NEUROPSYCHOLOGICAL CHARACTERIZATION OF
COGNITIVELY-IMPAIRED-NOT-DEMENTED INDIVIDUALS

by

KEVIN ROSS PETERS

B.A., Brock University, 1997

M.A., The University of British Columbia, 1999

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Department of Psychology

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

September 2003

© Kevin Ross Peters, 2003
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology

The University of British Columbia
Vancouver, Canada

Date Sept. 3, 2003
The focus of this dissertation was on Cognitively-Impaired-Not-Demented (CIND) individuals. CIND is a diagnostic label applied to individuals who present with cognitive impairment but do not meet formal criteria for dementia (Ebly, Hogan, & Parhad, 1995). The purpose of this research was to determine whether the cognitive heterogeneity of CIND could account for the prognostic heterogeneity of this condition. This dissertation is divided into two sections. The primary objective of the first section was to characterize the neuropsychological test performance of a large sample of clinic-referred CIND participants. Participants classified as Not-Cognitively-Impaired (NCI; n = 68) differed from CIND individuals (n = 205) on a number of demographic, clinical, and neuropsychological variables. Measures of learning and memory, visuoconstruction abilities, and cognitive flexibility provided the best discrimination between NCI and CIND participants. Clinical comparison data were generated for various demographically defined groups of CIND participants. The results supported the impression that CIND is a cognitively heterogeneous population.

The primary objectives of the second section of this dissertation were, first, to determine whether subgroups of CIND individuals with distinct neuropsychological profiles exist in two independent samples, and second, to determine whether subgroup membership was related to diagnostic outcome over periods of 2 to 5 years. A series of cluster analyses was performed on ipsative factor z-scores derived from principal component analyses. Five subgroups were identified in the Base Sample (n = 461): Verbal Dysfunction, Verbal/Visuospatial Dysfunction, Memory/Verbal Dysfunction, Memory Dysfunction, and Visuospatial Dysfunction. This 5-cluster solution was replicated in an independent sample of CIND individuals (n = 166). The highest rates of conversion to dementia were observed in the Memory Dysfunction and Memory/Verbal Dysfunction subgroups. The Verbal Dysfunction subgroup was most likely to show improvement in cognitive status. The results suggest that CIND is a syndrome of conditions rather than a
single population at risk for converting to dementia. This observation has implications for the
diagnosis and treatment of CIND individuals. A deeper understanding of the cognitive
heterogeneity of CIND will facilitate research to identify individuals who will and will not
progress to dementia.
TABLE OF CONTENTS

ABSTRACT ......................................................... ii

LIST OF TABLES .................................................. v

LIST OF FIGURES ................................................ vi

ACKNOWLEDGEMENTS ............................................. vii

CHAPTER 1 Introduction and Overview ................................ 1
Overview of Dissertation ........................................ 4

CHAPTER 2 Cognitive Impairment Without Dementia: A Literature Review ........ 6
An Overview of Diagnostic Labels .................................. 6
Rationale for Examining CIND ..................................... 12
Neuropsychological Prediction of Dementia: Contemporary Approaches ..................... 13
Neuropsychological Prediction of Dementia: The Findings .................................. 15
Overview of the Present Investigation ................................ 18
Appendix 2.1: Cluster Analysis: A General Overview ......................... 21

CHAPTER 3 Neuropsychological Characterization of Cognitively-Impaired-Not-Demented (CIND) Individuals: Clinical Comparison Data .................. 25
Method .................................................................. 29
Results and Discussion ............................................. 36
General Discussion .................................................. 52
Appendix 3.1: Example of How to Use Logistic Regression Results to Facilitate Diagnosing CIND .................................................. 58

CHAPTER 4 Subgroups of Cognitively-Impaired-Not-Demented (CIND) Individuals: Delineation, Reliability, and Predictive Validity .................. 61
Method .................................................................. 66
Results .................................................................. 73
Discussion ............................................................. 94

CHAPTER 5 General Discussion ......................................... 100
Summary of Major Findings and Contributions of This Research .................. 100
Theoretical and Practical Implications of This Research ............................ 102
Limitations of This Research ....................................... 106
Avenues for Future Research ........................................ 108

