderly are presented.

Landmark Contributions of Martin et al. (1986)

Martin and colleagues (1986) administered a wide range of neuropsychological measures to 42 early probable AD patients (with a mean symptom duration of 3 years). Particular attention was paid to the performance of these subjects on measures assessing accessibility of semantic knowledge [i.e., Boston Naming Test (BNT), Verbal Fluency subtest of Mattis Dementia Rating Scale, Associate Learning subtest of Wechsler Memory Scale (WMS)-easy items only] and visual-spatial skills [i.e., Block Design subtest of the Wechsler Adult Intelligence Scale (WAIS), Rey-Osterrieth Complex Figure-copy only, Mosaic Comparison Test]. Upon qualitative inspection of these data, the investigators observed that while the majority of patients appeared to exhibit deficits in all domains assessed, there were two distinct subgroups of patients deviating from this global pattern. That is, some patients demonstrated isolated preservation of visual-constructional skills, while another small group
of patients displayed spared word-finding ability in the context of otherwise disrupted cognitive functioning.

The researchers next decided to investigate these qualitatively distinct neuropsychological patterns statistically, employing cluster analytic methodology (i.e., Ward’s method). This resulted in the identification of three subgroups of patients: (a) those with relatively equivalent impairment of word-finding and visual-constructional abilities; (b) those with impaired access to semantic knowledge coupled with relatively spared visual-constructional functioning; and, (c) those with relatively intact access to semantic knowledge accompanied by visual-spatial/constructional impairment. The subgroups did not differ with regard to age, education, or duration of illness (Martin et al., 1986; Martin, 1990). Significance tests evaluating inter-subgroup differences in overall dementia severity were not conducted.

Martin et al. (1986) proceeded to examine Positron Emission Tomography (PET) scans of members from each of the 3 subgroups and found them to be consistent with the neuropsychological findings. That is, those with impairment in both neuropsychological domains showed bilateral hypometabolism of the temporal and parietal lobes; those with impaired access to semantic knowledge, yet relatively spared spatial abilities, exhibited greater hypometabolism in the left temporal region relative to other cortical regions; and those with the reverse cognitive pattern exhibited greater hypometabolism in the right temporal and parietal regions.

The patients involved in this research were followed and re-tested on the six measures of verbal and visual-spatial functioning at one- to two-year intervals after the initial
assessment in 1986. Martin (1990) reported that the re-evaluations indicated distinct patterns of deterioration based on subgroup membership. Specifically, those in Subgroup 1, who initially displayed equal impairment in both domains, continued with this general global decline in functioning. However, patients in Subgroups 2 and 3 generally exhibited significantly greater deterioration in the area (i.e., either verbal or visual-spatial) which was impaired at the initial evaluation. These results suggest that patterns of deterioration may be predicted with knowledge of initial subgroup membership. Subsequent neuropathological examination of the brains of those from the subgroups who died shortly after the 1986 study confirmed the diagnoses of AD in the majority of cases (Martin, 1990).

The research of Martin et al. (1986) demonstrated a double dissociation of function between semantic knowledge and visual-constructional systems among AD patients. The subsequent demonstration that corresponding asymmetrical metabolic dysfunction was associated with these subgroupings strongly suggested a right-left factor contributing to heterogeneous neuropsychological presentations. This implied that while the majority of AD patients appear to undergo a more or less global loss of neuropsychological functions, presumably reflecting underlying symmetrical neuropathology, subgroups of patients with pronounced deficits in one area of functioning and relatively spared abilities in the other could be identified, and these subgroups reflect asymmetrical neuropathology. The implications of this finding are far-reaching, not only clinically in terms of compensatory strategies of adaptation, but also with respect to fundamental scientific issues, such as research design.

**Other Neuropsychological Right-Left Subgrouping Studies**

Becker et al. (1988) partially replicated the study of Martin et al. (1986) utilizing a
larger sample of 86 patients. The measures employed were similar to those of Martin et al. (1986) (e.g., Benton’s Three-Dimensional Block Design and Visual Form Discrimination tests, copy of Rey-Osterrieth complex figure, word generation, modified BNT, easy verbal paired associates). These researchers identified small subgroups similar to groups 2 and 3 of Martin et al., by calculating composite scores for the two neuropsychological domains (e.g., semantic knowledge accessibility and visual-constructional functioning), and utilizing the following criteria: (a) one composite score must be within the normal range; (b) the other score must fall at least 2 standard deviations from it. The remaining patients were classified into Martin’s global subgroup. These researchers reported a trend for those with lexical semantic impairments yet spared visual-constructional functioning to be older than those in the other subgroups, although this finding did not reach statistical significance. Duration of illness did not differ between the 3 groups.

Although recognizing the work of Martin et al. (1986) as ground-breaking, Fisher and colleagues (1996) noted some methodological limitations with this research. Most notably, the Martin et al. (1986) study did not involve replication of the clustering results across different algorithms, approaches, or similarity measures. This was viewed as problematic, in that lone cluster analytic solutions have often been shown to be unreliable (Blashfield & Aldenderfer, 1988; Everitt, 1974; Morris, Blashfield, & Satz, 1981). As well, the sample studied by Martin and colleagues was small ($N = 42$) for a cluster analytic study, and the participants were preselected to include those with focal patterns. Thus, Fisher et al. (1996) sought to investigate whether subgroups of AD patients with qualitatively distinct neuropsychological profiles could be reliably (i.e., across methods) identified, and as such, set
out to replicate the results of the Martin et al. (1986) study utilizing data collected from 134 probable AD patients, selected only in terms of satisfaction of criteria (McKhann et al., 1984) allowing for maximal antemortem diagnostic accuracy. The measures employed were similar to those of Martin et al. (1986) [i.e., BNT, Controlled Oral Word Association, Verbal Fluency-Animals, Easy Associates (WMS), Block Design, Figure-Ground, Pentagon Copy task].

Fisher et al. (1996) hypothesized that 3 subgroups similar to those identified by Martin et al. (1986) would emerge in their sample, hypothetically labeling them GAD (Global AD), RAD (Right hemisphere AD), and LAD (Left hemisphere AD). Formulation of this hypothetical classification enabled the prediction of subgroup differences (Morris & Fletcher, 1988). Drawing on the Goldberg-Costa (1981) model of hemispheric specialization and the Rourke (1982) White Matter Model, it was theorized that: (a) LAD disrupts intramodal left-hemisphere gray matter systems housing overlearned knowledge representations; (b) RAD disrupts white matter interregional integration systems, which are more prominent in the right hemisphere; while (c) GAD disrupts both types of cerebral systems (Fisher, Rourke, & Bieliauskas, 1997).

As predicted, 3 neuropsychological subgroups resembling the ideal types GAD, LAD, and RAD were reliably identified across multiple methods in the Fisher et al. (1996) study. Subgroup 1 (GAD), the largest of the groups, was marked by moderate to severe anomia and constructional dyspraxia. Individuals in Subgroup 2 (LAD) displayed relatively spared visual-perceptual/constructional functioning but severe anomia. Members of Subgroup 3 (RAD) exhibited a pattern of intact naming, borderline-normal block construction ability, and
moderate difficulty in copying overlapping figures. Thus, the Fisher et al. (1996) study successfully replicated the work of Martin et al. (1986) with a larger, more representative sample. Consistent with the findings of Martin et al. (1986) and Becker et al. (1988), Fisher et al. (1996) reported no differences between the subgroups with respect to age or duration of illness. Additionally, Fisher et al. found no difference in educational level between the three subgroups. However, lower Mini Mental State Examination (MMSE) scores were obtained by the GAD subgroup, compared to the LAD and RAD groups.

In a follow-up case study series, Fisher, Rourke, Bieliauskas, et al. (1997) provided detailed descriptions of both randomly and hand selected cases from the three subgroups identified in the 1996 study. Data from multiple testings were available for several of these patients, allowing longitudinal analysis. Of note, for all three subgroups (i.e., GAD, RAD, & LAD), the pattern of impairment revealed at the initial examination remained consistent across time. That is, for GAD, progression of the disease resulted in continued homogeneous decline in all cognitive areas. Similarly, for RAD and LAD, the initially preserved cognitive domain remained relatively less affected at follow-up examinations. These findings of profile stability are consistent with those reported by Martin (1990).

Strite et al. (1997) as Fisher et al. (1996), also attempted to replicate the results of Martin et al. (1986), employing the following measures: WAIS-R Comprehension, Vocabulary, Object Assembly, and Block Design subtests; BNT; Verbal Fluency; WMS-R Visual Reproduction I. These researchers initially attempted to utilize the same clustering methodology employed in the original study of Martin and colleagues. As such, they employed Ward's method alone in cluster analyzing R-Factor scores of AD patients on verbal
and visual measures. Because Martin et al. (1986) reported 5 clusters of patients (akin to RAD, LAD, and three severity levels of GAD), Strite et al. forced the Ward’s analysis to classify their patient sample into 5 subgroups. These researchers reported that many patients classified to the “asymmetric” groups had comparable scores in the 2 domains of functioning, suggesting that they should have been placed into one of the symmetric subgroups. Of note, these researchers may have been unsuccessful because they forced a 5-cluster solution. Had they forced a 3 group solution (or not forced a solution at all) they may have met with more success. Utilization of an initial clustering strategy controlling for outliers (e.g., Q-Factor analysis) may also have had an impact on the success of their classificatory attempt. It is unclear why these strategies were not pursued, as the Strite et al. sample was more similar (in composition) to that of Fisher et al. (1996) than to that of Martin et al. (1986).

Nevertheless, while giving up perhaps too easily on cluster analytic methodology, Strite et al. (1997) subsequently decided to pursue an alternative subgrouping method. This approach involved the calculation of discrepancy index (DI) scores (verbal factor score minus visual-spatial factor score). Those patients with a DI greater than +1 were classified into the asymmetric spatial group (RAD-like) group, and those with DIs of less than -1 comprised the asymmetric verbal (LAD-like) group. The researchers then classified the remaining (and majority of) patients (i.e., “symmetric” patients) into 3 severity groupings (high, low, midlevel) on the basis of performance on the “global impairment index” (i.e., mean of the two factor scores—visual and verbal). The performance patterns of the 3 qualitatively distinct groups produced by this method highly resembled those of Martin et al. (1986) and Fisher et al. (1996). Consistent with the findings of the aforementioned studies, there were no
differences between the subgroups with respect to age or educational level, and the RAD- and LAD-like groups did not differ significantly from each other in terms of MMSE scores.

Strite et al. (1997) also demonstrated, in a cross-sectional fashion, that members of the spatial and global subgroups were identifiable across disease severity stratification levels (i.e., mild, moderate, and severe stages of dementia according to MMSE scores). However, all members of the language-impaired (i.e., LAD-like) subgroup were classified as moderately demented (i.e., none fell within the mildly or severely demented categories).

While a handful of other right-left subgrouping studies have been conducted, such efforts were less stringently designed in terms of subject selection criteria and classification methodology (e.g., Naugle, Cullum, Bigler, & Massman, 1985; Rasmusson & Brandt, 1995), and hence are not reviewed here. However, it should be mentioned that the results of such studies are not inconsistent with those reviewed above. Rasmusson and Brandt (1995) reported less subgroup-profile stability over a one year period than that suggested by Martin (1990) and Fisher et al. (1997). However, their methods of deriving subgroups (e.g., without correcting scores for age and educational differences) and atheoretical approach to measuring instability render their results inconclusive.

Of note, Massman and Doody (1996) reported a high proportion (63%) of atypical finger tapping test performance among their right-handed probable AD sample. Twenty-six percent of their sample demonstrated an exaggerated right hand tapping advantage, while 37% displayed a reversal of the expected asymmetry (i.e., left hand performance equal to that of the right hand). Moreover, this atypical motor asymmetry was associated with cognitive asymmetries: that is, those with exaggerated right hand performances performed better on the
BNT than they did on a figure copying task (WMS-R figures), while those with exaggerated left hand performances showed the reverse pattern. Patients with normal finger tapping performances performed equally well on the BNT and figure copying task. Similarly, the mean VIQ minus PIQ difference score was significantly higher for the exaggerated right group than that for the reversal group; the normal group’s score fell between that of the other two groups. Of note, those with exaggerated right hand tapping performances had higher premorbid verbal abilities.

Neuropsychological Subgrouping Studies of Normal Elderly

A reasonable argument could be put forth that the heterogeneous neuropsychological manifestations of AD simply reflect age independent inter-individual differences in brain organization (e.g., gender, congenital, genetic differences), differential changes in brain organization as a function of age and environmental exposure, and/or the heterogeneity of the normal aging process itself (i.e., on a neurodegenerative level), all and any of which may leave individuals differentially vulnerable to losses in certain areas, resulting in distinct patterns of breakdown corresponding to these differences (Joanette et al., 1992; Valdois, Joanette, Poissant, Ska, & Dehaut, 1990). It is quite possible that the cognitive decline associated with normal aging may not proceed in a homogeneous fashion. Thus, neuropsychologically identified subgroups of AD patients may merely reflect differences in premorbid levels of functioning (i.e., premorbid strengths and weaknesses). As such, distinct weak areas of neuropsychological functioning before AD strikes may become more pronounced when the disease begins. Alternately, areas of strength may be relatively resilient to the initial impact of the disease. If this explanation alone were true, the neuropsychological heterogeneity of
AD would not be attributable to underlying asymmetrical variants of the disease, but to a single generalized process which simply serves to exacerbate premorbid individual differences.

Valdois et al. (1990) investigated the above potentially confounding factor to AD subgrouping research in their study of neuropsychological data collected from "normal" elderly individuals. Cluster analysis of scores from 10 neuropsychological measures [Hooper, Visual Form Discrimination, Line Orientation, Logical Memory (delayed recall), Visual Reproduction (delayed recall), Word List Repetition (immediate recall), Rey Figure-Copy, WAIS Block Design, Verbal Fluency, Oral Comprehension] yielded 6 subgroups. Three of these groups were most differentiated by level of performance (i.e., quantitative) differences within the normal range of functioning. Overall, these groups did not display striking patterns of strengths and weaknesses: Subgroup 1 exhibited a strength on Verbal Fluency; subgroups 2 and 3 displayed no strengths or weaknesses, although the lowest level of performance for Subgroup 2 was exhibited on the Verbal Fluency measure, which may have contributed to the differentiation of this group from Subgroup 1.

The other three subgroups identified by Valdois et al. (1990) displayed profiles of spared and impaired neuropsychological functioning, that did not appear to represent exacerbations of the patterns observed among the normally functioning subgroups. One of these groups performed within normal limits on all measures save for the verbal memory (i.e., Logical Memory) task. The remaining two subgroups were comprised of clearly impaired subjects (i.e., in terms mean z-scores across all measures). The first of these subgroups was impaired on all measures save for Line Orientation, Word List Repetition, Verbal Fluency, and Oral Comprehension. Compared to their other performances, this group showed a distinct
weakness on the Rey copy task, and a strength on Oral Comprehension. The other group was impaired on all measures save for Visual Form Discrimination, Logical Memory, and Verbal Fluency. This group demonstrated a strength on the Visual Form Discrimination test and a weakness on Oral Comprehension. These groups bore some resemblance to, yet did not overlap with those identified by Martin et al. (1986), although differences between the measures employed across the studies make such comparisons difficult.

Of note, Valdois et al. (1990) failed to elaborate the extent of their exclusionary screening process beyond simply stating that the subjects were “without obvious clinical signs of brain damage or psychiatric disorders” (p. 589). Given these weak inclusionary criteria, it is likely that early AD patients with subtle signs were not screened out. In spite of this limitation, the Valdois et al. study is valuable as a caveat to overinterpretation of subgrouping study results. Moreover, the results of this study suggest that patterns observed among impaired (yet undiagnosed) elderly do not correspond to those who exhibit normal range performances. More stringently designed studies of normal elderly individuals (i.e., utilizing comprehensive screening criteria and similar measures/methodologies to those used in AD subgrouping research) would be helpful in shedding greater light on this issue.

In a more recent investigation of cognitive heterogeneity in the normal aging population, Mitrushina, Uchiyama and Satz (1995) studied a community sample of 156 normal elderly individuals. The participants in this study were more stringently screened than in the Valdois et al. (1990) study. Only those with MMSE scores ≥ 24 and without psychiatric and/or neurological histories were included. In addition, none of the subjects were identified as demented by their private physicians. Upon preliminary examination of the
neuropsychological data, the authors noted several outliers in their sample (i.e., individuals with neuropsychological test performances falling over 2 standard deviations below the mean on any 2 measures). Nevertheless, the authors decided to retain these subjects' data in the analyses, noting that in any large sample of normal elderly, a certain proportion is expected to be in the preclinical stages of AD.

Several neuropsychological measures were administered [the Logical Memory (LM) and Visual Reproduction (VR) subtests (immediate and delayed recall) of the WMS; the Rey Auditory Verbal Learning Test (RAVLT); the Rey-Osterrieth Complex Figure Copy (hereafter, Rey Copy) and Recall trials; the BNT; the Trail Making Test parts A & B; and the Digit Span, Comprehension, Arithmetic, and Similarities subtests from the WAIS-R]. Those of interest with regard to the Martin et al. (1986) and Fisher et al. (1996) studies include the Rey Copy task and the BNT.

As in the Valdois et al. (1990) study, cluster analyses revealed six stable subgroups, which did not differ significantly from each other in terms of educational level. Like Valdois et al. (1990), Mitrushina et al. (1995) utilized Euclidean distance as the similarity measure, an approach more sensitive to quantitative than qualitative (i.e., pattern) differences; as a result, three of these subgroups represented level of performance differences. Thus, only 4 distinct clinically meaningful qualitative patterns were revealed: (a) normal performance on all measures (3 subgroups; 70% of the sample); (b) normal performance on all but selected memory measures; (c) a RAD-like pattern; (d) a LAD-like pattern. The authors interpreted the three groups whose members performed normally on all measures [pattern (a) above] as representative of normal aging, while the remaining subgroups [patterns (b), (c), and (d)],
were taken to represent preclinical patterns of the dementing process (Mitrushina et al., 1995).

With specific respect to the Rey Copy and BNT performances, 3 different qualitative patterns emerged: (a) normal (and equivalent) performance on both measures; (b) normal performance on BNT but impaired Rey Copy performance (i.e., RAD-like pattern); (c) normal performance on Rey Copy but impaired BNT performance (i.e., LAD-like pattern). Four subgroups performed within normal limits on the Rey Copy and BNT (3 of them displayed normal performance on all measures administered; the other group performed within normal limits on all but the RAVLT). Two of the three groups who performed within normal limits on all measures nevertheless displayed patterns of relative strengths and weaknesses: group 1 demonstrated a pattern of strength (i.e., superior performance) in the area of visual memory; group 2 demonstrated a relative weakness in visual memory. Group 3 demonstrated relatively equal abilities in all cognitive realms. It should be emphasized that the levels of BNT and Copy performances were equivalent to each other for each of the 3 normal groups (i.e., straight line between these two variables on the profiles). These three profiles within the normal range of performance were interpreted as lifelong patterns of strengths and weaknesses. Group 4, who performed normally (and equivalently) on the BNT and Rey Copy, yet within the impaired range on the RAVLT (but within normal limits on the other memory measures) may represent a preclinical form of isolated memory impairment.

The two remaining subgroups were similar to RAD and LAD. The first, resembling RAD, was impaired on the Copy task and the VR delayed recall trial, was low average to borderline on the LM, RAVLT, Rey-Osterrieth Recall trial and VR immediate trial; yet
performed within normal limits on the BNT, Comprehension, Trails A and B, Digit Span, Arithmetic, and Similarities subtests of the WAIS-R. Thus this group was characterized by mild impairment of visual-constructional functioning and borderline to mild memory impairment (verbal and spatial, with worse spatial memory), in the face of spared language and executive functioning.

The second group, resembling LAD, was impaired on all measures save for the RAVLT Recognition trial, the Rey Copy task and Rey recall trial, and Trails A. Naming and Trails B were most notably impaired. Thus this group demonstrated spared visual constructional functioning, within the context of impaired naming, memory (more so in verbal realm), and executive functioning abilities. These patients were older and showed a trend toward being less educated than members of the other subgroups.

In sum, with specific respect to BNT and Rey Copy performances, the first group was impaired on the latter but not the former, while the second group performed within normal limits on the Copy task and was clearly impaired (1.5 standard deviations below the mean) on the BNT. Overall, the memory performance of the first group (i.e., resembling RAD) was generally within normal limits but was clearly below that of the normal groups, approaching one standard deviation below the mean. The LAD-like group’s memory performance was clearly impaired, falling at one standard deviation below the mean for all but the RAVLT recognition and Rey-Osterrieth recall trials. It would appear that these two subgroups may represent preclinical RAD and LAD patterns. As the authors point out, overall, these findings support the subtype model, rather than the possibility that AD subgroups simply represent pronouncement of premorbid cognitive strengths and weaknesses. Although variations within
the normal group were apparent, the patterns demonstrated by the 3 normal groups across the measures did not resemble those of the preclinical groups. In particular, each normal group performed at the same level on both the BNT and Copy.

**Integrative Summary**

It is apparent from the above review that the original conceptualization of AD as a homogeneous neuropsychological entity is in the process of undergoing a change, as evidenced by the research avenues pursued over the past decade or so. The landmark research of Martin and his colleagues (1986), demonstrating a double dissociation of function between semantic knowledge and visual-constructional systems amongst AD patients, and the subsequent demonstration that corresponding inter-hemispheric asymmetrical metabolic dysfunction were associated with these psychometric subgroupings, strongly suggested a right-left factor contributing to heterogeneous neuropsychological presentations. The results of this research suggested that whereas most AD patients undergo a more or less homogeneous dissolution of neuropsychological functions (presumably reflecting underlying symmetrical pathology), in a manner consistent with the global stage model, qualitatively distinct subgroups of patients with pronounced deficits in one area of functioning and relatively spared abilities in the other could be identified. Furthermore, these subgroups appear to reflect asymmetrical neuropathology in the early period of the disease. The Becker et al. (1988), Fisher et al. (1996), and Strite et al. (1997) consistent replications across subtyping methods and samples added weight to the above neuropsychological subtyping findings. As well, the findings that the 3 subgroups do not differ in terms of duration of illness (Becker et al., 1988, Fisher et al., 1996, Martin et al., 1986) reinforced the limitations of a
strict stage model conceptualization of AD.

The suggestion of neuropsychological subtypes of AD raises the issue of heterogeneity in terms of patterns of cognitive strengths and weaknesses among the normal aging population. In this manner, the apparent heterogeneity of AD may not be a function of variants of the disease *per se*, but may merely reflect the effects of a unitary disease process in exacerbating premorbid patterns. This potential confound, in turn, points to a limitation in the neuropsychological AD subgrouping research conducted to date. That is, none of the studies that employed objective classificatory methodology included normal control groups in their classification attempts (i.e., in a parallel fashion), thus failing to rule out the possibility that any identified AD subgroups simply reflected premorbid patterns of neuropsychological strengths and weaknesses.

Two studies have been conducted in which neuropsychological data of normal elderly were cluster analyzed. While the results of the first were inconclusive, the second study (Mitrushina et al., 1995) provides evidence for preclinical subtypes resembling RAD and LAD, which did not correspond with normal variability (i.e., patterns of strengths and weaknesses). However, longitudinal investigations of those suspected to be in the preclinical AD period are necessary in order to determine whether such individuals do indeed go on to develop AD. As well, longitudinal investigations of normal elderly aimed at tracking premorbid strengths and weaknesses and the impact of these on subsequent manifestations of AD and eventual subgroup assignment would be helpful. Perhaps the easiest way to evaluate this issue would be to conduct subgrouping studies employing both normal elderly and AD patients in which both groups are given the same measures and whose data are
subjected to the same empirical classification analyses. If qualitatively similar subgroups emerged from both samples, a strong case could be made that the identified AD subgroups represent nothing more than exacerbation of premorbid strengths and weaknesses.

The discussion now proceeds to applications of existing theoretical models in explaining the above neuropsychological subgrouping findings, and presentation of a rudimentary neuropsychological model of AD which integrates these earlier theorizations.

**Toward a Neuropsychological Model of AD**

**Goldberg-Costa Model**

Goldberg and Costa (1981) introduced a model of hemispheric specialization based on neuroanatomical distinctions between the cerebral hemispheres. They noted that the right hemisphere is more diffusely organized than the left, has a lower grey to white matter ratio (i.e., implying more long myelinated fibres), and has a greater representation of the association areas. Conversely, the left hemisphere is more focally organized; distinct modality specific cortical areas are more prominent. Following the suggestion of Gur et al. (1980), Goldberg and Costa interpreted the grey to white ratio as a marker of structural organization with respect to intra- versus inter-regional integration. As such, they theorized that the structural organization and connectivity patterns of the right hemisphere make it the hemisphere more suitably organized for interregional integration, whereas the neuroanatomical structure and connectivity of the left hemisphere make this hemisphere more suitable for intraregional integration.

On the basis of these structural asymmetries, Goldberg and Costa (1981) hypothesized that the right hemisphere has a greater capacity to deal with complex information (i.e., due
to its greater representation of association areas), and also has the capability to process several modes of representation within a single cognitive task (i.e., due to its interregional connectivity). One implication of this is that the right hemisphere may have a greater ability to process novel stimuli for which the individual has no pre-existing code or descriptive system (i.e., for which the individual does not have rules or codes to routinely apply). On the other hand, it was theorized that the left hemisphere functions more effectively during tasks requiring the utilization of a single mode of processing, due to its greater representation of unimodal sensory and motor areas and intramodal connectivity, and was thus particularly suited for the stereotypic application and storage of codes or descriptive systems that have already been learned.

The neuropsychological patterns of RAD and LAD can readily be interpreted, in terms of this model, as breakdowns of the respective hemispheres (initially in relative isolation). Members of the LAD group are not able to access their descriptive systems effectively, as such systems are in the process of disorganizing, or breaking down. Thus, overlearned material is gradually being lost, depending on the strength of the associations, with weaker associations (assuming less redundancy or compensatory systems) most vulnerable. This is likely the result of the degeneration of left hemisphere grey matter. Application of the Goldberg-Costa (1981) model leads one to predict that LAD patients will be able to perform tasks which do not rely on previously learned descriptive systems (i.e., those unreliant on the acquired knowledge base), depending on their level of attention and severity of memory deficits. On the other hand, the deficits of the RAD group may be explained in terms of the breakdown of the essential white matter patterns of connectivity in the posterior right
The designation of AD as a “cortical” dementia would appear to be misleading on several grounds. It is known that several subcortical structures are involved (e.g., nucleus basalis of Meynert, locus ceruleus, raphe nucleus) to varying degrees in AD (Bondareff et al., 1987; Mann & Esiri, 1988). Further, several investigators have reported white matter degeneration in AD brains (Brun & Englund, 1986; Gottfries, Bartfai, Carlsson, Eckernas, & Svennerholm, 1986; O'Brien, Desmond, Ames et al., 1996; for review see O'Brien, Ames, & Schwietzer, 1996). Indeed, there is a suggestion that white matter lesions are common in patients with AD, even when great effort is expended in ensuring that those with vascular risk factors, cerebrovascular and cardiovascular disease are specifically excluded (Bennett, Gilley, Wilson, Huckman & Fox, 1992), or that these vascular factors are controlled for (O'Brien, Desmond, Ames, et al., 1996). Bennett et al. (1992) noted that white matter lesions appear most commonly in the centrum semiovale and other subcortical regions surrounding this center. Disruption of this area is of great significance, as this mass of white matter contains commissural, association, and projection fibres (Carpenter, 1991).

The neuropathology of AD tends to target populations of projection neurons important for feedback/feedforward projections among the limbic system and association cortices, resulting in disconnections (Hyman, Arriagada, Van Hoesen, & Damasio, 1993). For example, the entorhinal cortex receives projections from the association cortices, and passes this information on to the hippocampus via the perforant pathway, and as such, is the major afferent link between the multimodal association areas and the hippocampus. During the course of AD, the entorhinal cortex is particularly hard hit by tangles (Hyman et al., 1993),
and this eventually destroys the perforant pathway, isolating the hippocampus from the cortex. Similarly, the CA1 zones and subiculum, which give rise to the major efferent pathways of the hippocampal formation, are also the targets of massive AD neuropathology (Hyman, Van Hoesen, Damasio, & Barnes, 1984). Thus, the high densities of tangles in the afferent neurons of the entorhinal cortex and the efferent neurons of the subiculum and CA1 zone serve to functionally disconnect the hippocampus from the rest of the cerebral cortex.

Of greater importance to the current investigation, is the other major area of the brain severely affected by the neuropathology of AD, the multimodal association cortex. This area is highly prone to neurofibrillary generation and neuritic plaques. In particular, the large pyramidal interassociational neurons of layers III, V, and VI are the primary recipients of damage (Chui, 1989; Lewis, Campbell, & Terry, 1987). This is of significance because it is the long axons of these neurons which form cortico-cortical connections. Given that these are the pivotal layers for cortico-cortical projections, some authors have suggested that, just as the pathology in the entorhinal cortex and subiculum disrupts the input and output of the hippocampus, pathology in layers III, V, and VI of the association cortices disrupts the input and output of each area (i.e., the feedforward and feedback projections that conjoin cortical areas of different hierarchies) (Damasio, Van Hoesen, & Hyman, 1990). Thus, it is argued that the association cortex becomes functionally disconnected. The parietal association cortex is the major posterior multimodal association area. Reductions in input to this area from unimodal association regions may account for the deficits observed in RAD.

**White Matter/NLD Model**

The above described white matter degeneration and disrupted connectivity of the
multimodal association region in AD allows for the application of the White Matter Model (Rourke, 1982, 1987, 1988, 1989, 1995). It appears that in RAD, the left hemisphere is spared for a period of time. Selective pathology affecting associational white matter (in addition to perturbation of grey matter) may disrupt the patterns of connectivity upon which right hemisphere functions are dependent, thus interfering with intermodal integration and resulting in the nonverbal learning disability pattern.

Rourke's (1982) model, building on the Goldberg and Costa (1981) conceptualization of hemispheric specialization, was designed to account for the deficits observed in one particular subtype of learning disability known as NLD (Nonverbal Learning Disability). Rourke (1982) contended that the deficits characterizing individuals with NLD could be explained in terms of the Goldberg-Costa model, as a result of dysfunctional right hemisphere systems. The patterns of neuropsychological assets and deficits that constitute the NLD syndrome are as follows: deficient visual-spatial-organizational and conceptual abilities; marked difficulties in dealing with novelty, informational complexity, and problem solving; all within the context of preserved rote verbal skills (Rourke, 1982).

Following Goldberg and Costa (1981), Rourke (1989, 1995) focused on the different neuroanatomical patterns between the cerebral hemispheres, theorizing that because the right hemisphere appears more dependent upon integrity of white matter connectivity, any neurological disease process that affects white matter would be expected to produce the syndrome of NLD. The relevant principles/deductions of the explanatory White Matter Model (Rourke, 1989, 1995) are listed below:

1. **Amount of white matter destroyed or dysfunctional.** In general, the more white
matter (relative to total brain mass) that is lesioned, removed, or dysfunctional, the more likely it is that the NLD syndrome will be in evidence.

2. **Development and maintenance of learned behaviour.** Right hemisphere white matter is crucial for the development and maintenance of its specific functions, such as intermodal integration, especially when novel information-processing situations are encountered. For example, significant destruction or permanent disruption of right-hemisphere white matter would be expected to pose a permanent handicap to the acquisition of new descriptive systems at any developmental stage. Left-hemisphere white matter is essential for the development, but not necessarily the maintenance of its specific penchants. Once natural language is acquired and automatized, specific functions presumably subserved by the prominent opercula of the left hemisphere would be expected to be relatively impervious to destruction or permanent disruption of white matter not immediately adjacent to and/or forming an integral part of the functioning of these opercula.

3. **Sufficiency.** A significant lesion confined to the right hemisphere may constitute a sufficient condition for the production of the NLD syndrome.

4. **Necessity.** The necessary (and "dose sensitive") condition for the production of the NLD syndrome is the destruction or dysfunction of white matter that is required for intermodal integration.

5. Left hemisphere systems, many of which are relatively "encapsulated" within the three major opercula (grey matter and short association fibers) may maintain enough stimulation within and between each other so that some fairly sophisticated
intramodal integrations can proceed with little or no input from the right hemisphere
(Rourke, 1989, p.113-115).

Given the progressive nature of AD (i.e., eventual global pathology), these principles require elucidation in the instance of RAD. Namely, in RAD, the NLD pattern would be expected to emerge for only a brief period. Thus, after a certain period of time, the spread of pathology to left hemisphere systems would be expected to give way to a more global pattern of cognitive deficits.

Reflection on the implications of the White Matter Model might lead one to ponder whether members of the RAD subgroup were exposed to a toxin or some other etiological agent to which white matter is particularly susceptible. Because the functioning of the right hemisphere is (theoretically) more dependent upon white matter integrity, this line of thought may explain the appearance of the NLD pattern despite global white matter pathology. Indeed, it has been pointed out that the subcortical structures affected in AD are not random; the disease is “partial” to those with extensive connectivity (i.e., white matter) patterns (Geula & Mesulam, 1994).

**Minimal Progression Model**

Much of the literature suggests that memory impairment is the most pronounced and initially appearing symptom in AD (Damasio et al., 1990; Grady et al., 1988; Price et al., 1993; Welsh, Butters, Hughes, Mohs, & Heyman, 1992), and may even occur in isolation (Berent et al., 1995). Studies investigating material-specific memory loss also suggest that memory impairment precedes other neuropsychological dysfunction (Becker, Lopez, & Wess,
1992; Martin et al., 1985). Based on the literature, it seems reasonable to assume that this impairment reflects underlying pathology in the region of the hippocampus and amygdala (Hyman et al., 1984; Milner, Corkin, & Teuber, 1968; Mishkin, 1978; Van Hoesen & Damasio, 1987). Martin (1990), in proposing his "Minimal Progression Model," noted the following empirical support for the hippocampus/amygdala as the initial locus of pathology in AD: (a) autopsy evidence of severe and perhaps universal involvement of this region; (b) the established role of these structures in learning and memory; (c) reports of isolated amnesic impairment in biopsy-proved cases of AD (e.g., Neary et al., 1986).

Indirect evidence supporting the amygdala/hippocampus as the initial locus of pathology in AD is provided by autopsy studies of the brains of Down's Syndrome (DS) patients who die before 50 years of age. It is well known that individuals with DS who live past the age of 50 years have distributions and quantities of plaques and tangles in their brains that are comparable to those observed in AD (Whalley, 1982; Wisniewski, Wisniewski, & Wen, 1985). However, those with DS who die at 20 years of age or younger rarely have any plaques and tangles at all (Mann & Esiri, 1988; Whalley, 1982). Thus, between the approximate ages of 20 and 50, the changes that occur in DS brains can be viewed as a pathological model of AD pathogenesis and progression (Mann & Esiri, 1988).

In order to investigate the topographic distribution of plaque and tangle formation corresponding to the different evolutionary periods, Mann and Esiri (1988) performed autopsies on nine patients with DS who died before age 50 (age range 13-49 years). Three groups of patients were identified (ages in years, in parentheses): (a) those with no plaques or tangles anywhere in the brain (13, 31), (b) those with numerous plaques and tangles
present in a distribution similar to that of AD brains (42, 48), and most interestingly, (c) those with an "intermediate" pathological picture, in which "mature" plaques were observed in the amygdala-hippocampal region and "primitive" plaques were seen in the neocortex (37, 40, 42, 43, 49). The tangles in group (c) brains were most heavily concentrated in layer II of the entorhinal cortex. This research suggests that initial plaque and tangle formation arises in the area of the hippocampus and amygdala, and then spreads to other areas (e.g., neocortex, nucleus basalis, locus ceruleus).

In accord with these findings, Martin (1990) theorized that from the medial temporal region, the disease progresses to the posterior temporal and parietal cortex, and then proceeds in a posterior to anterior direction, eventually reaching the frontal cortex and then also the subcortical nuclei (e.g., basalis, locus ceruleus, raphe). These posterior-to-anterior formulations are consistent with those from autopsy studies reporting that in AD brains, the frontal cortex is affected to a much lesser extent than are the other regions (e.g., Brun & Englund, 1986). As well, neuroimaging studies suggest that the parietal and temporal regions are more heavily affected early in the disease, and that frontal involvement is associated positively with disease severity (Chase, Burrows, & Mohr, 1987; Friedland, Budinger, Koss, & Ober, 1985; Haxby, Duara, Grady, Cutler, & Rapoport, 1985; Haxby et al., 1986). Furthermore, clinical observations that motor signs appear late in the course of the disease in most cases (Cummings & Benson, 1992) also suggest that the frontal cortex does not become crucially involved until late in the disease. Thus, as Martin (1990) suggests, a posterior to anterior pattern of progression appears to be a reasonable speculation.

We now theorize about RAD and LAD. Martin (1990) suggested that perhaps the
disease initially invades the medial temporal region asymmetrically in some individuals and continues to proceed in an asymmetrical fashion initially. This line of thought would imply subgroup-specific patterns of neuropsychological dysfunction.

An Integration of Theoretical Applications: The Subgroup-Specific Progression Model

Drawing on the Goldberg-Costa (1981) model of hemispheric specialization and the Rourke (1982) White Matter Model, Fisher et al. (1996) theorized that: (a) LAD disrupts intramodal left hemisphere grey matter systems mediating overlearned knowledge representations; (b) RAD disrupts white matter interregional integration systems, which are more prominent in the right hemisphere; (c) GAD disrupts both types of cerebral systems. Fisher, Rourke, and Bieliauskas (1997) went on to predict distinct neuropsychological progression patterns for the three AD subgroups, drawing on and extending the Minimal Progression Model offered by Martin (1990). These theorizations are briefly outlined below.

1. Typical AD (i.e., GAD). In most cases of AD, the original locus of pathology arises in the central region of the hippocampal formation. Then, it spreads out bilaterally, encompassing both of the distinct functional systems of the hippocampal formation: one left, one right. Thus, among global AD cases, anterograde memory loss will be general (i.e., there will be no evidence of material-specific impairment). Thereafter, the pathology progresses systematically, concentrating particularly in the entorhinal cortex, until it reaches connections to the association cortices. Eventually, it spreads to posterior temporal and parietal regions as Martin (1990) suggested. At the same time, the pathology radiates from the hippocampal formation to the basal forebrain. From there, it destroys projections to the prefrontal lobe. This could explain why the pre-central and post-central gyri are spared until the very late
period of the disease. As the pathology of GAD is theorized to affect both hemispheres relatively equally, global cortical deficits (i.e., visual-spatial and semantic) should be consistently apparent. That is, there should be minimal discrepancy between performance levels in these two domains of functioning at any time during the illness (i.e., with progression).

2. RAD. In RAD, the hippocampus/amygdala is also the origin of pathology. However, the pathology arises not medially, but laterally to the right. Thus, the putative right functional memory system is attacked first, resulting in material-specific memory loss for complex visual-spatial information. Several studies have demonstrated the role of the right hippocampus in tactile and visually-based learning tasks involving spatial information and orientation, and it has been reported that the greater the removal of tissue from this area, the greater the memory impairment (Corkin, 1965; Milner, 1965; Smith & Milner, 1981). This direct relationship between cell loss and increasing deficit would appear to explain the progressive nature of the memory impairment in AD. While proceeding to the right temporal and parietal association cortices, the disease spreads also to the contralateral (left) hippocampus/amygdala. From the contralateral hippocampus/amygdala, the disease spreads to the parietal and temporal cortices on this (the left) side as well. Thus, in RAD, the general pattern of neuropsychological deficits should be as follows: material-specific memory loss (visual-spatial), followed by visual-spatial deficits, then global memory impairment, and finally global cortical deficits (i.e., both visual-spatial, and semantic, but more severe visual-spatial). This formulation is consistent with the results of the material-specific memory loss research of Martin et al. (1985) and Becker et al. (1992). From here the disease progresses in the
typical pattern outlined in (1). The discrepancy between visual-constructional and semantic functioning should remain with time, being the largest initially, yet evincing reduction with time and progression of the disease. This group would be expected to demonstrate an NLD pattern early in the disease process. However, as the disease advances, the NLD pattern would become less and less apparent.

3. LAD. A pattern opposite to that of RAD is theorized for this subgroup. As with GAD and RAD, the hippocampus/amygdala is the site of pathological origin. In LAD however, this pathology arises in and remains disproportionate for a time in the left hippocampal region. The role of the left hemisphere hippocampus in verbal learning of a rote nature (e.g., recall of word lists) has been demonstrated and, as with the right hemisphere hippocampus, there seems to be a direct relationship between amount of hippocampus removed, and the extent of memory impairment (Petrides & Milner, 1982). Thus in LAD, the putative left functional memory system is affected first, resulting in material-specific memory loss for rote verbal information. The disease then progresses to the left temporal and parietal association cortices, and the right hippocampal formation. From the right hippocampus/amygdala, the disease spreads to the right parietal and temporal cortices. Thus, in LAD, the general pattern of neuropsychological deficits should be as follows: material-specific memory loss (rote verbal), followed by semantic knowledge deficits, then global memory impairment, and finally global cortical deficits (i.e., both semantic and visual-spatial, the former being more severe). From here the disease progresses as in (1). The most prominent deficits of this group are expected in the realm of access to acquired (i.e., overlearned) knowledge systems. With respect to visual-constructional and semantic
functioning, the discrepancy between the two should reduce with time and progression of the disease.

Where do we go from here?

It would appear that what is needed at this point to advance neuropsychological subgrouping research is (a) replication of the Fisher et al. (1996) reliability study, and (b) evaluation of the above subgroup-specific theoretical progression model. As the Fisher et al. (1996) study was limited in terms of sample size, sample composition, and lack of inclusion of a control group, replication attempts should aspire to correct these shortcomings. With respect to evaluation of the above presented neuropsychological model of AD, given the progressive nature of the disease, it is imperative that such evaluation include investigation of the dynamics of the model. As such, longitudinal data are needed. In sum, what is required to advance the AD subgrouping literature is a larger, more representative sample, containing both AD patients and comparable healthy elderly controls, for which longitudinal data are available. While the attainment of such a sample may sound an insurmountable task, the multicenter Consortium for the Establishment of a Registry for Alzheimer’s Disease (CERAD) project provides a source of data fulfilling the above mentioned requirements.

On CERAD

CERAD was established in 1986 with the goal of collecting uniform longitudinal clinical, neuropsychological, and neuropathological data on a large sample of patients with AD recruited from 24 American university medical centers (Heyman, 1996). The CERAD test battery was designed to provide a brief yet reliable standardized assessment procedure for patients with suspected AD, for research use by these and other university centers (Morris et
al., 1989). Annual re-testing of the patients allows for large scale tracking of the natural progression of AD (Heyman, 1996).

The mission statement of CERAD from the 1996 *Manual of Operations* (Heyman, 1996) is to do as follows:

(a) standardize procedures and instruments for the evaluation and diagnosis of cases with AD;

(b) assess the manifestations and natural progression of this disorder through standardized assessment, over time, of cases and controls enrolled at multiple clinical sites;

(c) provide information on the natural history of AD, the relationship of clinical, neuropsychological, and neuropathological changes, and identification of subgroups usually not obtainable from the relatively small number of cases seen at any one site;

(d) disseminate these findings through publication of peer-reviewed papers, presentations at national meetings, and the training of investigators in this area of research. (p.2)

**CERAD Neuropsychological Studies**

The CERAD neuropsychological battery includes measures of verbal fluency, confrontational naming, verbal list-learning (immediate free recall, delayed free recall, recognition), and constructional praxis. Studies of this battery have indicated that although all the measures in the battery distinguish *probable* AD patients from normal elderly, the delayed verbal recall measure is particularly sensitive to early AD, being the single best discriminator of all the CERAD measures in distinguishing early AD from normal aging (Welsh, Butters, Hughes, Mohs, & Heyman, 1991; Welsh et al., 1992). Stepwise linear discriminant function analysis has revealed that the only other measure that contributes to further discrimination between normal elderly and those with early AD is confrontational
naming (Welsh et al., 1992). However, confrontational naming had little predictive power as a lone discriminator, and when combined with delayed recall, did not appreciably improve the discriminability obtained by delayed recall alone.