REFERENCES .............................................................. 115
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Description of Selected Labels That Are Based on Formal Sets of Diagnostic Criteria</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>Demographic and Clinical Data for NCI and CIND Participants</td>
<td>38</td>
</tr>
<tr>
<td>3.2</td>
<td>Mean (and Standard Deviation) of Performance on Neuropsychological Tests with Univariate Logistic Regression Analyses Results for NCI ( (n=68) ) and CIND ( (n=205) ) Participants</td>
<td>39</td>
</tr>
<tr>
<td>3.3</td>
<td>Results of a Backward Stepwise Logistic Regression Analysis for the Base Sample of NCI and CIND Participants ( (n=223) )</td>
<td>41</td>
</tr>
<tr>
<td>3.4</td>
<td>Mean (and Standard Deviation) Neuropsychological Test Performance for Various Age Groups with ANOVA Results</td>
<td>44</td>
</tr>
<tr>
<td>3.5</td>
<td>Mean (and Standard Deviation) Neuropsychological Test Performance Arranged by Education Group with Independent t-test Results</td>
<td>45</td>
</tr>
<tr>
<td>3.6</td>
<td>Mean (and Standard Deviation) Neuropsychological Test Performance for Males and Females with Independent t-test Results</td>
<td>47</td>
</tr>
<tr>
<td>4.1</td>
<td>Number (and Percentage) of Participants in Each CIND Subcategory in the Base ( (n=461) ) and Replication ( (n=166) ) Samples</td>
<td>69</td>
</tr>
<tr>
<td>4.2</td>
<td>Neuropsychological Tests Grouped by DSM-III-R Criteria for Dementia in the Base and Replication Samples</td>
<td>70</td>
</tr>
<tr>
<td>4.3</td>
<td>Results of Principal Component Analysis with Varimax Rotation for the Base Sample ( (n=461) )</td>
<td>74</td>
</tr>
<tr>
<td>4.4</td>
<td>Results of Principal Component Analysis with Varimax Rotation for the Replication Sample ( (n=166) )</td>
<td>75</td>
</tr>
<tr>
<td>4.5</td>
<td>Subgroup Profiles on Demographic and Clinical Measures for Base ( (n=461) ) and Replication ( (n=166) ) Samples</td>
<td>84</td>
</tr>
<tr>
<td>4.6</td>
<td>Mean (and Standard Deviation) Neuropsychological Test Scores for Each Subgroup in the Base Sample ( (n=461) ).</td>
<td>87</td>
</tr>
<tr>
<td>4.7</td>
<td>Mean (and Standard Deviation) Neuropsychological Test Scores for Each Subgroup in the Replication Sample ( (n=166) )</td>
<td>88</td>
</tr>
<tr>
<td>4.8</td>
<td>Diagnostic Status at Follow-up in Base ( (n=225) ) and Replication ( (n=94) ) Samples</td>
<td>93</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>3.1</td>
<td>Frequency histogram of intertest scatter scores in CIND individuals</td>
<td>48</td>
</tr>
<tr>
<td>3.2</td>
<td>Neuropsychological profiles for CIND individuals with the smallest (squares with dashed line) and largest (diamonds with solid line) intertest scatter scores</td>
<td>50</td>
</tr>
<tr>
<td>3.3</td>
<td>Plot showing the cumulative percentage of CIND individuals for each level of cognitive impairment ranging from 0 to 14 measures</td>
<td>51</td>
</tr>
<tr>
<td>3.4</td>
<td>Neuropsychological profiles for CIND individuals with similar intertest scatter scores and impairment scores</td>
<td>53</td>
</tr>
<tr>
<td>4.1</td>
<td>Factor score profile for Verbal Dysfunction subgroup</td>
<td>78</td>
</tr>
<tr>
<td>4.2</td>
<td>Factor score profile for Verbal/Visuospatial Dysfunction subgroup</td>
<td>79</td>
</tr>
<tr>
<td>4.3</td>
<td>Factor score profile for Memory/Verbal Dysfunction subgroup</td>
<td>80</td>
</tr>
<tr>
<td>4.4</td>
<td>Factor score profile for Memory Dysfunction subgroup</td>
<td>81</td>
</tr>
<tr>
<td>4.5</td>
<td>Factor score profile for Visuospatial Dysfunction subgroup</td>
<td>82</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