Despite the above, due to floor effects, delayed recall scores are not generally as useful in terms of determining dementia severity or staging progression of the disease (Welsh et al., 1991). Rather, the measures of semantic processing and visual-constructional functioning in combination with memory measures best differentiates stages of AD progression (Welsh et al., 1992). For example, a combination of fluency, constructional praxis, and recognition memory best differentiates patients with moderate dementia from those with severe dementia (Welsh et al., 1992).

**Purpose of the Current Study**

The purpose of the study described herein was to replicate the findings of Fisher et al. (1996) on a more representative, independent sample. Furthermore, inclusion of a longitudinal design allowed for evaluation of the dynamics of the theoretical model presented above. Improvements in the current study design over those employed in past research include: (a) employment of a non-demented comparison group, (b) longitudinal analysis of patient data, and (c) and the utilization of a large, multi-center data source.

**Hypotheses**

**Hypothesis 1**: Patients with *probable* AD can be reliably differentiated into 3 empirically derived neuropsychological subgroups (i.e., GAD, RAD, and LAD).

**Rationale**: Past research has reliably demonstrated that when neuropsychological data from *probable* AD patients are subjected to empirical classification techniques, 3 qualitatively
distinct subgroups of patients emerge (Fisher et al., 1996; Martin et al., 1986).

**Hypothesis 2:** These 3 patient subgroups can be determined on the basis of performance in two domains of neuropsychological functioning: visual-constructional functioning and semantic knowledge accessibility.

**Rationale.** In accord with past research (Fisher et al., 1996; Martin et al., 1986), it is expected that the three subgroups of AD patients will resemble the theoretical ideal types (Fisher, Rourke, & Bieliauskas, 1997) (a) GAD, (b) RAD, and (c) LAD, respectively: (a) those with global impairment (impaired access to semantic knowledge and impaired visual-constructional functioning); (b) those with relatively spared accessibility of semantic knowledge yet impaired visual-constructional functioning; (c) those with impaired ability to access semantic knowledge, but relatively unimpaired visual-constructional functioning.

**Hypothesis 3:** Normal elderly control subjects do not display neuropsychological patterns resembling those of the RAD and LAD patient subgroups.

**Rationale.** Past research has demonstrated that when data from normal elderly samples are subjected to empirical classification techniques, significant interpretable subgroups resembling RAD and LAD do not emerge among the unimpaired subjects (Mitrushina et al., 1995; Valdois et al, 1990). Clearly, it is expected that normal elderly may show different patterns of relative strengths and weaknesses (i.e., on memory measures; Mitrushina et al., 1995). However, it is expected that they will demonstrate consistent levels of performance in both the semantic and visual-constructional domains (i.e., like GAD patients, yet within the normal range of functioning) (Mitrushina et al., 1995).

**Hypothesis 4:** In the early and middle period of the disease, initial neuropsychological
discrepancy patterns of the RAD and LAD subgroups will be consistently apparent, yet will gradually diminish as the disease progresses. On the other hand, the GAD pattern will remain consistent with advancement of the disease.

As such, it is expected that those with the GAD pattern will continue to show impairment in both the semantic knowledge and visual-constructional realms of functioning, consistently exhibiting a pattern of relatively equivalent decline in the two domains. Those with RAD will continue to show sparing of semantic knowledge accessibility for a period of time, and though impairment in this domain will eventually occur, it will be relatively less severe than that of visual-constructional functioning until late in the disease, when all functioning is profoundly disrupted.

The prediction for LAD patients is the opposite of that for the RAD group; that is, those classified with LAD will initially display visual-constructional functioning within normal limits, in the face of difficulty in the area of semantic knowledge. While, with time, visual-constructional functioning will eventually show decline into the impaired range, it will continue to be relatively less severely affected than word finding ability, until the latest period of the disease, when all three subgroups will be rendered indistinguishably globally impaired, as floors effects of the neuropsychological measures are reached. As such, longitudinal analyses of performance discrepancies between semantic functioning and visual-constructional functioning will reveal that such discrepancies diminish gradually over time for the RAD and LAD groups, but remain relatively stable for the GAD group.

Rationale. Past research involving neuropsychological subgroups of AD has suggested that the initial neuropsychological pattern (i.e., GAD, RAD, & LAD) remains detectable
across time, during the early and middle period of the disease (Fisher et al., 1997; Martin, 1990). It is known however that AD is progressive, and eventually results in widespread cognitive impairment (Price et al., 1993).
CHAPTER III
Methodology

Subjects

Clinical, demographic, and neuropsychological CERAD data from 960 AD patients (564 female, 396 male) and 465 non-demented controls (305 female, 160 male) were utilized in this investigation. The discrepancy in number of patients versus controls was due to data availability. All subjects had agreed to participate in the longitudinal data collection process and had consented to the terms of the CERAD study. All patients were diagnosed with probable AD (McKhann et al., 1984) after having been referred to a CERAD site for inpatient/outpatient consultation regarding suspected dementia. [A list of participating CERAD sites and their respective referral areas is provided in Appendix A (Heyman, 1996).]

All patients had experienced gradual onset of symptoms and progressive memory loss for a period of at least 12 months (Morris et al., 1989). The average duration of time since onset of first AD symptoms was 4.21 years ($SD = 2.64$). With respect to the staging of AD, the mean Clinical Dementia Rating (CDR; see Appendix B for description) fell directly between the mildly and moderately demented stages ($M = 1.47; SD = .64$).

None of the AD patients met the following exclusionary criteria: (a) history of a major stroke prior to onset of dementia; (b) history of abrupt onset of dementia; (c) history of substance abuse; (d) presence of hypertension with current systolic blood pressure > 200 or current diastolic blood pressure > 110; (e) presence of comorbid Parkinson's disease; (f) evidence of other major central nervous system disorders or injuries; (g) presence of metabolic disorder(s), delirium, or altered consciousness; (h) inability to complete all the
CERAD neuropsychological tests on at least one occasion; (i) presence of serious cardiac disease (e.g., severe angina, heart failure) or cancer; (j) severe hearing, visual, physical, or language disturbance affecting the ability to complete testing; (k) current major depression or other major psychiatric illness; (l) absence of caregiver or significant other willing to provide adequate history and monitor the patient's progress; (m) uncooperativeness; (n) poor comprehension of or inability to speak and read the language of the test battery; (o) unlikelihood to return for annual follow-up observations.

The comparison (i.e., control) group comprised spouses or caregivers/friends of the AD patients; they were selected so as to be similar to the cases in terms of demographic characteristics (Heyman, 1996). All controls met comparable exclusionary criteria as the patients, and none had a history of cognitive impairment (Heyman, 1996). Inclusionary criteria for both patients and controls were as follows: (a) must be at least 50 years of age; (b) must be free of major health problems, such as cancer, cardiac or respiratory disease, current major depression, and/or recent alcohol or drug abuse; (c) must have normal consciousness with no signs of delirium; (d) must be willing and cooperative research participants; (e) must be able to read and comprehend the language on the neuropsychological forms (Heyman, 1996).

Table 1 presents the demographic characteristics of the patient and control samples and their average levels of performance on neuropsychological measures of interest. The control group was slightly younger ($p < .05$) and more educated ($p < .05$) than the patient sample, although the age and educational ranges and frequency distributions were similar. Both groups were highly educated and predominantly Caucasian.
Table 1

Descriptive Statistics for Demographic and Neuropsychological Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>AD</th>
<th>NC</th>
<th>Range</th>
<th>AD</th>
<th>NC</th>
<th>AD</th>
<th>NC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics (raw)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>72.07(7.93)</td>
<td>68.38(8.04)</td>
<td>50-97</td>
<td>50-93</td>
<td>960</td>
<td>465</td>
<td></td>
</tr>
<tr>
<td>Education (years)</td>
<td>12.30(3.70)</td>
<td>13.67(3.31)</td>
<td>0-26</td>
<td>2-25</td>
<td>960</td>
<td>465</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>396</td>
<td>160</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>564</td>
<td>305</td>
</tr>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>784</td>
<td>429</td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>158</td>
<td>34</td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td><strong>General Impairment Ratings (raw)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDR</td>
<td>1.47(64)</td>
<td></td>
<td>0.5-5</td>
<td></td>
<td></td>
<td>960</td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>17.40(5.77)</td>
<td>28.73(1.62)</td>
<td>1-29</td>
<td>17-30</td>
<td>926</td>
<td>461</td>
<td></td>
</tr>
<tr>
<td>Blessed</td>
<td>4.57(2.48)</td>
<td></td>
<td>0-14</td>
<td></td>
<td></td>
<td>959</td>
<td></td>
</tr>
<tr>
<td><strong>Neuropsychological Tests (T scores)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BNT</td>
<td>18.44(20.59)</td>
<td>48.10(12.37)</td>
<td>0-57</td>
<td>0-57</td>
<td>960</td>
<td>465</td>
<td></td>
</tr>
<tr>
<td>VF</td>
<td>28.34(9.83)</td>
<td>49.47(9.96)</td>
<td>9-60</td>
<td>28-85</td>
<td>960</td>
<td>465</td>
<td></td>
</tr>
<tr>
<td>PRAX</td>
<td>28.18(19.41)</td>
<td>49.39(10.83)</td>
<td>0-62</td>
<td>7-62</td>
<td>960</td>
<td>465</td>
<td></td>
</tr>
<tr>
<td>IVBM</td>
<td>2.36(7.46)</td>
<td>48.85(24.49)</td>
<td>0-71</td>
<td>0-117</td>
<td>960</td>
<td>465</td>
<td></td>
</tr>
<tr>
<td>DVBM</td>
<td>15.50(8.34)</td>
<td>49.30(10.63)</td>
<td>1-54</td>
<td>9-71</td>
<td>960</td>
<td>464</td>
<td></td>
</tr>
<tr>
<td>RECOG</td>
<td>7.03(15.02)</td>
<td>51.39(14.13)</td>
<td>0-62</td>
<td>0-62</td>
<td>960</td>
<td>465</td>
<td></td>
</tr>
</tbody>
</table>
Note. AD = Alzheimer's patients; NC = Normal Controls; CDR = Clinical Dementia Rating; MMSE = Mini Mental State Exam; BNT = Boston Naming Test; VF = Verbal Fluency (Animal Category); PRAX = Constructional Praxis; IVBM = Immediate Verbal Recall; DVBM = Delayed Verbal Recall; RECOG = Recognition Memory.
Semantic Memory (SM) and Visual Construction (VC) composite T scores were calculated for the patients', averaged, and expressed graphically (see Figure 1) to demonstrate the goodness of fit of this sample (as a whole), compared to past AD samples employed in subgrouping research. As can be easily seen, akin to the Fisher et al. (1996) study, the current sample was substantially less impaired than that of Martin et al. (1986) in these two domains of functioning. However, the slightly better performance in the VC as compared to the SM realm of functioning observed by Martin et al. (1986) was also present in the current sample.

For comparison of the current patient sample to those of Fisher et al. (1996) and Martin et al. (1986) on demographic and symptom severity indicators, see Figure 2. Mean age of the current sample closely resembles that of Fisher et al. (1996). The current sample more closely resembles the population at large in terms of educational level compared to the previous two samples. Mean estimate of duration since disease onset, in addition to MMSE and Blessed scores for the current sample approximate those of the sample employed by Fisher et al. (1996). However, compared to the Fisher et al. (1996) sample, the overall degree of the current patients' dementia seems slightly less severe.

**The CERAD Assessment Battery**

The CERAD standardized battery was established in order to ensure that data collected at different research centers are comparable; prior to its development, differing test

---

1 Age-, gender-, and education-corrected T scores were calculated for the BNT, VF, and PRAX variables utilizing normative data provided by Welsh et al. (1994). The Clock Drawings were rated on a three-point scale with 0, 1, and 2, reflecting normal, impaired, and severely impaired performances, respectively. These ratings were reassigned to T scores of 55, 40, and 25, without demographic correction, as such data are not available. Formulas for deriving the current composites were as follows: SM = (T score on BNT + T score on VF) /2; VC = (T score on PRAX + T score on CLOCK)/2. Composites for the other 2 studies were based on similar formulas, utilizing data provided by the authors (Fisher et al., 1996; Martin et al., 1986).
Figure 1. Comparison of current sample to previous samples on Visual Construction and Semantic Memory composites.
Figure 2. Comparison of current and past samples on demographic and symptom severity indicators.
protocols made across-research-center comparisons difficult to evaluate (Morris et al., 1989). Indeed, CERAD was developed in order that large scale, standardized data on AD patients could be made available, in a united attempt toward a better understanding of the disorder. The general goal of the CERAD developers was to provide a brief yet reliable assessment battery which discriminates the cognitive changes characteristic of normal aging from those associated with mild AD (Morris et al., 1989). As well, the battery was designed to be sensitive to (i.e., measure or quantify) progression of the disease process over time (Morris et al., 1989).

There are two parts to the CERAD assessment battery: (a) the clinical evaluation, and (b) the neuropsychological assessment. Each are re-administered to patients and controls annually (Morris et al., 1989). Both have been outlined in detail previously (Morris et al., 1989); a synopsis is provided here.

Clinical Assessment

The clinical assessment component of the CERAD battery is conducted by experienced and certified clinicians. It involves administration of the following: semistructured interviews conducted with AD patients and their caretakers (i.e., informants); medical, physical, and neurological examinations; a depression scale; a general medical history; a drug inventory. The informant provides information regarding the patient’s decline in cognitive abilities from the premorbid level. The following are grossly assessed by the clinician via the Blessed Dementia Scale (Blessed, Tomlinson, & Roth, 1968; see Appendix B for fuller description) and the Short Blessed Test (Katzman et al., 1983): memory, orientation, insight, language, personality, calculations, visual-construction. A Clock Drawing measure is also
administered, and clinically judged for accuracy on a 3-point scale (see description in Appendix B). This measure was used in the current study in the calculation of composite VC scores for comparison of this sample to other samples (i.e., Figure 1). The Clinical Dementia Rating Scale (CDR; Berg, 1984; Hughes, Berg, Danziger, Coben, & Martin, 1982; Morris, 1993) is also administered in order to stage dementia severity; this measure is utilized in the present study as a staging measure (see Appendix B for description of the CDR).

**Neuropsychological Measures**

The CERAD neuropsychological battery is administered by trained and certified psychometrists after the clinical assessment (Morris et al., 1989). The following five measures comprise the battery: Animal Category Verbal Fluency (Isaacs & Kennie, 1973), Modified Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1978), Mini-Mental State Exam (Folstein, Folstein, & McHugh, 1975), Word List Memory (Atkinson & Shiffrin, 1971), and Constructional Praxis (Rosen, Mohs, & Davis, 1984). Each is described in Appendix B. Normative data for the CERAD neuropsychological battery has been collected from normal elderly volunteers and community samples (Heyman, 1996; Welsh et al., 1994).

In the current study, the Animal Category Verbal Fluency Test (VF) and the Modified Boston Naming Test (BNT) were utilized as measures of semantic knowledge accessibility (i.e., word-finding ability); the Constructional Praxis (PRAX) test was used as a measure of visual-constructional functioning; and the Word List Memory test was used to measure immediate (IVBM), delayed (DVBM), and recognition (RECOG) verbal memory. The Mini Mental State Exam (MMSE) was utilized as a measure of dementia severity.
Psychometric Properties of the CERAD Neuropsychological Battery

In a study involving data from 16 CERAD sites, interrater scoring reliability was demonstrated to be high; intraclass correlation coefficients ranged from .92 (Constructional Praxis) to 1.0 (Word List Memory—immediate free-recall) (Morris et al., 1989). Test-retest reliability after a one month period was shown to be more than adequate for each of the different neuropsychological measures. Such reliability coefficients for patients on the semantic and visual-constructional measures ranged from .73 to .90 (Morris et al., 1989). The lower coefficients reported for controls on these measures (i.e., ranging from .54 to .67) are likely due to a restricted range of scores due to ceiling effects (Morris et al., 1989). Nevertheless, all reliability coefficients in the Morris et al. study (for both patients and controls) were statistically significant at the .01 level.

The CERAD battery has been demonstrated to accurately distinguish demented from non-demented subjects (Morris et al., 1989; Welsh et al., 1992), suggesting discriminative validity. In the Morris et al. and Welsh et al. studies, elderly controls significantly outperformed the AD patients on all the neuropsychological measures ($p < .01$). In the Welsh et al. study, discriminant function analysis correctly classified 91% of the subjects into patient and control groups. The battery has also been shown to be sensitive to increasing cognitive decline over time (Morris et al., 1989; Morris et al., 1993; Welsh et al., 1992), suggesting longitudinal validity. In this regard, patients show decline in their test scores over time (i.e., at follow-up assessments), while controls do not (Morris et al., 1989; Morris et al., 1993; Welsh et al., 1992).

An initial factor analysis of the CERAD neuropsychological data from 354 AD
patients revealed three factors accounting for 73% of the total variance among the measures: (I) Memory; (II) Language (i.e., semantic knowledge accessibility); (III) Constructional Praxis (Morris et al., 1989). These factors accounted for 44%, 16%, and 13% of the total variance, respectively. Factor I, Memory, had high loadings from Word List Immediate Free-Recall, Word List Delayed Free-Recall, and Word List Recognition. Factor II had high loadings from the Modified Boston Naming and Animal Fluency measures. Factor III had a high loading only from Constructional Praxis. The Mini Mental State Exam (MMSE) which was developed to tap memory, language, and constructional functioning, had moderate loadings on each factor (Morris et al., 1989).

In a more recent similar factor analytic study of CERAD data from 202 AD patients, 2 factors (i.e., memory and nonmemory) were reported, accounting for 51% of the variance (Larrain & Cimino, 1998). These findings appear to conflict with those originally reported by Morris et al. (1989).

**On Cluster Analysis**

Cluster analysis is a term applied to a set of objective multivariate techniques utilized for grouping a large number of individuals into smaller subgroups (i.e., clusters), so that those in the same subgroup are more similar to each other on some predetermined criteria than they are to individuals in other subgroups (Hair, Anderson, Tatham, & Black, 1992). The resulting clusters should demonstrate high within-cluster homogeneity, and high between-cluster heterogeneity, reflective of natural relationships within the data (Morris et al., 1981). The application of cluster analysis involves many subjective decisions by the researcher, including the choice of similarity measure, algorithm, and overlying clustering approach/method to be
utilized. Several different approaches are available, including the following: hierarchical agglomerative methods, hierarchical divisive methods, iterative partitioning techniques, factor-analytic variants, graphic methods, clumping procedures, and density searching techniques (Blashfield & Aldenderfer, 1988; Everitt, 1974; Morris et al., 1981).

Each of these approaches, in turn, have various specific algorithms and procedural options. As well, many different measures of inter-individual similarity/dissimilarity are available, each varying in sensitivity to profile elevation, shape, and scatter (Blashfield & Aldenderfer, 1988). Euclidean distance is the most commonly used measure of similarity between two individuals (Hair et al., 1992). In essence, this is the measure of the length of a straight line drawn between the two observations, and is sensitive to differences in severity, shape, and scatter between subgroup profiles (Hair et al., 1992). The Pearson product moment correlation is another commonly utilized similarity measure (Blashfield & Aldenderfer, 1988). This measure is disproportionately sensitive to the pattern (i.e., shape) of subgroup profiles, at the expense of loss of information regarding scatter and elevation (Aldenderfer & Blashfield, 1984; Everitt, 1974; Morris & Fletcher, 1988).

Because of the multitude of approaches, algorithms, and similarity measures available, and the lack of appropriate criteria for their selection in particular situations, different solutions are often obtained with the same data set, depending on the method employed (Blashfield & Aldenderfer, 1988; Everitt, 1974; Morris et al., 1981). There is no consensus in the literature as to the criterion by which to select the appropriate number of clusters in a data set. For all the above reasons, investigators must clearly outline the methods utilized, in addition to the rationale behind their decisions. Furthermore, due to the unreliability of one
particular analysis, clustering studies should necessarily include replication of the initial solution across two or more similarity measures, algorithms, and methods (Everitt, 1974; Hair et al., 1992; Morris et al., 1981).

**Clustering Methods Utilized in the Data Analysis Strategy**

**Q-Type Factor Analysis**

This procedure involves the factoring of an intercorrelation matrix of subject profiles, in order to condense large numbers of heterogeneous individuals, into distinct subgroups of similar individuals (Hair et al., 1992). The intercorrelation matrix of test data is first inverted, allowing correlations to be computed between subject profiles rather than between the test variables. Following this, factors with eigenvalues ≥ the ratio of subjects to the number of variables entered into the analysis (Del Dotto & Rourke, 1985) are extracted, using standard procedures (e.g., principal components analysis) and rotated. Each factor is interpreted as a cluster; individuals are assigned to clusters (i.e., factors) on the basis of their highest single factor loading (Blashfield & Aldenderfer, 1988; Everitt, 1974).

**Hierarchical Agglomerative Methods**

Hierarchical agglomerative clustering methods begin with the computation of an interindividual similarity matrix. Initially, each subject forms his/her own cluster (Afifi & Clark, 1990; Norusis, 1993). Next, the 2 most similar cases are combined into a cluster. This continues in subsequent steps; at each, the candidate case that is most similar to one of the existing clusters is combined into this pre-existing cluster, thus reducing the number of clusters by one, in each step (Hair et al., 1992). Once a cluster is formed, it cannot split, but can combine with other clusters. Thus, cases cannot separate from the clusters to which they
have been assigned. These successive fusions are continued until all subjects eventually form
a single group (Afifi & Clark, 1990; Everitt, 1974).

Several hierarchical agglomerative techniques are readily available to the researcher. For example, one of the most common statistical software packages used in psychological research, the Statistical Package for the Social Sciences (SPSS; Norusis, 1993), offers the following methods: single linkage (also known as the nearest neighbour technique), complete linkage (or the furthest neighbour technique), the average linkage between groups (or unweighted pair-group method using arithmetic averages), the average linkage within groups method, Ward’s method, the centroid method, and the median method.

While it is beyond the scope of this discussion to describe each hierarchical agglomerative technique, it should be pointed out that several problems have been noted with the single linkage (Aldenderfer & Blashfield, 1984; Edelbrock, 1979; Everitt, 1974; Hair et al., 1992; Sneath & Sokal, 1973), average linkage between groups (Edelbrock, 1979), and centroid methods (Norusis, 1993). For example, the single linkage technique has been reported to be sensitive to the production of “chaining” effects (resulting in elongated heterogeneous clusters) due to outliers in the data. This is due to the fact that the only criterion for assignment to a cluster by this method is that the case under consideration be similar to only a single member of the existing cluster (Aldenderfer & Blashfield, 1984; Edelbrock, 1979; Everitt, 1974; Norusis, 1993; Sneath & Sokal, 1973). Likewise, due to weak cluster inclusion criteria, the average linkage between groups method has been known to be overly sensitive to outliers in the data (Edelbrock, 1979; Williams, Clifford, & Lance, 1972), and the centroid method has similar undesirable properties (Norusis, 1993).
On the other hand, average linkage within groups, complete linkage, and Ward’s method have been successfully accurate in past subtyping investigations (DeLuca, Rourke, & Del Dotto, 1990; Fisher et al., 1996). The average linkage within groups method represents a refining variant of the average linkage between groups method (Norusis, 1993), reducing the impact of outliers. In a similar manner, complete linkage corrects some of the shortcomings of the single linkage procedure (Aldenderfer & Blashfield, 1984). Ward’s method is one of the most popular methods, and in multimethod clustering studies, is often one of the most accurate (Aldenderfer & Blashfield, 1984; Fisher et al., 1996). For these reasons, the average linkage within groups, complete linkage, and Ward’s techniques were selected for use in the current study. Each is briefly described as follows.

**Average Linkage Within Groups.**

This method combines clusters in a manner such that the mean similarity between all cases in the new cluster is as high as possible. It does this by considering the similarity between two clusters as the average of the similarities between all possible pairs of cases in the resulting cluster (Norusis, 1990). This technique is considered a refinement over the unweighted paired-group method (i.e., average linkage between groups) because the latter method only considers similarities between pairs of cases in different clusters, while average linkage within groups considers all possible pairs of cases (Norusis, 1993).

**Complete Linkage.**

In this method, cases are assigned according to the requirement that each cluster member must be more similar to the most dissimilar member of that cluster, than to the most dissimilar member of any other cluster (Edelbrock, 1979). This method is valuable in that it
bypasses the "chaining problem" (Hair et al., 1992) of single linkage, by requiring that any candidate for inclusion into an existing cluster must be within a certain level of similarity to all members of that cluster (Aldenderfer & Blashfield, 1984; Sokal & Michener, 1958).

**Ward's Method.**

This method combines clusters which result in the smallest increase in overall sum of the squared within-cluster variance (Hair et al., 1992; Morris et al., 1981; Ward, 1963). As such, at each stage in the clustering procedure, the within-cluster sum of squares is calculated for all possible combinations obtainable by combining two clusters from the previous stage. The case is then assigned to (or the cluster is then merged with) the cluster resulting in the least increase in variance (Ward, 1963).

**K-Means Iterative Partitioning**

This cluster analytic method produces clusters by estimating cluster centers based on the values of the variables and assigning cases to the centers that are the nearest (measured by squared Euclidean distance) (Aldenderfer & Blashfield, 1984). As each case is added to a cluster, the center is updated to a mean for the cases that are thus far included in the cluster. After each case is classified in this manner, the algorithm reassigns each case to the nearest cluster centers: if the case is closest to the centroid (mean) of it's own cluster, it is left in that cluster, otherwise, it is reassigned to the cluster with the closest centroid (MacQueen, 1967; Norusis, 1988). This procedure is repeated until a stable solution is reached, in which no individual case changes cluster membership (Afifi & Clark, 1990).

**Overview of Data Analysis Process**

The first step of data analysis in this project involved conversion of raw patient scores
on the CERAD measures of semantic accessibility, visual-construction, and memory [modified Boston Naming Test (BNT), Animal Fluency (VF), Constructional Praxis (PRAX), Word-List immediate free recall (IVBM), Word-List delayed recall (DVBM), Word-List recognition (RECOG)] to demographically corrected standard $T$ scores for comparative purposes, utilizing normative data provided by Welsh et al. (1994). Subsequently, a Q-type factor analysis (i.e., utilizing an inverted matrix with columns representing the subjects and rows the variables) was undertaken, in which product moment correlations were calculated between all possible column pairs (i.e., between subject profiles). Factors extracted were retained and rotated to Varimax criterion, if the eigenvalues equaled or exceeded the ratio of number of subjects to the number of variables (Del Dotto & Rourke, 1985). An initial Q-type factor analysis was thought beneficial in that this alternative “clustering” method has no bias against creating relatively small clusters; many cluster analytic techniques (including Ward’s and complete linkage) are biased toward creating clusters of equal size (Aldenderfer & Blashfield, 1984; Edelbrock, 1979; Hair et al., 1992). Furthermore, Q-analysis controls for the presence of outliers, as every case is not necessarily required assignment to a cluster.

The Q-analysis identified significant qualitatively distinct subgroups of patients (as evaluated via multivariate analysis of variance and profile analysis procedures), and as such, various cluster analyses were performed on the $T$ score data, in an attempt to generate similar subgroupings across multiple methods (i.e., establish internal validity). Only those clusters emerging via the majority of methods were accepted as reliable. The product moment correlation coefficient was chosen as the primary similarity measure (i.e., utilized in the 3 hierarchical agglomerative analyses), given that the prime interest of this study involved
qualitative pattern (i.e., shape) variability (Aldenderfer & Blashfield, 1984; Everitt, 1974; Morris et al., 1981). The above described hierarchical agglomerative methods (average linkage within groups, complete linkage, and Ward's method) were utilized first, followed by an iterative partitioning procedure (i.e., K-Means), in which squared Euclidean distance served as the similarity measure. The appropriate number of clusters in the data was determined by inspection of the cluster fusion coefficients at each stage (i.e., for sudden drop), in addition to the corresponding dendrograms (Everitt & Dunn, 1991).

The solutions derived by the hierarchical and iterative cluster analytic techniques were subsequently compared to those generated by the Q-factor analysis. This was done 3 ways: (a) visual comparison of mean profiles generated by each procedure; (b) misclassification analysis; (c) computation of a multimethod-multiprofile matrix of intercorrelations.

Demographic and psychometric data were provided for the Q-derived subgroups. In addition, comparisons were made between the subgroups to determine whether age, education, duration of illness, and severity of the illness (i.e., Blessed, CDR, & MMSE scores) differed between the groups. Following the patient data analyses, the same procedures outlined above were performed on the control sample data.

Longitudinal analyses aimed at evaluating stability of initial patient profiles in the early period of the disease were also conducted. In meeting this end, five subjects from each subgroup were randomly selected, and their patterns of discrepancy between semantic and visual-constructional functioning were qualitatively examined across time.

Following this, discrepancy scores (i.e., differences between BNT and PRAX performances) were calculated for the patients at each testing. Correlations between
discrepancy scores and testing number (i.e., first testing, second testing, etc.) were then calculated. For subgroups LAD and RAD, diminishing discrepancy scores were expected across time, while for GAD, this was not expected to be the case. Following this, a between-subgroup ANOVA was conducted with discrepancy score serving as the dependent variable. Due to restrictions with the data, separate ANOVAs were also conducted on the data from each testing, with subgroup (i.e., RAD, LAD, GAD) serving as the independent variable and mean discrepancy score, the dependent variable. [Note: The originally planned repeated measures ANOVA with the independent variables being (a) the assessment number (i.e., initial, 1 yr. follow-up, 2 year follow-up, etc.), and (b) subgroup (i.e., RAD, LAD, GAD); and mean discrepancy score serving as the dependent variable, were not carried out. This repeated measures design could not be employed due to limitations in consistency of patients returning for follow up visits, which resulted in empty cells and great variability in terms of inter-follow up assessment intervals.]

In further investigating progression patterns, the patient sample was cross-sectionally divided into severity groupings (i.e., stage levels) on the basis of CDR ratings. Correlations between initial discrepancy scores and CDR ratings were calculated for the subgroups. Following this, the mean discrepancy scores for each subgroup were compared across the severity levels (i.e., utilizing ANOVA). With respect to GAD, no substantial differences across stage were expected. In regard to RAD and LAD however, it was expected that the differences between semantic and visual-constructive functioning would be greatest for the mildly impaired patients, become less prominent for the moderately impaired cases, and evince the least discrepancy in the severely demented groups (i.e., the discrepancy was predicted to
diminish as severity of the disease increases).
CHAPTER IV

Results

Preliminary Examination of the Data

Initially, two correlation matrices (PxP) of the T scores from the six test variables (BNT, VF, PRAX, IVBM, DVBM, RECOG), one for the AD patients and one for the controls, were computed and examined. For the AD patients, forty percent of the coefficients were above .29, and each variable had appreciable correlations (i.e., \( r \geq .29 \)) with at least one other variable; 14 of the 15 coefficients were significant \( (p < .01) \). Examination of the matrix created from the control subjects' data revealed similar correlations between the variables; 13 of the 15 coefficients were significant at the .01 level. Bartlett's test of sphericity revealed that both correlation matrices were unlikely to be identity matrices \( (p < .01) \) (Norusis, 1993). As well, the anti-image correlation matrices generally had low coefficients, and the values of the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy were acceptable for both matrices (.70 for the patients; .69 for the controls). All of these findings suggested the appropriateness of the factor model (Norusis, 1993).

R-Factor Analyses

 Patients. The T score data from the AD patients on the 6 test variables were submitted to a principal components factor analysis using Varimax orthogonal rotation of the principal components solution (SPSS® FACTOR procedure, version 7.5). The criterion for factor extraction was presence of an eigenvalue greater than 1 (i.e., \( \text{Mineigen} = 1 \)); the latent root criterion; Afiﬁ & Clark, 1990; Hair et al., 1992; West, 1991). This criterion for selecting the appropriate number of factors was observed to correspond with that suggested by
examination of the scree plot, in addition to the commonly utilized rule of thumb regarding selection of only those components which explain at least 100 + N (N = the number of variables) percent of the total variance (100 + 6, or at least 16.7% in this sample) (Affi & Clark, 1991).

A two-component solution emerged, accounting for 57.3% of the total variance (see Table 2 for component loadings). The eigenvalues for the first and second components were 1.79 and 1.65, respectively. Variables loading highly on Factor I were as follows: DVBM, RECOG, IVBM. Those with significant loadings on Factor II included: PRAX, VF, BNT. These 2 factors resembled the memory and nonmemory factors identified recently by Larrain and Cimino (1998a) in their analysis of CERAD data from an independent patient sample. Thus, the dissociation between accessibility of semantic knowledge and visual-constructional functioning amongst AD patients reported originally by Morris et al. (1989) in their analysis of CERAD data, and later by Martin and colleagues (1986) and Fisher et al. (1996) on non-CERAD yet similar data, was not replicated on this AD sample. Nevertheless, it should be noted that inspection of the factor plot of the variables in rotated factor space did suggest 3 clusters of variables: (a) PRAX; (b) BNT and VF; (c) IVBM, DVBM, RECOG (see Figure 3).

Controls. When the T score data from the control group were submitted to an R-type factor analysis employing the same procedure and criteria outlined above for the patients, a two-component solution emerged, accounting for 57.4% of the total variance (see Table 3 for component loadings). The eigenvalues for the first and second components were 1.98 and 1.47, respectively. Variables loading highly on Factor I were as follows: DVBM, IVBM,
Table 2

**Principal Components R-Type Factor Loadings for the Patients (N = 960)**

<table>
<thead>
<tr>
<th>Test Variable</th>
<th>Factor I</th>
<th>Factor II</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNT</td>
<td>.227</td>
<td>.708</td>
</tr>
<tr>
<td>VF</td>
<td>.330</td>
<td>.729</td>
</tr>
<tr>
<td>PRAX</td>
<td>-.118</td>
<td>.740</td>
</tr>
<tr>
<td>IVBM</td>
<td>.615</td>
<td>.213</td>
</tr>
<tr>
<td>DVBM</td>
<td>.810</td>
<td>.150</td>
</tr>
<tr>
<td>RECOG</td>
<td>.761</td>
<td>-.019</td>
</tr>
</tbody>
</table>

*Note.* BNT = Boston Naming Test; VF = Verbal Fluency; PRAX = Constructional Praxis; IVBM = Immediate Verbal Memory; DVBM = Delayed Verbal Memory; RECOG = Verbal Recognition Memory.
Figure 3. R-Factor plot of patient $T$ score variables in rotated factor space. [BNT = Boston Naming Test; VF = Verbal Fluency; PRAX = Constructional Praxis; IVBM = Immediate Verbal Memory; DVBM = Delayed Verbal Memory; RECOG = Recognition Verbal Memory.]
Table 3

Principal Components R-Type Factor Loadings for the Controls \((N = 465)\)

<table>
<thead>
<tr>
<th>Test Variable</th>
<th>Factor I</th>
<th>Factor II</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNT</td>
<td>.117</td>
<td>.666</td>
</tr>
<tr>
<td>VF</td>
<td>.307</td>
<td>.615</td>
</tr>
<tr>
<td>PRAX</td>
<td>-.071</td>
<td>.716</td>
</tr>
<tr>
<td>IVBM</td>
<td>.784</td>
<td>.246</td>
</tr>
<tr>
<td>DVBM</td>
<td>.840</td>
<td>.247</td>
</tr>
<tr>
<td>RECOG</td>
<td>.738</td>
<td>-.115</td>
</tr>
</tbody>
</table>

Note. BNT = Boston Naming Test; VF = Verbal Fluency; PRAX = Constructional Praxis; IVBM = Immediate Verbal Memory; DVBM = Delayed Verbal Memory; RECOG = Verbal Recognition Memory.
RECOG. Those with significant loadings on Factor II included: PRAX, BNT, VF. These two factors were the same as those identified in the R-factor analysis of the patient data. However, the factor plot of the variables in rotated factor space did not suggest the same clusters for the controls as it did for the patients: PRAX, BNT, and VF appeared to form one cluster; IVBM, DVBM, and RECOG comprised a second cluster. The RECOG variable was removed enough from the IVBM and DVBM variables to suggest that it may even represent a third dimension of its own (see Figure 4).

**Q-Factor Analysis**

**Patients**

Q-type factor analysis involves the factoring of individuals rather than variables (i.e., as opposed to tests, as in the above R-method) (McKeown & Thomas, 1988). This type of analysis was accomplished by creating a correlation matrix in which the individual patients represented the rows, and the test variables the columns. The resulting matrix was then transposed, and correlation (i.e., similarity) coefficients were calculated between the profiles of each pair of subjects in the sample, creating a 960 x 960 matrix of intercorrelation (SPSS® procedure PROXIMITIES). Subsequently, this new matrix was submitted to a principal components factor analysis (SPSS® procedure FACTOR). Only those factors with eigenvalues ≥ the ratio of the number of subjects to the number of variables (i.e., 960 ÷ 6, or 160) were retained (Del Dotto & Rourke, 1985) and subjected to Varimax orthogonal rotation. This yielded three factors, accounting for 93.8% of the common variance. Patients were assigned to each subgroup on the basis of the factor for which they demonstrated a loading at or above .50. Subjects without at least one loading of .50 on a factor, in addition
**Figure 4.** R-Factor plot of control T score variables in rotated factor space. [BNT = Boston Naming Test; VF = Verbal Fluency; PRAX = Constructional Praxis; IVBM = Immediate Verbal Memory; DVBM = Delayed Verbal Memory; RECOG = Recognition Verbal Memory.]
to those with significant loadings (i.e., ≥ .50) on more than one factor, were not considered in the determination of subgroups. This assignment procedure resulted in the subgrouping of 720 patients, or 75% of the sample. Approximately forty-three percent of these subjects were assigned to Subgroup I (i.e., Factor I); Subgroups II and III were comprised of 34.3% and 22.4% of the reduced sample, respectively (see Table 4).

Following this, T score means from the six test variables used in the R- and Q-type factor analyses were computed for each subgroup and plotted graphically (see Figure 5). Examination of these profiles revealed that all three subgroups performed similarly on some of the measures. Namely, all three groups demonstrated a similar qualitative pattern of performance on the three memory measures employed in this investigation. In terms of levels of performance, the three groups were all severely impaired on the memory measures. Thus, with respect to both profile shape and levels of performance, it would appear that these memory variables have little or no utility as discriminators between the groups.

Despite their similar memory performances, the 3 subgroups did demonstrate qualitatively distinct patterns on the remaining variables (i.e., BNT, PRAX, VF). Subgroup 1 demonstrated impaired performance on BNT and VF, but performed within borderline-normal limits on PRAX. Subgroup 2 performed within normal limits on BNT, but evinced impairment on VF and PRAX. Members of Subgroup 3 were impaired on all three of these measures (i.e., BNT, VF, PRAX).

Multivariate analysis of variance (MANOVA) was undertaken to investigate whether the three groups differed significantly, utilizing the Q-defined subgroups as the independent
Table 4

Results of Q-Type Factor Analysis--Patients

<table>
<thead>
<tr>
<th>Loading(s)</th>
<th>Number of Patients</th>
<th>% of Entire Patient Sample (N = 960)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor I only</td>
<td>312</td>
<td>32.50</td>
</tr>
<tr>
<td>Factor II only</td>
<td>247</td>
<td>25.73</td>
</tr>
<tr>
<td>Factor III only</td>
<td>161</td>
<td>16.77</td>
</tr>
<tr>
<td>Factors I &amp; II</td>
<td>84</td>
<td>8.75</td>
</tr>
<tr>
<td>Factors I &amp; III</td>
<td>59</td>
<td>6.15</td>
</tr>
<tr>
<td>Factors II &amp; III</td>
<td>5</td>
<td>.52</td>
</tr>
<tr>
<td>Factors I, II, &amp; III</td>
<td>2</td>
<td>.21</td>
</tr>
<tr>
<td>No SL on any Factor</td>
<td>90</td>
<td>9.38</td>
</tr>
</tbody>
</table>

Note. SL = Significant Loading (i.e., ≥ .50).
Figure 5. Mean performance profiles of patient subgroups derived by Q-Factor analysis. [Q1 = Q-group 1 (n = 312); Q2 = Q-group 2 (n = 247); Q3 = Q-group 3 (n = 161); BNT = Boston Naming Test; VF = Verbal Fluency; PRAX = Constructional Praxis; IVBM = Immediate Verbal Memory; DVBM = Delayed Verbal Memory; RECOG = Recognition Verbal Memory.]
variable, and the six neuropsychological measures as the dependent variables. Utilizing Wilks' Lambda as the criterion, such analysis revealed that the groups were significantly different from each other on the combined dependent measures \( F(12, 1424) = 380.32, p < .001 \). A commonly employed equation for calculating the proportion of variance accounted for by this significant subgroup effect \( (\eta = 1-\lambda; \text{ Tabachnick & Fidell, 1989}) \), indicated that 94.3% of the variance in the best linear combination of test scores was accounted for by subgroup assignment. Subsequent univariate \( F \) tests\(^1\) revealed that the BNT \( F(2, 717) = 2120.06, p < .001 \), PRAX \( F(2, 717) = 401.40, p < .001 \), VF \( F(2, 717) = 75.98, p < .001 \), and DVBM \( F(2, 717) = 9.80, p < .001 \) variables were significant contributors to these overall group differences.

Because there were significant pooled within-group correlations among the test variables, a stepwise analysis was also performed, as suggested by Tabachnick and Fidell (1989)\(^4\). This involved an original univariate \( F \) test for the first variable (i.e., BNT), followed by analyses of covariance (i.e., a series of ANCOVAs) for each of the remaining variables, one at a time. Thus, the remaining five variables were evaluated as to the amount of new information (i.e., in terms of variance accounted for) they added to the combination of

---

\(^1\) Prior to such analysis, the graphic distributions of test data for each subgroup were examined to ensure the assumptions of MANOVA were adequately met. No major deviations from the assumptions that would invalidate the procedure were noted.

\(^2\) Due to the increased Type I error rate resulting from computation of multiple ANOVAs, more stringent \( \alpha \) levels were set via a Bonferroni type adjustment. Conservative \( \alpha \) values of .002 (.001667, to be precise; i.e., .01 \div 6) were assigned for each of the 6 variables, and the overall level considered necessary for a significant result for the set of dependent variables (i.e., .01) conformed to the following equation: \( \alpha = 1 - (1-\alpha_1)(1-\alpha_2)...(1-\alpha_6) \) (Tabachnick & Fidell, 1989).

\(^4\) The problem with performing multiple univariate \( F \) tests on correlated dependent variables involves the likelihood that correlated measures assess overlapping aspects of the same construct. Thus, the univariate \( F \)'s are not independent, making adjustment of the error rate difficult.
dependent variables already tested, controlling for the effects of all previously entered variables, in a manner analogous to hierarchical stepwise analysis of the independent variables in multiple regression. Retaining the previous α adjusted for inflated Type I error rate (i.e., .002), the Fs for the DVBM (p = .028) and IVBM (p = .048) measures failed to reach significance, indicating that these variables were not significant in explaining additional variance.

Subsequent profile analysis also indicated that the overall patterns of performance on the 6 tests were significantly different between the 3 groups [(Wilks' Λ = .066) F(10, 1426) = 411.88, p < .001]. Indeed, eta square (i.e., the strength of the association; \( \eta^2 = 1 - \text{Wilks' Lambda} \)) revealed that 93.4% of the variance about adjacent line segments of the profiles was accounted for by the varying shapes of the profiles.

**Controls**

A Q-factor analysis identical to that outlined above was next conducted on the control data. Retention of those factors with eigenvalues ≥ the ratio of the number of subjects to the number of variables (i.e., 465 ÷ 6, or 77.5) (Del Dotto & Rourke, 1985) yielded two factors, accounting for 76.8% of the common variance. The two emerging factors were subjected to Varimax orthogonal rotation. Controls were assigned to each subgroup on the basis of the factor for which they demonstrated a loading at or above .50. Those without at least one loading of .50 on a factor, in addition to those with significant loadings (i.e., ≥ .50) on both factors, were not considered in the determination of subgroups. This assignment procedure resulted in the subgrouping of 207 control subjects, or 44.5% of the sample. Seventy-seven percent of these subjects were assigned to Subgroup I; Subgroup II was comprised of 23%
of the reduced sample (see Table 5).

Following this, T score means from the six test variables used in the R- and Q-type factor analyses were computed for each Q subgroup (Q1 and Q2) in addition to the remaining controls not assigned to a group by the Q analysis (i.e., designated as group Q0) and plotted graphically (see Figure 6). Examination of these profiles revealed that the three groups generally performed within normal limits on the measures. However, Q1 demonstrated isolated impairment on IVBM.

Multivariate analysis of variance (MANOVA) was undertaken to investigate whether the two true Q groups (i.e., Q1 and Q2) differed significantly, utilizing the Q-defined subgroups as the independent variable, and the six neuropsychological measures as the dependent variables. Utilizing Wilks' Lambda as the criterion, such analysis revealed that the groups were significantly different from each other on the combined dependent measures \( F(6, 200) = 54.05, p < .001 \). The proportion of variance accounted for by this significant subgroup effect (\( \eta^2 = 1 - \lambda \)), indicated that 62\% of the variance in the best linear combination of test scores is accounted for by subgroup assignment (Tabachnick & Fidell, 1989).

Subsequent univariate F tests\(^5\) revealed that all of the test variables were significant contributors to these overall group differences: IVBM \( F(1, 205) = 70.98, p < .001 \), VF \( F(1, 205) = 24.93, p < .001 \), DVBM \( F(1, 205) = 23.20, p < .001 \), RECOG \( F(1, 205) = 22.57, p < .001 \), PRAX \( F(1, 205) = 19.37, p < .001 \), and BNT \( F(1, 205) = 18.11, p < .001 \). [Of note, BNT, the variable with the highest \( F \) for the patients, had the lowest \( F \) for

\(^5\) Due to the increased Type I error rate resulting from computation of multiple ANOVAs, more stringent \( \alpha \) levels were set via a Bonferroni type adjustment. Conservative \( \alpha \) values of .002 (i.e., .01/6) were assigned for each of the 6 variables.
Table 5

Results of O-Type Factor Analysis—Controls

<table>
<thead>
<tr>
<th>Loading(s)</th>
<th>Number of Subjects</th>
<th>% of Entire Control Sample (N = 465)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor I only</td>
<td>160</td>
<td>34.41</td>
</tr>
<tr>
<td>Factor II only</td>
<td>47</td>
<td>10.11</td>
</tr>
<tr>
<td>Factors I &amp; II</td>
<td>26</td>
<td>5.59</td>
</tr>
<tr>
<td>No SL on either Factor</td>
<td>232</td>
<td>49.89</td>
</tr>
</tbody>
</table>

Note. SL = Significant Loading (i.e., ≥ .50).
Figure 6. Mean performance profiles of normal control subgroups derived by Q-Factor analysis. [QO = non-classified controls (n = 255); Q1 = Q-group 1 (n = 160); Q2 = Q-group 2 (n = 47); BNT = Boston Naming Test; VF = Verbal Fluency; PRA = Constructional Praxis; IVBM = Immediate Verbal Memory, DVBM = Delayed Verbal Memory, RECOG = Recognition Verbal Memory.]
the controls. Also, the variable with the highest $F$ for the controls, IVBM, failed to reach univariate significance for the patients.] Because there were significant pooled within-group correlations among the test variables, a stepwise analysis was also performed, retaining the previous $\alpha$ adjusted for inflated Type I error rate. Similar results emerged, with all variables reaching significance at at least the .001 level.

Subsequent profile analysis indicated that the overall patterns of performance on the 6 tests were significantly different between the 2 groups [(Wilks' $\lambda = .415$) $F (5, 201) = 56.75, p < .001$]. Calculation of eta squared (i.e., the strength of the association; $\eta^2 = 1$-Wilks' Lambda) revealed that 58.5% of the variance about adjacent line segments of the profiles was accounted for by the varying shapes of the profiles.

When an additional Q-analysis of the control subjects’ data was undertaken, eliminating those with MMSEs < 25 (reducing the control sample $N$ to 455 from 465), in an attempt to make the control sample more stringently normal (Mitrushina et al., 1995), the same results obtained. That is, the Q-analysis again identified 2 subgroups, accounting for 77% of the variance. Subgroup I comprised 78% of the subjects. These subgroups highly resembled those derived on the entire control sample (see Figure 7). As such, the full control sample was retained in the subsequent cluster analyses.

**Cluster Analyses**

In order to evaluate the reliability of the Q-derived patient and control subgroups mentioned above, four different clustering algorithms were utilized on the same data, in an attempt to replicate identification of the Q-groups across various methods. Such replication is required, because cluster analysis often generates inconsistent results across different
Figure 7. Mean performance profiles of normal control subgroups (all included [Q1, Q2]; reduced sample with those having MMSE scores < 25 excluded [RQ1, RQ2] derived by Q-Factor analysis. [BNT = Boston Naming Test; VF = Verbal Fluency; PRAX = Constructional Praxis; IVBM = Immediate Verbal Memory; DVBM = Delayed Verbal Memory; RECOG = Recognition Verbal Memory.]
approaches (i.e., Q-analysis, hierarchical agglomerative methods, iterative partitioning procedures), algorithms (e.g., Ward's versus average linkage), and similarity measures (e.g., correlation versus Euclidean distance) (Affifi & Clark, 1990). In this manner, it was reasoned that if indeed subgroups exist, a design incorporating different approaches, algorithms, and similarity measures, should produce consistent results. Only then, can one be assured that results obtained are not simply an artifact of the particular methodology utilized (Campbell and Fiske, 1959).

Four widely used techniques were chosen, given their availability, and success in previous studies (see Blashfield & Aldenderfer, 1988; Everitt & Dunn, 1991): Ward's method (Ward, 1963), average linkage within groups (Sneath & Sokal, 1973), complete linkage (Sokal & Michener, 1958), and K-Means iterative partitioning (MacQueen, 1967). The correlation coefficient was selected as the similarity measure for the hierarchical analyses, whereas squared Euclidean distance was employed in the K-Means procedure. The SPSS® version 7.5 implementations of these methods were utilized, the hierarchical algorithms with procedure “CLUSTER,” and the K-Means iterative algorithm with the “QUICK CLUSTER” procedure. Data from all 960 patients were used in the patient analyses, and those from all 465 controls were used in the control analyses (i.e., outliers were not removed; these analyses were not performed on the Q-reduced samples).

Patients

Inspection of the amalgamation coefficients and dendrograms generated by the three hierarchical methods suggested the presence of distinct homogeneous clusters of patients. All three hierarchical agglomerative algorithms suggested a three group solution (Note: Even
rescaled to the maximum reduction SPSS® will output, the plots were too large to reproduce here).

The SPSS® K-Means procedure requires the researcher to select the number of clusters to be derived from the data. As the Q-analysis and the hierarchical procedures clearly suggested the presence of three groups, \( k \) was set at 3 for the iterative partitioning procedure. The Q-derived means of the six variables for each of the 3 subgroups were utilized as the initial cluster seeds (QINIT). When the procedure was re-run utilizing initial cluster seeds randomly selected by the program (RINIT), the results were identical (Subgroup 1 \( r_{\text{QINIT}} = 1.00 \); Subgroup 2 \( r_{\text{QINIT}} = 1.00 \); Subgroup 3 \( r_{\text{QINIT}} = 1.00 \)).

**Controls**

Inspection of the amalgamation coefficients and dendrograms generated by the Ward and average linkage hierarchical methods suggested the presence of 2 clusters of control subjects. The amalgamation coefficients and dendrogram for the complete linkage algorithm suggested either a two or three group solution.

Because two of the three hierarchical methods and the Q-factor analysis had clearly indicated 2 clusters, \( k \) was set at 2 for the iterative partitioning procedure. The means of the six variables from the 2 Q-derived subgroups were utilized as the initial cluster seeds (QINIT). When the procedure was re-run utilizing initial cluster seeds randomly selected by the program (RINIT), the results were identical (Subgroup 1 \( r_{\text{QINIT}} = 1.00 \); Subgroup 2 \( r_{\text{QINIT}} = 1.00 \)).

**Evaluation of Clustering Results**

For both the patient and control groups, the results of the five subgrouping analyses
were evaluated three ways, in a manner similar to that of previous well-designed clustering studies (Del Dotto & Rourke, 1985; Fuerst, Fisk, & Rourke, 1989): (a) across-method visual comparison of graphic profiles constructed by plotting the means of each subgroup on the six measures, (b) misclassification analysis, and (c) construction of a multiprofile-multimethod matrix (Campbell & Fiske, 1959).

**Visual Comparison of Mean Subgroup Profiles Generated by each Clustering Method**

**Patients.** For all 5 clustering methods, mean $T$ scores on each of the 6 measures were calculated for each subgroup. Mean profiles for each subgroup by each method were then plotted graphically (see Figure 8). Inspection of the mean subgroup profiles generated by each clustering method revealed that the same 3 qualitative patterns were apparent across the methods, with only minor occasional deviations. As such, the pattern exhibited by each subgroup was judged as highly similar across methods.

The average linkage and K-Means procedures produced profiles most closely resembling the Q-factor profiles. In comparison to the Q-profiles, Ward’s profiles were similar for Subgroups 1 and 2. However, Ward’s Subgroup 3 profile showed a RECOG score at about the same level as DVBM, with IVBM inferior to both these other memory scores. This deviates from the Q-profile for Subgroup three, which depicted RECOG performance as relatively equivalent to IVBM and inferior to DVBM. The complete linkage algorithm also generated a Subgroup 3 mean profile that varied from that produced by the Q analysis in a manner similar to Ward’s profile 3, yet more extreme, suggesting intact RECOG performance for Subgroup 3. Also, the complete linkage Subgroup 1 profile deviated from the Q-subgroup 1 profile in terms of depicting PRAX performance as roughly equivalent to VF. The complete
Figure 8. Comparative figures depicting mean patient subgroup profiles generated by each clustering method. (A) Average Linkage (ALWG); (B) Ward's; (C) Q-Factor Analysis; (D) Complete Linkage (COMPLT); (E) K-Means. [BNT = Boston Naming Test; VF = Verbal Fluency; PRAX = Constructional Praxis; IVBM = Immediate Verbal Memory; DVBM = Delayed Verbal Memory; RECOG = Recognition Verbal Memory.]
linkage profile for Subgroup 2 more closely resembled the Q-subgroup 2 profile.

**Controls.** The profiles for the control subgroups were generally similar across methods (see Figure 9). The three hierarchical methods and the K-Means procedure generated Subgroup 1 profiles highly similar to the Q-Factor derived Subgroup 1 profile. The Subgroup 2 profiles generated by these methods were similar to each other in their main deviation from the Q2 profile: they all showed peaks on the IVBM variable (Q2 clearly does not peak on this variable). The Q-factor subgroup profile graph (middle of Figure 9) also includes a mean profile plot for those (i.e., more than half of the control sample) not classified by the Q-Factor procedure (i.e., Q0). It would appear that the deviation of Subgroup 2 on IVBM shown by the hierarchical and K-Means algorithms was due to these non-Q classified subjects, who generally performed very well on IVBM. Thus, it would seem that the alternate clustering procedures classified the Q-outliers into the Q2 subgroup.

**Misclassification Analysis**

Misclassification analysis involved the use of the Q-defined groups as the criterion subgroups. Counts were made as to the number of subjects misclassified to the Q-factor groups by the other four clustering methods (see Table 6).

**Patients.** As can be seen in Table 6(a), misclassifications were evident for each of the methods. However, the vast majority of subjects (74-97%) were correctly classified by each method. The average-linkage-within-groups and Ward’s algorithms were most successful in correctly classifying the subjects. The complete linkage procedure performed the worst (although still, quite well) of all the algorithms.

**Controls.** Although misclassifications were evident for each of the methods, the
Figure 9. Comparative figures depicting mean control subgroup profiles generated by each clustering method. (A) Average Linkage (ALWG); (B) Ward's; (C) Q-Factor Analysis (Q1 = Q-Group 1; Q2 = Q-Group 2; Q0 = those not classified by the Q-analysis); (D) Complete Linkage (COMPLT); (E) K-Means. [BNT = Boston Naming Test; VF = Verbal Fluency; PRAX = Constructional Praxis; IVBM = Immediate Verbal Memory; DVBM = Delayed Verbal Memory; RECOG = Recognition Verbal Memory.]
Table 6

**Misclassification Analysis**

(a) Patients

<table>
<thead>
<tr>
<th>Clustering Method</th>
<th>Misclassifications</th>
<th>Total Misclassifications (N=720)</th>
<th>% Correctly Classified</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q1 (n=312)</td>
<td>Q2 (n=247)</td>
<td>Q3 (n=161)</td>
</tr>
<tr>
<td>ALWG</td>
<td>6</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Ward</td>
<td>59</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>Complete</td>
<td>34</td>
<td>0</td>
<td>153</td>
</tr>
<tr>
<td>K-Means</td>
<td>45</td>
<td>32</td>
<td>3</td>
</tr>
</tbody>
</table>

(b) Controls

<table>
<thead>
<tr>
<th>Clustering Method</th>
<th>Misclassifications</th>
<th>Total Misclassifications (N=207)</th>
<th>% Correctly Classified</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q1 (n=160)</td>
<td>Q2 (n=47)</td>
<td></td>
</tr>
<tr>
<td>ALWG</td>
<td>1</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>Ward</td>
<td>2</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>Complete</td>
<td>12</td>
<td>31</td>
<td>43</td>
</tr>
<tr>
<td>K-Means</td>
<td>16</td>
<td>21</td>
<td>37</td>
</tr>
</tbody>
</table>

**Note.** ALWG=Average Linkage Within Groups; Q1=Q-group 1; Q2=Q-group 2; Q3=Q-group 3.
majority of control subjects (79-94%) were correctly classified to the Q-criterion groups (see Table 6b). As was the case with the patients, Ward’s method and average-linkage-within-groups were most successful in classifying the control subjects to the appropriate Q-factor groups, and the complete linkage algorithm was least accurate.

**Multiprofile-Multimethod Matrices**

As a third, and perhaps most powerful means of assessing the consistency of the subgroups generated by each of the five clustering approaches, separate multiprofile-multimethod matrices of intercorrelations were constructed for the patient and control groups (see Tables 7 & 8). The multitrait-multimethod matrix (Campbell & Fiske, 1959) involves reliability and validity coefficients computed between several traits assessed by several methods. The validity coefficients represent the correlations obtained for the same trait across methods. In this manner, independent measures of the same construct are correlated. Obviously, the higher the correlation coefficients, the more valid the measures (and the construct). The matrix also contains correlations between different traits assessed by the same method, and correlations between different traits measured by different methods. Campbell and Fisk (1959) argued that in order to demonstrate satisfactory validity of a construct, one is obligated to demonstrate high correlations between the same traits measured by different methods (i.e., validity coefficients). As well, one must also demonstrate that these coefficients are higher than those between different traits measured by distinct measures, in addition to those between different traits measured by the same method.

**Patients.** Correlations computed between the same subgroup profiles across different methods were generally high, ranging from .33 to .99, with the majority (i.e., 25 out of 30)
<table>
<thead>
<tr>
<th></th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>K1</th>
<th>K2</th>
<th>K3</th>
<th>C1</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td></td>
<td>.40</td>
<td>.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note. \( Q_1 = \text{Q-Factor Subgroup 1}; Q_2 = \text{Q-Factor Subgroup 2}; Q_3 = \text{Q-Factor Subgroup 3}; W_1 = \text{Ward's Subgroup 1}; W_2 = \text{Ward's Subgroup 2}; W_3 = \text{Ward's Subgroup 3}; A_1 = \text{Average Linkage Subgroup 1}; A_2 = \text{Average Linkage Subgroup 2}; A_3 = \text{Average Linkage Subgroup 3}; K_1 = \text{K-Means Subgroup 1}; K_2 = \text{K-Means Subgroup 2}; K_3 = \text{K-Means Subgroup 3}; C_1 = \text{Complete Linkage Subgroup 1}; C_2 = \text{Complete Linkage Subgroup 2}; C_3 = \text{Complete Linkage Subgroup 3}. \)
Table 8

**Multiprofile-Multimethod Matrix: Controls**

<table>
<thead>
<tr>
<th></th>
<th>Q1</th>
<th>Q2</th>
<th>W1</th>
<th>W2</th>
<th>A1</th>
<th>A2</th>
<th>K1</th>
<th>K2</th>
<th>C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2</td>
<td></td>
<td>-.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1</td>
<td>.98</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W2</td>
<td>-.89</td>
<td>.42</td>
<td>-.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>.98</td>
<td>.04</td>
<td>.95</td>
<td>-.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>-.96</td>
<td>.23</td>
<td>-.98</td>
<td>.95</td>
<td>.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K1</td>
<td>.99</td>
<td>.02</td>
<td>.97</td>
<td>-.87</td>
<td>.99</td>
<td>.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K2</td>
<td>-.95</td>
<td>.25</td>
<td>-.98</td>
<td>.97</td>
<td>-.92</td>
<td>.99</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>.95</td>
<td>.26</td>
<td>.90</td>
<td>-.73</td>
<td>.95</td>
<td>-.85</td>
<td>.96</td>
<td>.84</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>-.98</td>
<td>-.01</td>
<td>-.97</td>
<td>.89</td>
<td>.95</td>
<td>.94</td>
<td>-.98</td>
<td>.95</td>
<td>-.95</td>
</tr>
</tbody>
</table>

**Note.** Q1 = Q-Factor Subgroup 1; Q2 = Q-Factor Subgroup 2; W1 = Ward’s Subgroup 1; W2 = Ward’s Subgroup 2; A1 = Average Linkage Subgroup 1; A2 = Average Linkage Subgroup 2; K1 = K-Means Subgroup 1; K2 = K-Means Subgroup 2; C1 = Complete Linkage Subgroup 1; C2 = Complete Linkage Subgroup 2.
exceeding r values of .90 ($M_r = .90$). In addition, the majority of heteroprofile-heteromethod and heteroprofile-homomethod coefficients were lower ($M_r = .38$).

**Controls.** Correlations between the same subgroup profiles generated by the different methods averaged .81. The heteroprofile-heteromethod and heteroprofile-homomethod coefficients were substantially lower ($M_r = -.73$).

**Inter-Subgroup Comparisons**

Given that the findings from the three methods evaluating the reliability of the Q-factor analysis derived subgroups strongly supported the predicted existence of three qualitatively distinct groups of patients, difference tests (i.e., ANOVAs; adjusted $\alpha = .001$) were performed between the groups on measures not employed in their derivation (Aldenderfer & Blashfield, 1984; Fletcher, 1985) (see Table 9). The groups did not differ significantly in terms of duration of illness [$F(2, 707) = .801, p = .445$]. However, significant between-subgroup differences were detected on the following variables: age [$F(2, 717) = 52.52, p < .001$], education [$F(2, 717) = 15.43, p < .001$], MMSE [$F(2, 693) = 81.58, p < .001$], CDR [$F(2, 717) = 42.08, p < .001$], Blessed [$F(2, 716) = 34.19, p < .001$].

Subsequently, multiple pairwise comparisons were conducted for the variables which yielded significant differences between the groups via the univariate $F$ tests, in order to determine specifically which of the three groups differed from each other. The Games-Howell procedure was selected as an appropriately conservative method for these multiple comparisons, due to the unequal subgroup sizes and occasional mildly heterogeneous variances (Howell, 1992; Jaccard, Becker, & Wood, 1984). The Games-Howell multiple comparisons indicated significant subgroup differences for all of these variables (see Table 9).
Table 9

Means and Standard Deviations for the O-Factor Derived Patient Subgroups on Measures not Employed in their Derivation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Subgroup 1 $M(SD)$</th>
<th>Subgroup 2 $M(SD)$</th>
<th>Subgroup 3 $M(SD)$</th>
<th>Significant ($p &lt; .05$) Paired Comparisons†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age*</td>
<td>75.12(6.68)</td>
<td>70.38(7.51)</td>
<td>68.43(8.42)</td>
<td>1 &amp; 2; 1 &amp; 3</td>
</tr>
<tr>
<td>DOI</td>
<td>4.25(2.65)</td>
<td>4.22(2.76)</td>
<td>4.54(2.48)</td>
<td>--</td>
</tr>
<tr>
<td>Education*</td>
<td>11.48(3.99)</td>
<td>12.85(3.41)</td>
<td>13.14(3.00)</td>
<td>1 &amp; 2; 1 &amp; 3</td>
</tr>
<tr>
<td>MMSE*</td>
<td>16.84(5.67)</td>
<td>19.87(4.28)</td>
<td>12.97(5.81)</td>
<td>1 &amp; 2; 1 &amp; 3; 2 &amp; 3</td>
</tr>
<tr>
<td>CDR*</td>
<td>1.52(0.67)</td>
<td>1.24(0.50)</td>
<td>1.81(0.69)</td>
<td>1 &amp; 2; 1 &amp; 3; 2 &amp; 3</td>
</tr>
<tr>
<td>Blessed*</td>
<td>4.70(2.55)</td>
<td>3.80(2.01)</td>
<td>5.81(2.66)</td>
<td>1 &amp; 2; 1 &amp; 3; 2 &amp; 3</td>
</tr>
</tbody>
</table>

**Note.** DOI = Duration since onset of illness; MMSE = Mini Mental State Exam; CDR = Clinical Dementia Rating.

†Games-Howell test.

*indicates significant ANOVA ($p < .001$).
Members of Subgroup 1 were older and less educated than those in Subgroups 2 and 3. In terms of the dementia severity measures, those in Subgroup 3 were most impaired, followed next by those in Subgroup 1; those in Subgroup 2 were least impaired on these global measures.

The control subgroups (i.e., Q1 and Q2) did not differ in terms of age \( t(205) = -.628, p = .53 \), education \( t(205) = 1.29, p = .20 \), or MMSE scores \( t(205) = .08, p = .93 \).

**Longitudinal Analyses of Patient Data**

Longitudinal analyses of the patient sample included the following: (a) examination of 5 randomly selected cases from each subgroup with regard to stability of semantic versus visual-constructional discrepancy; (b) computation of correlations between discrepancy scores and testing number (i.e., initial testing, first follow-up, second follow-up, etc.); (c) conduction of a two-way ANOVA with subgroup and assessment number serving as the independent variables and discrepancy score serving as the dependent variable; (d) conduction of separate ANOVAs on the data from each testing, with subgroup serving as the independent variable and discrepancy score the dependent variable.

**Stability of Profiles**

In selecting patients from each Q-subgroup for longitudinal profile stability analysis, the initial Q-sample \( (N = 720) \) was reduced to include only those patients with 3 or more follow-up assessments. This resulted in a reduced total sample comprised of 113 subjects \( (Q1n = 37; Q2n = 66; Q3n = 10) \). Five cases were randomly selected from each of the reduced subgroup samples, and the BNT and PRAX scores were plotted graphically for each
testing session (see Figures 10-12)⁶.

The longitudinal discrepancy profiles for Q1 were generally indicative of stability: four out of the five clearly suggested consistently superior PRAX performance (i.e., normal range) as compared to BNT performance (impaired range), especially in the earlier years (see Figure 10). The longitudinal profiles for the Q2 group were less stable, with only three out of the 5 cases showing stability of the initial profile across time (see Figure 11). Indeed, one of the patients appeared to show a reversal of their profile. Profiles for the Q3 group were, in the main, stable (see Figure 12).

**Discrepancy Score Changes with Time**

In further investigating BNT/PRAX discrepancies across time, discrepancy scores were calculated for each of the Q-classified subjects. In order to eliminate negative values and control the direction of the discrepancy (i.e., in the predicted direction) for Subgroups 1 and 2, adjustments were made in calculating this score for each subgroup. For Subgroup 1, this score was calculated by subtracting the patient’s BNT T score from their PRAX T score. For Subgroup 2, the discrepancy score was computed by subtracting the PRAX T score from the BNT T score. For the third subgroup, the discrepancy score was the absolute value of the BNT T score minus the PRAX T score. These scores were computed for every subject who appeared at each testing interval.

A Pearson product moment correlation was calculated between discrepancy scores and testing number, revealing a significant negative correlation between these two variables.

---

⁶ The BNT and PRAX scores were used in examining semantic accessibility versus visual-constructional discrepancies in these and subsequent longitudinal analyses as they have been suggested by past research as the most sensitive to such domain-specific differences (Fisher et al., 1996; Fisher et al., 1997). In addition, these variables had the highest univariate Fs following the Q-MANOVA in the current investigation.
Figure 10. Histograms depicting BNT/PRAX T Score discrepancies across time for 5 randomly selected Q1 patients. [BNT = Boston Naming Test; PRAX = Constructional Praxis; Q1 = Q-Factor Derived Subgroup 1].
Figure 11. Histograms depicting BNT/PRAX T Score discrepancies across time for 5 randomly selected Q2 patients. [BNT = Boston Naming Test; PRAX = Constructional Praxis; Q2 = Q-Factor Derived Subgroup 2].
Figure 12. Histograms depicting BNT/PRAZ T Score discrepancies across time for 5 randomly selected Q3 patients. [BNT = Boston Naming Test; PRAZ = Constructional Praxis; Q3 = Q-Factor Derived Subgroup 3].
for the patient sample as a whole \( r = -.31, p < .01; n = 1588 \). To detect between-subgroup differences, mean BNT/PRAX discrepancies at each assessment were calculated for each subgroup, and plotted graphically (see Figure 13). As diminishing discrepancy scores across time were suggested by these plots, especially for Subgroups 1 and 2, separate Pearson correlations between testing number and discrepancy scores were calculated for each subgroup. These coefficients revealed significant negative correlations for Subgroup 1 \( r = -.36, p < .01; n = 647 \) and Subgroup 2 \( r = -.34, p < .01; n = 656 \), and a nonsignificant correlation for Subgroup 3 \( r = .05, p = .43; n = 284 \).

The findings of negative mean discrepancy scores for Subgroup 2 (from follow-up testings 4 through 7; see Figure 13) suggested a reversal of the initial BNT > PRAX pattern with time. To investigate this further, mean BNT and PRAX scores were plotted by testing number for this subgroup, and for the other subgroups for comparative purposes (see Figures 14-16). Inspection of Figure 14, revealed the expected consistency of PRAX > BNT over time, even at the 7th year follow up assessment, for the Q1 group. Similarly, Q3 (Figure 15) showed little discrepancy in terms of BNT and PRAX performances across testings. However, as suggested by the negative discrepancy scores obtained after follow up testing number 3, Q2 did evince little stability of the initial BNT > PRAX pattern (see Figure 16). At follow up assessment #1, this subgroup’s mean performance on the 2 measures appeared to equalize until the fourth year follow up assessment, at which time mean PRAX scores began to exceed mean BNT scores, although both performances were clearly within severely impaired limits. It should be noted that the number of subjects decreased markedly with each follow up testing, and as such, these results should be interpreted cautiously.
Figure 13. BNT–PRAX Discrepancy T Scores for the patients across time.
Figure 14. Mean BNT and PRAX T scores across testings for the Q1 group.
Figure 15. Mean BNT and PRAX T scores across testings for the Q3 group.
Figure 16. Mean BNT and PRAX T scores across testings for the Q2 group.
A follow up 3x8 factorial (two-way) ANOVA with subgroup and assessment number serving as the independent variables, and discrepancy score serving as the dependent variable was next conducted. This analysis revealed significant main effects for subgroup \( F(2, 1563) = 76.7, p < .001 \) and assessment number \( F(7, 1563) = 13.05, p < .001 \), in addition to a significant two-way interaction effect between the two independent variables \( F(14, 1563) = 5.4, p < .001 \). While this analysis violated the ANOVA assumption of independent groups, the appropriate repeated measures ANOVA design could not be employed due to limitations in consistency of patients returning for follow-up visits, which resulted in empty cells, and great variability in terms of inter-follow up assessment intervals. Nevertheless, the above methodologically suboptimal analysis suggested that any differences between discrepancy scores at different assessment dates should be evaluated separately for each of the Q-derived subgroups.

Given the above limitations, separate ANOVAs were conducted on the data from each testing, with subgroup (i.e., 1, 2, 3) serving as the independent variable and mean discrepancy score serving as the dependent variable. These analyses revealed significant between-group differences in discrepancy scores for the initial assessment \( F(2, 717) = 372.96, p < .001 \), the first year follow up \( F(2, 277) = 55.87, p < .001 \), in addition to the 2nd \( F(2, 279) = 31.65, p < .001 \), 3rd \( F(2, 132) = 6.25, p < .003 \), 4th \( F(2, 74) = 18.66, p < .001 \), and fifth year \( F(2, 46) = 14.0, p < .001 \) follow up assessments. However, between-subgroup differences in discrepancy scores from the 6th \( F(2, 25) = 2.51, p = .101 \) and 7th \( F(2, 14) = 2.24, p = .144 \) year follow up assessments were not significant (see Table 10).

Post hoc multiple comparisons (i.e., Games-Howell) revealed that for the initial
assessment, significant differences between all three subgroups were apparent, with Subgroup 1 showing the most discrepant scores, followed by Subgroup 2, and finally, Subgroup 3. At the first, 2nd and 3rd follow up testings, significant differences were apparent between Subgroups 1 and 2, and Subgroups 1 and 3, but not between Subgroups 2 and 3. At the fourth year follow up testing all the groups' discrepancy scores differed significantly from each other but the difference between Subgroups 2 and 3 was due to Subgroup 2's negative discrepancy scores (i.e., reversal effect). At this testing the discrepancy score of Subgroup 1 was significantly greater than that of subgroups 2 or 3. Following this fourth follow up testing, the groups did not differ with regard to their discrepancy scores (i.e., at follow up assessments 5, 6, and 7) (see Table 10 & Figure 13).

Discrepancy Score Changes as a Function of Dementia Severity/Stage

To examine the stability of discrepancy scores across stage of AD, the following cross-sectional analyses were conducted: (a) computation of correlations between discrepancy scores and CDR ratings; (b) computation of ANOVAs for each subgroup, with CDR scores serving as the independent variable, and discrepancy scores the dependent variable.

The patient sample was first divided into severity groupings (i.e., stage levels) on the basis of CDR ratings. A Pearson product moment correlation computed for the entire Q-sample \(N = 720\) between CDR stage and initial BNT/PRAX discrepancy scores (i.e., from the initial assessment), suggested no relationship between these two variables \(r = -.06, p = .14; n = 720\).

Following this, separate Pearson correlations between CDR ratings and initial discrepancy scores were calculated for each subgroup. These coefficients revealed significant
Table 10

Mean BNT/PRA X Discrepancy T Scores across Assessments

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Subgroup 1 M(SD)</th>
<th>Subgroup 2 M(SD)</th>
<th>Subgroup 3 M(SD)</th>
<th>Significant (p &lt; .05) Paired Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial*</td>
<td>35.75(12.19)</td>
<td>14.60(17.82)</td>
<td>2.63(4.21)</td>
<td>1 &amp; 2, 1 &amp; 3, 2 &amp; 3</td>
</tr>
<tr>
<td>1 year*</td>
<td>26.94(17.20)</td>
<td>2.10(21.47)</td>
<td>8.57(14.73)</td>
<td>1 &amp; 2, 1 &amp; 3</td>
</tr>
<tr>
<td>2 years*</td>
<td>23.85(20.72)</td>
<td>1.24(25.06)</td>
<td>7.41(14.34)</td>
<td>1 &amp; 2, 1 &amp; 3</td>
</tr>
<tr>
<td>3 years*</td>
<td>13.92(16.24)</td>
<td>2.26(20.35)</td>
<td>2.77(8.24)</td>
<td>1 &amp; 2, 1 &amp; 3</td>
</tr>
<tr>
<td>4 years*</td>
<td>23.77(19.92)</td>
<td>-7.63(20.21)</td>
<td>2.71(5.72)</td>
<td>1 &amp; 2, 1 &amp; 3, 2 &amp; 3</td>
</tr>
<tr>
<td>5 years*</td>
<td>23.12(16.93)</td>
<td>-8.34(20.52)</td>
<td>0(0)</td>
<td>N.S.</td>
</tr>
<tr>
<td>6 years</td>
<td>13.10(24.71)</td>
<td>-6.80(17.89)</td>
<td>0(0)</td>
<td>--</td>
</tr>
<tr>
<td>7 years</td>
<td>8.51(20.84)</td>
<td>-10.11(14.91)</td>
<td>0(0)</td>
<td>--</td>
</tr>
</tbody>
</table>

Note. BNT = Boston Naming Test; PRAX = Constructional Praxis; N.S. = None Significant. *indicates significant (p < .005) between-subgroup ANOVA difference.
(yet modest) negative correlations for Subgroups 1 \([r = -.18, p < .01]; n = 312\) and 3 \([r = -.24, p < .01]; n = 161\), and a significant positive correlation for Subgroup 2 \([r = .23, p < .01]; n = 247\). This suggests that with more severe dementia in terms of stage, discrepancy scores decline for Subgroups 1 and 3, but not for Subgroup 2.

Next, ANOVAs comparing mean discrepancy scores across severity levels (i.e., CDR stages) for each subgroup were conducted. The analysis was significant for the Q1 subgroup \([F (3, 307) = 4.11, p < .05]\). A graphical plot of mean discrepancy score by CDR stage suggested that this subgroup displays a trend for consistently decreasing discrepancy scores across stages 1 (mild dementia) through 3 (severe dementia) (see Figure 17). However, post-hoc comparisons revealed the decrease in discrepancy scores to be significant only between CDR stages 1 and 3 (i.e., not between stages 1 and 2 or 2 and 3). Significant Fs were also obtained for Q2 \([F (3, 243) = 4.81, p < .01]\) and Q3 \([F (2, 156) = 4.23, p < .05]\). Examination of discrepancy by stage plots (Figure 17) suggested that the discrepancy scores of Subgroup 2 increased with CDR stages of dementia. Paired comparisons revealed that these increases were significant \((p < .05)\) between stages .5 and 2, .5 and 3, 1 and 2, and 1 and 3, but not between .5 and 1 or between 2 and 3. With respect to Subgroup 3, discrepancy scores decreased slightly with advancing CDR stage (see Figure 17). Post-hoc comparisons revealed that these declines were significant \((p < .05)\) only between stages 1 and 3.
Figure 17. Plots of initial BNT/PRAX mean discrepancy scores by CDR stage for each subgroup. [BNT = Boston Naming Test; PRAX = Constructional Praxis; Q1 = Subgroup 1; Q2 = Subgroup 2; Q3 = Subgroup 3. CDR .5 = questionable dementia; CDR 1 = mildly demented; CDR 2 = moderately demented; CDR 3 = severely demented.]
CHAPTER V

Discussion

The current investigation involved the utilization of factor and cluster analytic classification techniques in a large scale examination of neuropsychological profiles obtained from patients diagnosed with probable AD. The primary goal of this research was to replicate the results of an earlier study, in which three neuropsychological subgroups of AD patients were reliably identified across clustering methods (Fisher et al., 1996). The current study employed methodological improvements over the previous one, by inclusion of a demographically similar healthy (i.e., non-demented) comparison group, utilization of a much larger, multicenter sample (i.e., CERAD), and incorporation of a longitudinal design.

Inclusion of a non-demented comparison group provided a means for determination of whether any identified patient subgroups were indeed specific to AD (i.e., apart from the normal cognitive heterogeneity of aging). The multicenter CERAD sample, the largest and most geographically representative of all studied to date, was utilized to counteract the statistical limitations of small n’s employed in previous subgrouping investigations. The longitudinal design allowed for evaluation of an earlier proposed subgroup-specific neuropsychological model of AD (Fisher, Rourke, & Bieliaskas, 1997). This model incorporates both the Subgroup and Stage Models of the disease, in outlining progression patterns for each of three neuropsychological subgroups.

Summary of Results

Subtyping Results

The results of this study demonstrated that subgroupings of patients diagnosed with
mild to moderate probable AD, each displaying qualitatively distinct patterns of performance on CERAD neuropsychological measures, can be identified via Q-type factor analysis. These Q-factor subgroups were replicated via three different hierarchical agglomerative clustering algorithms, in addition to a non-hierarchical iterative partitioning technique, establishing their reliability.

In accordance with previous research, three patient subgroups were identified. All three groups displayed a similar pattern of performance on immediate (free recall), delayed (free recall), and recognition trials of a verbal list learning task. However, clinically meaningful distinctions between the subgroup patterns on other measures were evident. Overall, Subgroup 2 appeared as the highest functioning of the three groupings. This subgroup demonstrated preserved naming abilities, although encountered mild to moderate difficulty copying figures on the constructional praxis measure. In contrast, Subgroup 1 demonstrated severe anamia, in the context of relatively spared (i.e., borderline-normal) visual-constructional functioning. The performance of Subgroup 3, the most impaired group, was marked by severe anamia and constructional dyspraxia. These subgroups resembled the hypothetical groups RAD, LAD, and GAD, respectively. Those in the RAD subgroup performed relatively better than those in the LAD and GAD groups on the Verbal Fluency test, although the mean performances of all 3 subgroups were clearly impaired on this measure. The naming and constructional praxis measures best discriminated the subgroups, in terms of both statistical and clinical significance.

Less than half of the control sample was classified into subgroups by Q-factor analysis, a largely unsatisfactory result. Two control subgroups emerged from the Q-analysis and 2 of
the 3 hierarchical agglomerative cluster analytic replication attempts. These subgroups did not resemble those of the patients, each demonstrating a flat profile across the BNT, PRAX, and VF variables. Those in Subgroup 1 performed within normal limits on all measures save for demonstrating isolated impairment on the immediate recall trials of the verbal list learning task. This subgroup appeared to perform relatively better on the visual constructional and semantic tasks compared to their performance on the memory measures. Subgroup 2 was characterized by a relatively flat profile, with slightly better performance on the memory measures compared to the semantic and visual-constructional variables. Nevertheless, all performances were within normal limits. When the group means on the 6 test variables were calculated for those control subjects not classified by the Q-analysis, average to high average levels of performance were observed on all measures. These mean scores yielded a flat profile, save for a prominent peak on the immediate verbal memory measure.

**Inter-Subgroup Comparisons**

There were no differences between the 3 patient subgroups in terms of duration of illness. However, Subgroup 1 (i.e., LAD) was older and less educated than Subgroups 2 and 3. As well, Subgroup 3 (i.e., GAD) was more advanced in terms of CDR stage, dementia severity (i.e., MMSE scores) and functional (ADL) impairment (i.e., as assessed by Blessed scores) than Subgroup 1 (LAD). In turn, Subgroup 1 (LAD) was more advanced in terms of CDR stage, and demonstrated greater impairment with respect to functional activities of daily living and MMSE scores compared to Subgroup 2 (RAD). The control subgroups did not differ with respect to age, education, or MMSE scores.
**Longitudinal Results**

Qualitative examination of the semantic (i.e., naming) and visual-constructional performances of randomly selected cases from each subgroup suggested stability of the Subgroup 1 (LAD) and 3 (GAD) profiles over time. Subgroup 2 (RAD) demonstrated less stability across assessments, although many of the cases did demonstrate consistency of their initial pattern.

Correlational analyses suggested that constructional-semantic discrepancies across time decreased slightly for the LAD subgroup, but did not decrease for the GAD subgroup. The RAD subgroup displayed a dynamic relationship with respect to semantic-constructional discrepancy scores across testing sessions. While these scores did decrease up to the 3rd follow-up testing, following that time, a reversal effect was observed, in which this subgroup performed slightly better on the visual-constructional compared to the semantic measure. Discrepancies between performance in the visual constructional versus semantic realms of functioning were greatest for the LAD subgroup and least for the GAD subgroup at the initial testing. Discrepancies remained greatest for the LAD group through to the fourth year follow up assessment. After this time, the groups did not differ in terms of their discrepancy scores. Following the initial testing, the RAD and GAD subgroups basically did not differ in terms of their discrepancy scores.

**CDR Stage Findings**

Cross-sectional evaluation of changes in discrepancy scores across CDR stage revealed a tendency for Subgroups 1 and 3 (i.e., LAD and GAD) to display declining discrepancy scores with increasing stage of dementia. However, Subgroup 2 (RAD) displayed
increasing discrepancy scores with more advanced stage of dementia.

**Evaluation of Hypotheses**

The first hypothesis of this study predicted that the current sample of patients with *probable* AD could be reliably differentiated into 3 empirically derived neuropsychological subgroups. This prediction was based on past research which has reliably demonstrated that when neuropsychological data from *probable* AD patients are subjected to empirical classification techniques, 3 qualitatively distinct subgroups of patients emerge (Fisher et al., 1996; Martin et al., 1986). This hypothesis was fully supported: the Q-factor analysis and the hierarchical techniques all clearly indicated 3 subgroups of patients. Furthermore, when the K-Means procedure was run with $k$ set at 3, the resulting subgroups resembled those produced by the Q-factor analysis and the hierarchical agglomerative techniques.

It was further hypothesized (i.e., hypothesis 2) that the 3 patient subgroups would be distinguishable on the basis of performance in two domains of neuropsychological functioning: visual-constructional ability and semantic knowledge accessibility. In this regard, it was expected that in accord with past research (Fisher et al., 1996; Martin et al., 1986), the three subgroups of AD patients would resemble the theoretical ideal types (Fisher, Rourke, & Bieliauskas, 1997) (a) GAD, (b) LAD, and (c) RAD, respectively: (a) those with global impairment (i.e., impaired access to semantic knowledge and impaired visual-constructional functioning); (b) those with impaired ability to access semantic knowledge, but relatively unimpaired visual-constructional functioning; (c) those with relatively spared accessibility of semantic knowledge but impaired visual-constructional functioning.

This second hypothesis was also supported; the three subgroups did indeed resemble
the hypothetical subtypes outlined above with respect to their performance patterns on visual-constructional and confrontational naming measures. Subgroup 1, the largest subgroup, resembled LAD. Patients in this group earned a mean naming T score of 4 (severely impaired) and a mean T score of 40 (borderline-normal) on the constructional praxis measure. Subgroup 2, the second largest group, earned T scores of 45 (average) and 30 (mildly-moderately impaired), on the naming and constructional measures respectively, displaying a pattern consistent with RAD. The third subgroup clearly resembled GAD, earning mean T scores of 2 on both measures, indicating severe impairment in both the naming and visual-constructional domains.

The RAD subgroup performed significantly better than the other two subgroups on the Verbal Fluency (i.e., Animal name generation task), which is not inconsistent with hypothesis two. However, all 3 subgroups displayed impairment on this measure: The GAD and LAD groups exhibited moderately to severely impaired scores; the RAD group displayed mildly to moderately impaired performance on this measure. This finding of impairment across the subgroups on animal name generation accords with those of previous studies (e.g., Fisher et al., 1996). It appears that performance on this task becomes quickly impaired for members of all AD subgroups, perhaps for different neuropsychological reasons (see Fisher et al., 1997, for discussion).

Hypothesis three predicted that the non-demented comparison group employed in this investigation would not display neuropsychological patterns resembling those of the LAD and RAD patient subgroups. It was expected that normal elderly might show different patterns of relative strengths and weaknesses (e.g., on memory measures; Mitrushina et al., 1995).
However, it was predicted that their cognitive profiles would not evince RAD and LAD patterns; that is, their level of performance in the visual-constructional and semantic knowledge realms would remain within the same range of functioning (Mitrushina et al., 1995).

Hypothesis three was generally supported. Unlike that with the patient sample, the Q-analysis of the control data was largely unsuccessful, classifying less than half the sample into two subgroups. These two subgroups did not resemble those derived from the patient data. With respect to the semantic and visual constructional measures, the control subgroups displayed patterns of flat lines, revealing equivalent (and normal) levels of performance in the two domains. The flat lines ran across all three non-memory measures (e.g., BNT, VF, PRA), rendering the control profiles dissimilar even to the GAD profile. (The GAD profile had a peak on VF).

The comparison group profiles on the memory measures were not flat, with both control subgroups performing best on recognition, then delayed recall, and least well on the immediate recall measure. This pattern is inconsistent with that observed for the three patient subgroups, in which delayed verbal recall was slightly better than immediate recall and recognition performance, although all three were severely impaired. It should be noted that control Subgroup 1 performed within the impaired range on the immediate recall measure, an unexpected finding. Aside from this isolated impairment of Subgroup 1 on the immediate recall task, both control subgroups performed within normal limits on the memory measures.

The final hypothesis concerned the specific progression patterns of AD for each subgroup. It was predicted that, in the early and middle period of the disease, the initial
impairment pattern of each subgroup would remain consistent. This hypothesis arose from past AD subgrouping research which has suggested that the initially presenting neuropsychological pattern (i.e., GAD, RAD, or LAD) remains detectable across time, during the early and middle period of the disease (Fisher et al., 1997; Martin, 1990). As such, it was expected that those with the GAD pattern would continue to show impairment in both the semantic knowledge and visual-constructional realms of functioning, consistently exhibiting a pattern of relatively equivalent decline in the two domains. Those with RAD were expected to continue to show sparing of semantic knowledge accessibility for a period of time, and though impairment in this domain was theorized to eventually occur, it was predicted to be relatively less severe than that of visual-constructional functioning until late in the disease, when all functioning is profoundly disrupted.

The prediction for LAD patients was the opposite of that for the RAD group. Those classified with LAD were expected to initially display visual-constructional functioning within normal limits, in the face of difficulty in the area of semantic knowledge accessibility. While, with time and progression of the disease, visual-constructional functioning was expected to eventually decline into the impaired range, it was predicted to continue to be relatively less severely affected than word finding ability, until the latest period of the disease. Indeed, it was expected that late in the disease course, all three subgroups would be rendered indistinguishably globally impaired, as floors effects of the neuropsychological measures were reached. Operationally, it was predicted that longitudinal analyses of performance discrepancies between semantic and visual-constructional functioning would reveal that such discrepancies diminish over time for the RAD and LAD groups only, but remain relatively
consistent for the GAD group. Similarly, it was expected that cross-sectional analyses of discrepancies between the two domains for each CDR stratification would reveal stable discrepancies across stage for the GAD subgroup, and declining discrepancies for the RAD and LAD subgroups with successive stages.

This set of hypotheses was largely supported for the GAD and LAD groups. The performance of the GAD subgroup on the semantic and visual-constructional measures (i.e., BNT and PRAX) remained consistent across testing sessions. This was apparent upon examination of the performances of 5 randomly selected GAD patients across time, in addition to inspection of the mean naming and constructional performance patterns of this subgroup at each testing session. As well, correlations computed between discrepancy scores and assessment number were not significant for this subgroup, suggesting minimal change over time from the initial pattern. With respect to the cross sectional analysis of discrepancy scores across CDR stage, the GAD subgroup showed a significant yet slight decrease in discrepancy scores with advanced stage. Post-hoc inspection of these data suggests that this finding was contributed to by floor effects.

The discrepancy findings for the LAD subgroup were in accord with those predicted. Examination of the individual profiles from the 5 randomly selected subgroup members revealed consistent PRAX > BNT patterns across assessment intervals. The mean performance patterns of this subgroup at each testing session also displayed consistent patterns of relatively superior constructional as opposed to semantic functioning. Correlations computed between assessment number and discrepancy scores were significantly negative for this subgroup, suggesting declining discrepancies with progression of the disease. With
respect to the cross sectional analysis of discrepancy scores across CDR stage, the LAD subgroup showed a significant yet slight decrease in discrepancy scores with advanced stage.

The longitudinal and cross-sectional discrepancy findings for the RAD group were generally inconsistent with those predicted. Examination of the individual profiles from the 5 randomly selected subgroup members revealed consistent BNT > PRAX patterns across assessment intervals for 3 of the 5 subjects. However, only two of the five subjects retained their initially normal level of naming performance at the first year follow up assessment. This finding is in distinct contrast to that of the LAD subgroup; 80% of the randomly selected LAD cases retained their PRAX performance within normal limits for at least one year following the initial assessment.

Examination of the mean performances of the RAD subgroup at each testing session revealed consistent patterns of slightly superior semantic as opposed to constructional functioning up until follow up assessment 4. After the fourth assessment, the scores of this subgroup demonstrated a reversal effect, with PRAX performance becoming superior to BNT. Correlations computed between assessment number and discrepancy scores were significantly negative for this subgroup, but were confounded by negative discrepancy scores following the fourth assessment. Thus while the discrepancy scores did decrease in the expected direction up until the fourth assessment, after this testing, negative scores were obtained, as mean PRAX performance became superior to mean BNT performance. With respect to the cross sectional analysis of discrepancy scores across CDR stage, the RAD subgroup displayed an increase in discrepancy scores with advanced stage, a clearly unexpected finding.
Compared to RAD and GAD, the LAD subgroup had a significantly larger initial discrepancy score and maintained this initial discrepancy pattern across time, up until the 5th follow up assessment. RAD on the other hand, had a mean initial discrepancy score falling between that of LAD and GAD. However, the RAD group did not retain this midway status. From assessment 2 onward, the discrepancy scores of the RAD group (when in the expected direction) were not significantly different from those of the GAD subgroup. This finding was in stark contradiction to that predicted for the RAD group, as members of this group (as LAD) were expected to retain their initial pattern of discrepancy for a period of time. Clearly, they were not expected to show comparable discrepancy scores to the GAD group. At the 6th and 7th year follow-up examinations, there were no between subgroup differences in discrepancy scores. This is in accord with the prediction that eventually all 3 groups would become indistinguishable, as global impairment is reached in the final phases of the disease.

In sum, it appears that whereas the longitudinal findings for the LAD and GAD subgroups were in accord with prediction, the progression patterns of the RAD subgroup did not conform to those expected (Fisher, Rourke, & Bieliauskas, 1997). Nevertheless, it should be pointed out that some members of the RAD subgroup did indeed retain their presenting pattern of normal word finding ability within the context of impaired visual-construction across time, and did not demonstrate rapid decline in the initially preserved naming domain or an eventual reversal effect. It may be that the RAD group is itself heterogeneous, comprised of different variants (e.g., rapid versus slowly progressive). Indeed, the within-subgroup variability of discrepancy scores was much higher for this subgroup compared to the others. Clearly, further research on the RAD subgroup alone is necessary to sort out the
apparent variability within this group of patients.

The cross-sectional CDR findings also generally supported hypothesis four for LAD and GAD but not RAD. Both the LAD and GAD groups showed a mild decrease in discrepancy scores with advancing CDR stage. These findings are in accord with prediction for LAD, which was expected to show less of a discrepancy with advancement of cognitive deterioration, as both domains become affected and global impairment eventuates. It was not expected that GAD discrepancy scores would decrease with advancing CDR stage, but post hoc inspection of the data revealed that scores in both domains reached floors at higher CDR stages, decreasing variability and thus leading to lower mean discrepancies. The finding for the RAD group that discrepancy scores increase with successive CDR stage was clearly unexpected and inconsistent with prediction. Research focusing on the RAD group alone, to sort out the variability within this group, should yield light on this seemingly peculiar cross sectional finding which would appear in discord with the longitudinal findings for this group.

Integration with the Literature

The three patient subgroups identified in this study are consistent with those reported in earlier studies employing similar measures and at times differing methodologies (Becker et al., 1988; Fisher et al., 1996; Martin et al., 1986; Strite et al., 1997), suggesting robustness of these findings. The current finding that the subgroups did not differ in terms of mean duration of illness is also consistent with the results of previous studies (Becker et al., 1988; Fisher et al., 1996; Martin et al., 1986). Similarly, the finding that the GAD subgroup is more impaired in terms of overall dementia severity compared to the other two subgroups is also consistent with earlier reports (Becker et al., 1988; Fisher et al., 1996).
However, there were some findings of this study that are inconsistent with those of earlier similar research. Namely, past studies have reported no differences between the 3 subgroups with respect to age (Fisher et al., 1996; Martin et al., 1986; Strite et al., 1997) and educational attainment (Fisher et al., 1996; Martin, 1990; Strite et al., 1997). However, the current study found the LAD group to be older and less educated than the other two subgroups. Previous studies have also reported that the RAD and LAD groups do not differ from each other in terms of MMSE scores (Fisher et al., 1996; Strite et al., 1997) or functional ADL measures (i.e., Blessed Test; Fisher et al., 1996). In contrast, although the current study found that GAD was indeed more impaired than both RAD and LAD on the MMSE and Blessed tests, LAD was noted to be more impaired on these measures relative to RAD.

These discrepancies from past research may be related to the relatively small sample sizes employed in previous studies, which may have failed to detect the above mentioned intersubgroup differences. The CERAD sample is much larger (i.e., at least 6 times larger than any previously employed sample) and more representative (e.g., in terms of geography, age, and educational levels) than past samples, making findings from the current investigation much more powerful. This may perhaps explain the earlier reports of nonsignificant findings regarding age, educational, and overall impairment differences between the subgroups, yet occasional mention of trends in the direction consistent with the findings of the current study (e.g., Becker et al., 1988; Fisher et al., 1996). The current finding that the LAD group scored significantly lower compared to the RAD group on the MMSE may be related to the LAD group’s lower educational level and older mean age; age and educational attainment have
been shown to be related to MMSE scores (Crum, Anthony, Bassett, & Folstein, 1993; Marcopulos, McLain, & Giuliano, 1997; Tombaugh, McDowell, Krisjansson, & Hubley, 1996). This difference may not have been detected in earlier studies due to the higher mean educational levels and more restrictive age ranges of past samples in comparison to the current sample. It should also be mentioned that, in their study of normal elderly in which LAD- and RAD-like preclinical AD groups were identified, Mitrushina et al. (1995) found their LAD-like group to have lower MMSE scores and to be less educated and older than any other subgroup.

The findings of this study also vary with regard to those of earlier studies in terms of the size of each patient subgroup. Strite et al. (1997) made a point of emphasizing a need for incidence data regarding the three AD subgroups. For comparative purposes, percentages of the samples classified into each subgroup in the current and past subgrouping studies are presented in Table 11. It can be seen from this table that typically, in past research, the GAD subgroup was the largest, followed in size by LAD and then RAD. However, in the current study, LAD was the largest subgroup, followed by RAD and then GAD. These inconsistent incidence rates are difficult to explain.

Whereas all of the studies included in Table 11 employed samples of only those individuals diagnosed with probable AD, there were differences across studies in the measures and methodologies employed. Namely, none of the studies aside from the current one and the Fisher et al. (1996) investigation removed outliers from their sample, forcing all subjects into one of the three subgroups. Furthermore, the Becker et al. (1988) and Strite et al. (1997) subgrouping methodologies involved discrepancy indexes, which are not truly
Table 11

Comparison of Incidence Rates of the Three Subgroups Across Studies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n/N$</td>
<td>%</td>
<td>$n/N$</td>
<td>%</td>
</tr>
<tr>
<td>GAD</td>
<td>71/86</td>
<td>82.60</td>
<td>45/134</td>
<td>33.58</td>
</tr>
<tr>
<td>LAD</td>
<td>11/86</td>
<td>12.79</td>
<td>29/134</td>
<td>21.64</td>
</tr>
</tbody>
</table>

Note. $n = $ number of subjects assigned to the subgroup; $N = $ total number of subjects in the sample; $\% = $ percent of sample classified into each subgroup; GAD = Global Alzheimer's Disease; LAD = Left Alzheimer's Disease; RAD = Right Alzheimer's Disease. [Data from the Martin et al. (1986) study are not provided here as the sample employed in that research was preselected to include those with focal patterns.]
objective classification techniques. This may account for some of the variability across studies. However, the methodology employed in the current investigation was identical to that utilized in the Fisher et al. (1996) study, yet the incidence rates across these two studies still vary, with GAD reported as the largest subgroup in the 1996 study, and the smallest subgroup in the current investigation. Thus, these differences must be due to factors other than classificatory methodology.

The current study used measures which were shorter (i.e., briefer) and hence may be less reliable than those of past studies (Larrain & Cimino, 1998b). On the other hand, the current sample was much larger, spanning many geographical regions, perhaps leading to increased accuracy of incidence rates. In support of this possibility, it should be noted that Joanette et al. (1992), in reviewing their comprehensive multiple single-subject research on AD patients, noted homogeneous impairment across cognitive domains (i.e., the GAD pattern) as apparent in only a minority of patients. As an alternate explanation, it may be that the GAD group is underrepresented in the current (i.e., CERAD) sample due to terms of involvement in the CERAD project. These terms require agreement to return for follow-up evaluations. Those with GAD, being the most impaired, may have been less willing to agree to these terms. Indeed, with respect to likelihood of return for follow-up evaluations, there did seem to be a relationship between overall rate of dementia severity and probability of returning for follow-up evaluations. In the current study, RAD was least impaired in terms of MMSE scores, and members of this subgroup were most likely to return for follow up visits.

With respect to the relative sizes of the LAD and RAD subgroups, the current findings
accord with those of 2/3 of past studies reporting that LAD is more prevalent than RAD (see Table 11). This finding is also consistent with neuroimaging reports which suggest that left hemisphere dysfunction is more common than right hemisphere dysfunction among those with AD (Jamieson et al., 1987; Lowenstein et al., 1989). It must be remembered however, as pointed out by Strite et al. (1997), that incidence rates reported in subgrouping studies do not reflect the general AD population at large. Given the nature of the referral patterns to university clinics, it may be that certain presentations of AD are more likely to be referred for evaluation (i.e., those with predominant, salient language impairments). As well, individuals who agree to participate in research studies are not representative of the population at large.

There is little past research with which to compare the longitudinal and stage findings of the current study. Martin et al. (1990) and Fisher et al. (1997) reported stable subgroup patterns in small numbers of patients studied individually, and similar findings were reported herein for the randomly selected cases from each subgroup. No previous attempt has been made to study larger numbers of members from each subgroup over time. In general, the current findings of stability of the mean GAD and LAD profiles over time are consistent with the earlier case reports.

The longitudinal findings for the RAD patients are less consistent with these earlier case reports, and the reversal effect noted after the 4th testing has not been previously documented. It may be that in the current study, the small numbers of subjects returning after the 4th assessment, and the selective drop out factors influencing those who did not return, have served to distort the RAD group's mean profile. It may also be that the PRAX measure has a higher floor than the BNT measure or that the task becomes overlearned. Moreover,
it is probable that, as mentioned previously, the RAD group is a heterogeneous group in itself. It is clear from case examples in the current and past investigations that some members of the RAD group retain their initial pattern over the first few years. It appears though from the current findings that many do not, and a GAD pattern may quickly emerge. Thus, there may be a more rapid variant of RAD. This is a reasonable speculation in that there have been several reports in the literature suggesting rapidly and slowly progressive forms or variants of AD (e.g., Jacobs et al., 1994; Salmon, Lineweaver, Galasko, & Hansen, 1998), including a study reporting that patients with temporal lobe presentations (i.e., impaired language functions in the context of spared visual-constructional skills and executive functions; akin to LAD) have a slower rate of dementia progression (Butters, Lopez, & Becker, 1996). In addition, there has been a preliminary report in the literature that as a group, patients classified as "High Verbal" (i.e., akin to RAD), show a greater rate of decline on the BNT compared to "High Spatial" patients (i.e., akin to LAD) (Norman, Delis, Salmon, & Bigler, 1998). It may be that the High Verbal group's rate of decline in this latter study was influenced by a predominant subset of members having a more rapid variant of RAD.

With respect to the model proposed by Fisher, Rourke, and Bieliauskas (1997), the current longitudinal findings are thus consistent with the predictions for the GAD and LAD subgroups, yet less consistent with those for the RAD group. Thus, the model requires adjustment/refinement in the case of RAD. Future research focusing on heterogeneity within the RAD group is needed in this regard.

Whereas no other studies have cross-sectionally investigated stability of subgroup profiles across stages of AD, there is one cross-sectional report in the literature which
investigated presence of AD subgroups across stages of dementia. After identifying groups of AD patients similar to the three subgroups reported herein, Strite et al. (1997) sought to determine cross-sectionally whether members of the three subgroups were apparent across stages of the disease. As such, they examined the distribution of subgroup members stratified by MMSE scores into mildly, moderately, and severely demented groups. Inspection of the resultant frequency distribution revealed that members of subgroups comparable to GAD and RAD were identifiable across the different AD stage stratified groups. However, all of their LAD-like patients \( n = 16 \) were classified as moderately demented. Although replication of this result was not an original goal of the current study, post-hoc analyses using CDR ratings and MMSE scores were undertaken in this regard, as it was felt that the absence of LAD members in the mild and severe stratified groups reported by Strite et al. may have been due to the small sample size employed in that study. Such analyses revealed that members of the three subgroups were apparent across mild, moderate and severe CDR and MMSE stratifications.

With respect to CDR ratings (.5 = questionably demented; 1 = mild; 2 = moderate; 3 = severe; 4 = profound; 5 = terminal), members of the GAD and LAD subgroups were apparent across stages .5 to 5, although none were classified at stage 4. Members of RAD were identified in stages .5 through 3 only, with none classified at stages 4 or 5, reflecting the fact that this subgroup was less demented in terms of mean MMSE and Blessed scores. Thus, members of the three subgroups were apparent across mild, moderate, and severe, (i.e., stages 1-3) CDR stages. Because of the potential difference in CDR and MMSE scores—the CDR is clinician rated and the MMSE is psychometric—and the fact that this alone could
account for differences between the current findings and those of Strite et al. (1997), it was
next decided to repeat the above analysis using MMSE stratification, employing the same
criteria for MMSE stratification utilized in the Strite et al. study (i.e., Haxby et al., 1988), in
classifying the patients into mild, moderate, and severe groups. Again, patients from each of
the three subgroups were classified into all 3 severity groups. Thus, it would appear that the
subgroups are indeed apparent across stages of dementia. The fact that Strite et al. found no
LAD patients in their mild and severe MMSE classified groups was likely related to their
small sample size; they had only 16 LAD patients, compared to the 312 LAD patients in the
current study.

Strite et al. (1997) reported asymmetric (i.e., LAD or RAD) profiles in 17.6 %, 35.6
%, and 13.3% of their MMSE-classified mildly, moderately, and severely demented patients,
respectively. The figures for the same analysis with the current sample were: 93%, 78%, and
51%. Differences in these figures across studies are no doubt related to methodological
measurement and subtyping distinctions, differential control of outliers, and sample size
differences (see above). However, it is more interesting to note that the above decreases in
percentages of asymmetric profiles with successive severity of the dementia (i.e., for the
current sample) supports hypothesis four of this study.

The current subgrouping findings for the control group are not inconsistent with those
of Mitrushina et al. (1995) although the samples, methodologies, and measures varied, making
comparisons difficult. These researchers identified a subgroup of “normals” who performed
within the average range on naming and copying tasks yet demonstrated low average to
mildly impaired immediate word list recall performance, and mildly to moderately impaired
performance on delayed and recognition word-list memory trials. Mitrushina et al. labelled this group a memory impaired preclinical form of AD. This group resembles the first subgroup identified in the current study in some respects. Like the Mitrushina et al. group, the CERAD control Subgroup 1 displayed average performance on the semantic and visual constructional measures but some memory impairment. However, the pattern of memory impairment differed, as members of the CERAD control Subgroup 1 demonstrated moderately impaired immediate recall, low average delayed recall and average recognition performances. Clearly, the neuropsychological measures employed across the 2 studies were not the same and the sample characteristics varied. It is possible that these two groups are manifestations of the same process. However, given the different apparent memory pattern (measurement and sampling differences aside), other explanations are in order. This memory impaired group in the current study may not be a preclinical form of AD but may represent a group of normals with attentional problems, as their memory pattern of impaired immediate recall performance in the face of normal delayed and recognition performance is commonly interpreted as due to attentional difficulties (Delis, Kramer, Kaplan, & Ober, 1994; Delis, Kramer, Kaplan, & Ober, 1987). This pattern is often noted among depressed elderly as well (Moss & Albert, 1992). While the normals in the current study were screened for Major Depressive Disorder, it is possible that those with milder depressive conditions were not screened out.

The second normal subgroup identified in the current study also resembles one found by Mitrushina et al. (1995), in which all performances were within normal limits but a memory pattern on the list learning task of recognition > delayed recall > immediate recall was
apparent. As with Mitrushina et al.'s group, control Subgroup 2 in the current study performed within normal limits on the measures and showed the same general memory pattern. Again, however, due to differences between the measures, any seeming similarities across the two studies may be artifactual.

Limitations of Present Study

The referral patterns of the various CERAD sites have been demonstrated to provide largely white, upper-class, well-educated patients and controls (Morris et al., 1989). Hence, the results of this investigation are limited to this specific population; the utility of the CERAD battery with minorities and those of lower socioeconomic status is yet to be determined (Morris et al., 1989). On the other hand, as others have pointed out, the large number of subjects available in the CERAD study and the multi-center nature (i.e., geographical representativeness) of the database provides diversity not possible from single center research (Galasko et al., 1995). Thus, although the generalizability of the current findings is limited by the CERAD sample characteristics, the results of studies done on this sample are the most representative to date.

The longitudinal portions of this study were subject to the inherent limitations of longitudinal designs (Pyke & Agnew, 1991; Rybash, Roodin, & Hoyer, 1995). Namely, attrition due to selective drop out likely affected sample composition at the follow up assessment intervals. This resulted in successively smaller n’s, inconsistency in subjects returning at each visit, and unequal cell sizes preventing repeated measures analyses. Practice effects may have also had an impact on the longitudinal results, as alternate test forms were not used at follow up testing sessions.
The control sample was significantly younger and more highly educated than the patient sample, despite the fact that the ranges and distributions of each group were similar on these variables. Although age and education norms were utilized in calculating T scores for the neuropsychological measures, the current method could be improved upon by reducing the samples to equate the groups in terms of age and education; it is never known for certain whether normative corrections fully control for age and education effects.

Outliers were not removed from the data prior to the cluster analytic replications of the Q-findings. As a result, it is probable that such deviant individuals may have distorted the patterns obtained in the replications. This was apparent among the control group cluster analytic replications of the Q-findings, in which the unclassified Q subjects seemed to inflate the replicated Q-2 profile in terms of performance on the IVBM variable. However, given that the same three patient Q-groups were identifiable across clustering approaches, algorithms, and similarity measures, it is unlikely that the outliers interfered in any significant way. Nevertheless, detection and removal of outliers would serve to increase the reliability of the results reported herein. Utilization of only those subjects classified by the Q-analyses in the cluster analytic replication attempts would likely have increased the accuracy of the replication attempts.

**Directions for Future Research**

Future research should be directed toward improving upon the above noted limitations of the current investigation. In such a manner, this research requires replication on a more representative (i.e., of minorities and low SES groups) sample. In addition, it would be beneficial to repeat the current analyses utilizing data collected from cognitively intact normal
elderly controls who more closely resemble the patients in terms of age and educational attainment, and who are more closely screened for depressive disorders and cognitive impairments. Future subgrouping studies should include patients with possible AD, in order to prevent exclusion of atypical early presentations of AD. Neuroimaging studies on members of the individual subgroups would allow for external validation of the classification scheme suggested herein.

Another beneficial avenue of research would be longitudinal examination of the stability of AD subgroup profiles quantitatively. This could be accomplished with repeated Q-factor analyses conducted at each follow-up interval. In this manner, it could be determined whether patients classified into one subgroup after the initial evaluation change subgroup membership. While members of the LAD subgroup would be expected to change to a GAD pattern eventually, the results of this study suggest continued LAD membership for some time. A subset of RAD patients would be expected to show the same pattern over time, whereas the remaining RAD patients would be predicted to quickly demonstrate a GAD pattern. One would predict that GAD patients would remain within the same subgroup throughout the course of the disease.

Given the apparent additional heterogeneity within the RAD subgroup, research efforts should be devoted to further delineating the variants of this group. It would appear that a subset of RAD patients may have a more rapidly advancing form of the disorder. This hypothesis could be evaluated by subclassifying those in the RAD group into those who have stable patterns over time and those who do not, and studying these two RAD groups separately. These groups could be compared on demographic variables, progression indices,
staging measures, and so on. It would also be of value to examine the performances of the variants of RAD, in addition to the performances of the GAD and LAD subgroups on neuropsychological measures that are not semantic or visual-constructional in nature. This would assist in further elaborating the various subgroups.

Yet another area requiring investigation is longitudinal following of normal elderly individuals via yearly neuropsychological evaluations. This would allow monitoring of premorbid strengths and weaknesses, and determination of the extent to which such differences predict subgroup assignment with the development of AD.

Concluding Comments: Significance and Implications of the Current Findings

The fact that the three neuropsychological subgroups of AD patients reported in the past were again identified in this project, utilizing the largest and most geographically/demographically representative AD sample studied to date, provides strong confirmatory evidence for the existence of qualitatively distinct patterns among this heterogeneous group of patients. The subgroups do not differ in terms of duration of illness. Further, they are identifiable across stages of dementia, whether staging criteria are psychometric or based on clinician’s ratings. Moreover, the initially presenting patterns remain over time for the LAD and GAD groups, in addition to a subset of RAD patients. Taken together, such results are in sharp contradiction to a strictly exclusive “stage model” approach to the conceptualization of AD, in which all patients are characterized as exhibiting global decline across domains of functioning. Whereas it is clear that the GAD subgroup does show homogeneous impairment across semantic and visual-constructional domains, members of the LAD and RAD subgroups are identifiable who demonstrate preservation in one of these areas.
of functioning for a significant period of time. Thus, there is evidence for different patterns of symptom progression across subgroups of patients. As such, it seems appropriate that in conceptualizing the nature of AD, one engage in a theoretical meshing of the stage and subgroup models of the disease.

It appears that the LAD subgroup is older and less educated than the other two groups, suggesting that age at disease onset and educational attainment may have an impact on neuropsychological presentation. In regard to the former, past CERAD research has suggested that those with a younger age of AD onset perform worse on constructional praxis and better on confrontational naming tasks compared to those with a later age of onset, and that younger AD patients appear to progress more rapidly on all CERAD neuropsychological measures (Koss et al., 1996). Non-CERAD studies have also suggested that those with younger age of AD onset evidence more rapid cognitive decline, and better baseline naming performance compared to those with older ages of onset (Jacobs et al., 1994). Neuroimaging and neurochemical investigations have also suggested differences between early- and late-onset AD patients (Koss, Friedland, Ober, & Jagust, 1985; Rossor, Iversen, Reynolds, Mountjoy & Roth, 1984). For example, Koss et al. (1985) reported a greater likelihood of right than left hemisphere metabolic reductions in early onset cases, and Rossor and colleagues (1984) noted that neurochemical changes in late onset cases were most predominant in the temporal lobe, whereas the neurotransmitter deficits in early onset cases were more diffuse and severe. All of these findings accord with the results of the current investigation, and suggest age of onset as a potentially important subgrouping variable, which may overlap with naming/constructional discrepancies.
The possible protective effects of education will be difficult to ascertain, given that age and educational attainment are highly correlated variables, with older individuals less likely to have had access to higher education. Thus, members of the LAD subgroup may be less educated because they are older; educational attainment may have no impact on subgroup assignment whatsoever. It should be emphasized that in the current investigation, the RAD and GAD groups did not differ significantly with regard to educational attainment, yet in terms of level of performance the GAD subgroup performed a mean of 43 T scores lower than the RAD subgroup on the BNT, a level of performance falling below that of the LAD group. Thus, any protective effects of education on naming performance would appear negligible.

With respect to the comparison group findings of the current investigation, there is a strong suggestion that the three patient subgroups which have been repeatedly identified are indeed specific to AD. While premorbid strengths and weaknesses surely have some impact on presenting symptomatology, there is a strong suggestion that these factors alone do not account for the differing AD patterns.

Given these now established subgroups of AD patients, research endeavours involving AD patients should refrain from studying this group as a whole (Martin, 1990). Rather, in order to allow further differences between these subgroups to be realized, researchers should first classify their samples into the three qualitatively distinct groupings, before proceeding to carry out their investigations. This would allow individual study of each subgroup in its own right, with respect to the research question under investigation. Retained areas of functioning apparent among the LAD group and some RAD patients should be targeted in developing compensatory strategies for these individuals in the early course of the disease.
References


is apparent in motor functioning. *Journal of Clinical and Experimental Neuropsychology, 18*, 110-121.


APPENDIX A

Participating CERAD Sites and Principal Investigators

Baylor College of Medicine, Houston, TX
Burke Rehabilitation Center, White Plains, NY
Columbia College of Physicians and Surgeons, NY
Duke University Medical Center, Durham, NC
Emory University/Dekalb Clinic, Atlanta, GA
Group Health, Seattle, WA
Jewish Hospital/Washington University, St. Louis, MO
Johns Hopkins University, Baltimore, MD
Massachusetts General Hospital, Boston, MA
Mount Sinai Medical Center, New York, NY
New York University, New York, NY
Rush-Presbyterian-St. Luke’s Hospital, Chicago, IL
University Hospitals, Cleveland, OH
University of Alabama, Birmingham, AL
University of California, San Diego, CA
University of Kentucky, Lexington, KY
University of Michigan, Ann Arbor, MI
University of Pennsylvania, Philadelphia, PA
University of Pittsburgh Hospitals, Pittsburgh, PA

R. Doody
J. Blass
R. Mayeux
K. Welsh
A. Auchus/H. Karp
E. Larson/W. Kukull
J. Morris
D. Mellits
J. Growdon
R. Mohs
S. Ferris
P. Gorelick/D. Bennett
P. Whitehouse
L. Harrell
W.C. Wiederholt
W. Markesbery
N. Foster
C. Clark
S. DeKosky
University of Rochester/Monroe, Rochester, NY

University of Southern California, Los Angeles, CA

University of Texas/Southwestern, Dallas, TX

University of Washington, Seattle, WA

VA Medical Center, Minneapolis, MN

P. Tariot

V. Henderson

M. Weiner

M. Raskind

J. Mortimer
APPENDIX B

Descriptions of the Tests Comprising the CERAD
Clinical and Neuropsychological Battery

Clinical Assessment

Blessed Dementia Rating Scale. (Blessed et al., 1968) This 22 item behavioural rating scale comprises questions regarding changes in the patient’s habits (e.g., sphincter control), memory (e.g., ability to recall recent events, shopping lists etc.), ability to care for him/herself (e.g., eating, dressing), activities of daily living (e.g., ability to perform household tasks, deal with small sums of money), and personality/emotional functioning (e.g., degree of emotional control, egocentricity), over the previous 6 months. It generates a "dementia" score, which may be viewed as a measure of AD severity. Total inability to complete an activity is given a score of 1, intermittent incapacity is assigned a score of ½, and complete retention of an ability generates a score of zero. Questions regarding eating, dressing, and continence are rated on a 0-3 point scale, with 0 representing no dysfunction, and 3, major impairment. Occurrences of the listed changes in personality, interests, or drives, are scored as 1 for each item present; absences of such alterations are scored as zero. The dementia score is calculated by simply summing the scores allocated for each item. Special scoring allowances are made when the presence of a recently acquired physical disability served to restrict activities. The total score ranges from 0 (fully preserved functional capacity/socioemotional functioning) to 28 (extreme impairment of self-care and personality functioning). This questionnaire was completed in an interview format with the aid of a close family member and/or friend (ideally, the primary caregiver). It is utilized in this research as an overall
estimate of dementia severity in terms of functional competence.

Clinical Dementia Rating Scale. (CDR; Berg, 1984; Hughes et al., 1982; Morris, 1993)

This measure allows clinicians to classify AD patients in terms of overall disease severity, or stage of dementia. It is based on information obtained from structured interview with the patient and informant and encompasses the following categories of functioning: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. On the basis of ratings in terms of the above six domains of functioning, the CDR stages each patient at one of the following levels: (0) non-demented; (0.5) questionably impaired; (1) mildly demented; (2) moderately demented; (3) severely demented; (4) profoundly demented; (5) terminal. Clinician guidelines for rating patients at each stage are provided below.

**Clinical Dementia Rating Scale (CDR) Criteria**
(Berg, 1984; Hughes et al., 1982; Morris, 1993)

<table>
<thead>
<tr>
<th>CDR Stage</th>
<th>Memory</th>
<th>Orientation</th>
<th>Judgment - problem solving</th>
<th>Community affairs</th>
<th>Home - activities</th>
<th>Personal care</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No memory or slight impairment</td>
<td>No cognitive impairment</td>
<td>Slight impairment in problem solving</td>
<td>Independent of cognitive impairment</td>
<td>Self-care activities maintained</td>
<td>Fully capable of self care</td>
</tr>
<tr>
<td>0.5</td>
<td>Mild cognitive impairment</td>
<td>Partial cognitive impairment</td>
<td>Moderate impairment in problem solving</td>
<td>Some impairment in community activities</td>
<td>Home activities maintained</td>
<td>Needs prompting</td>
</tr>
<tr>
<td>1</td>
<td>Moderate cognitive impairment</td>
<td>Severe cognitive impairment</td>
<td>Severe impairment in problem solving</td>
<td>Impaired in community activities</td>
<td>Home activities severely impaired</td>
<td>Requires assistance in daily activities</td>
</tr>
<tr>
<td>2</td>
<td>Severe cognitive impairment</td>
<td>Extremely severe cognitive impairment</td>
<td>Extreme impairment in problem solving</td>
<td>Completely dependent in community activities</td>
<td>Home activities completely impaired</td>
<td>Requires much help with personal care, often assistance</td>
</tr>
<tr>
<td>3</td>
<td>Severe cognitive impairment</td>
<td>Severe cognitive impairment</td>
<td>Extreme impairment in problem solving</td>
<td>Completely dependent in community activities</td>
<td>Home activities completely impaired</td>
<td>Requires much help with personal care, often assistance</td>
</tr>
<tr>
<td>4</td>
<td>Severe cognitive impairment</td>
<td>Severe cognitive impairment</td>
<td>Extreme impairment in problem solving</td>
<td>Completely dependent in community activities</td>
<td>Home activities completely impaired</td>
<td>Requires much help with personal care, often assistance</td>
</tr>
<tr>
<td>5</td>
<td>Severe cognitive impairment</td>
<td>Severe cognitive impairment</td>
<td>Extreme impairment in problem solving</td>
<td>Completely dependent in community activities</td>
<td>Home activities completely impaired</td>
<td>Requires much help with personal care, often assistance</td>
</tr>
</tbody>
</table>
### More Severe Stages of Dementia (usually institutionalized)

<table>
<thead>
<tr>
<th>Process</th>
<th>Profound (4.0)</th>
<th>Terminal (5.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech and Language</td>
<td>Speech usually unintelligible or irrelevant: unable to follow simple instructions or comprehend commands</td>
<td>No response: no comprehension</td>
</tr>
<tr>
<td>Recognition</td>
<td>Occasionally recognizes spouse or caregiver</td>
<td>No recognition</td>
</tr>
<tr>
<td>Feeding</td>
<td>Uses fingers more than utensils, requires such assistance</td>
<td>Feeds to be fed: may have NG tube, may have swallowing difficulties</td>
</tr>
<tr>
<td>Contienniece</td>
<td>Frequently incontinent despite assistance or training</td>
<td>Total incontinence</td>
</tr>
<tr>
<td>Mobility</td>
<td>Able to walk a few steps with help; usually chair-bound; rarely out of home or residence; purposeless movements often present</td>
<td>Bedridden: unable to sit or stand; contractures</td>
</tr>
</tbody>
</table>
Clock Drawing. (Heyman, 1996) This rating measure involves asking the examinee to draw a clock, with the hands of the clock showing the time “10 before 11” or “2:45.” The clinician grades the examinee’s performance on a 3-point scale, to indicate normal, mildly to moderately impaired, or severely impaired performance. The rating is based on the examinee’s ability to place the numbers correctly and symmetrically within the clock face, in addition to placing the hands in a generally accurate position (Heyman, 1996). This measure is utilized in the current study as a measure of visual-constructional functioning.

Neuropsychological Measures

Animal Category Verbal Fluency. (Isaacs & Kennie, 1973) This test requires the examinee to name aloud as rapidly as possible, as many types of animals (e.g., lion, elephant, snake, etc.) he/she can generate within a 60 second time period. Each member of the animal kingdom (real or imaginary) produced is given a score of one, except repetitions and proper nouns. This measure is utilized in this study as a reflection of accessibility of semantic knowledge.

Constructional Praxis. (Rosen et al., 1984) This test involves four line drawn figures (i.e., circle, diamond, intersecting rectangles, cube) of increasing spatial complexity. The examinee is asked to copy each figure; erasing is allowed. There is a two minute time limit per figure. Specific scoring criteria are provided (Heyman, 1996); the maximum score for perfect drawings is 11. This measure is utilized in the current study as a measure of visual-constructional functioning.
Mini-Mental State Exam. (Folstein et al., 1975) The Mini-Mental State Exam is a gross screening instrument containing 30 items designed to assess areas such as orientation, recall, language, perceptual-motor functioning (e.g., copy design), and attention/concentration. It yields a single score as an estimate of dementia severity. A cut-off score of 23 is recommended in determining the presence of cognitive impairment (normal range = 24-30; maximum score = 30) (Folstein et al., 1975). This instrument was utilized as a measure of dementia severity in the current investigation, given its widespread use in research and its established psychometric properties (Galasko et al., 1990; Salmon, Thal, Butters, & Heindel, 1990).

Modified Boston Naming Test. (Kaplan et al., 1978) This test comprises 15 line drawings ranging from familiar items (e.g., house) to more difficult (i.e., less common) objects (e.g., funnel). One third (i.e., five) of the items are of high, medium, and low frequency of occurrence in the English language, respectively (Morris et al., 1989). The examinee is asked to name each object depicted in the drawings. A maximum exposure time of ten seconds is allowed per item (Morris et al., 1989). No semantic or phonemic cuing is provided (Heyman, 1996). Each correctly named item is given a score of 1; perfect performance results in a total score of 15. This test is used in the current study as a measure of semantic knowledge accessibility.

Word List Memory. (Atkinson & Shiffrin, 1971) This is a ten word list learning measure, involving free-recall, delayed free-recall, and recognition components (i.e., each subtask is
qualitatively distinct and receives a separate score). On the first trial, each of the ten nouns are presented visually for 2 seconds, and the examinee is instructed to read each word aloud as they are shown. Following this, the subject is asked to recall as many of the words as possible (90 second time limit). On the next two learning trials the 10 words are again presented visually, but are done so in different orders. Again, after each trial, the subject is asked to recall as many of the words as possible. After a delay period (5 minutes), a delayed free-recall trial is given (90 second time limit). On the recognition trial, the target words are presented interspersed with 10 distractor word cards, and the subject is instructed to state whether or not each word had appeared on the original list. The examinee's recognition score is calculated as the total number of correct words recognized minus 10, to account for chance effects (i.e., random guessing); resulting calculations of less than zero are scored/recorded as zero (Morris et al., 1989). For the immediate free-recall trials (i.e., first 3 trials) the maximum score is 30; the maximum correct responses for the delayed recall trial is 10 (Morris et al., 1989). The immediate recall, delayed recall, and recognition scores are utilized in this study as measures of verbal memory.
APPENDIX C

Glossary of Terms

anemia: Difficulty finding words (Ayd, 1995).

constructional dyspraxia: Difficulty with visual-spatial-motor activities such as manipulating objects, or drawing, often associated with right-hemisphere lesions (Goodwin, 1989).

GAD: Global Alzheimer’s disease (Fisher, Rourke, & Bieliauskas, 1997). A subgroup of AD patients with relatively equivalent impairment in naming and constructional functioning.

LAD: Left Hemisphere Alzheimer’s disease (Fisher, Rourke, & Bieliauskas, 1997). A subgroup of AD patients exhibiting impaired word finding skills in the context of relatively preserved constructional functioning.

probable AD: Designation used to refer to AD patients meeting the following criteria established by the NINCDS-ADRDA (McKhann et al., 1984): (a) Alzheimer’s dementia established by clinical examination and documented by the Mini-Mental Test, Blessed Dementia Scale, or some similar examination, and confirmed by neuropsychological tests; (b) deficits in two or more areas of cognition; (c) progressive worsening of memory and other cognitive functions; (d) no disturbance of consciousness; (e) onset between ages 40 and 90, most often after age 65; and (f) absence of systemic disorders or other brain diseases that in and of themselves could account for the progressive deficits in memory and cognition.

NLD: Nonverbal Learning Disability (Rourke, 1982). A learning disability syndrome characterized by highly developed rote verbal skills within the context of poor psychomotor, tactile-perceptual, visual-spatial, organizational, and nonverbal problem-solving abilities (Harnadek & Rourke, 1994).
**RAD:** Right Hemisphere Alzheimer's Disease (Fisher, Rourke, & Bieliauskas, 1997). A subgroup of AD patients exhibiting impaired constructional functioning in the context of relatively preserved word finding skills.

**VIQ:** The Verbal Intelligence Quotient (VIQ) is a score calculated from the verbal subtests of the Wechsler Adult Intelligence Scale.

**PIQ:** The Performance Intelligence Quotient (PIQ) is a score calculated from the performance subtests of the Wechsler Adult Intelligence Scale.
VITA AUCTORIS

Nancy Justina Fisher was born on June 9th, 1968, in Toronto, Ontario. She graduated from high school an Ontario Scholar in January of 1987, and was granted early admission, on academic scholarship, to the undergraduate Liberal Arts programme at the University of Guelph. In June of 1992, she received her Bachelor of Arts (Specialized Honours-Psychology) Degree from York University, granted the distinction *Summa Cum Laude* upon graduation. Nancy received her Master of Arts degree in Clinical Neuropsychology from the University of Windsor in 1995. Her doctoral training/research at the University of Windsor was supported by awards from the Alzheimer Society of Canada. She is scheduled to receive her Ph.D. in Clinical Neuropsychology in October of 1998. Her post-doctoral fellowship in Geriatric Neuropsychology will be completed at Sunnybrook Health Science Centre in Toronto, Ontario, under the supervision of Dr. Mary Tierney.