While completing this dissertation, I received considerable support from a number of people who deserve to be acknowledged. First of all, I would like to thank the members of my thesis committee for taking the time to share their knowledge and expertise with me over the years that I have been at U.B.C. As my supervisor, Peter Graf has been the one most responsible for helping me to develop my own way of thinking about research and for how I express my ideas in both oral and written formats. Sherri Hayden, my co-supervisor, has taught me about the clinical side of neuropsychology and has helped me to understand that there is more to research than simply collecting data. John Pinel has always been available to talk to me about the finer points of teaching and giving presentations and for just talking about things that interest us both.

I would also like to acknowledge the hard work and dedication of the participants and study personnel from the Canadian Study of Health and Aging (CSHA) and the Canadian Collaborative Cohort of Related Dementias (ACCORD) study. I would like to thank Howard Feldman (Head, Division of Neurology, Department of Medicine at U.B.C) especially for allowing me to take such an active role in the neuropsychological component of the ACCORD study. Under the training and supervision of Sherri Hayden, I performed neuropsychological testing for many of the ACCORD study participants at the Vancouver Study Centre and I was responsible for creating the neuropsychological database for the ACCORD study.

The members of the Memory and Cognition Lab (Claudia Jacova, Beat Meier, Noam Butterfield, Erin Rennie, Todd Woodward, and Stephanie Cassie) also deserve to be acknowledged. These individuals have probably heard more about my research than anyone else at U.B.C. and I thank them all for their helpful feedback over the years.

Two of the chapters in this dissertation are self-contained manuscripts that will be submitted for publication. I was the primary investigator on the work reported in these manuscripts and will be the senior author on the to-be-submitted manuscripts. The following
individuals will appear as co-authors: Peter Graf, Sherri Hayden and Howard Feldman. These individuals assisted me by contributing valuable feedback on my ideas and in the editorial process of preparing these manuscripts for submission.

Throughout the completion of my doctorate, I have been very fortunate to receive financial support through a Postgraduate Scholarship from the Natural Sciences and Engineering Research Council of Canada and from a Doctoral Training Award jointly funded by the Alzheimer Society of Canada and the Canadian Institutes of Health Research.

Lastly, I would to acknowledge my partner Anne-Marie and my daughter Sarah, who was born during the completion of my dissertation. Anne-Marie has always been very kind and supportive of my obsessive tendencies to keep working on things until they were just right (well, almost just right). She knew when to give me space, but she also knew when to remind me about having perspective. I thank her for all the times that she picked up the slack around the house when I was swamped with deadlines. As for Sarah, she has not been around for very long, but she has been the greatest source of inspiration in my life. She has taught me about the importance of having fun and for just being and for that I thank her.
CHAPTER 1

Introduction and Overview

Is it possible to identify distinct neuropsychological subgroups among individuals diagnosed as Cognitively-Impaired-Not-Demented (CIND)? Do reliable profiles of neuropsychological test data exist for predicting the development of dementia? The purpose of this exploratory investigation was to answer these questions using data from two large cohorts of CIND individuals. The first objective was to characterize the neuropsychological test performance of CIND individuals and provide data that would enhance the accurate identification of these persons. The second objective was to try to identify subgroups with distinct profiles of neuropsychological test performance in this population. The third objective was to determine whether subgroup membership has diagnostic value, that is, whether subgroups are associated with differential outcomes (e.g., dementia versus no dementia).

Dementia is a general diagnostic label used to characterize individuals who demonstrate progressive and global cognitive decline which is often associated with some form of neurodegeneration (Lezak, 1995). It has been estimated that approximately 8% of Canadians over 65 have dementia according to the Canadian Study of Health and Aging (CSHA Working Group, 1994a). Similar prevalence rates have been reported for other countries (Fratiglioni, De Ronchi, & Aguero-Torres, 1999; Jorm, Korten, & Henderson, 1987; Rockwood & Stadnyk, 1994). There is also consistent evidence that the prevalence of dementia increases with age, approximately doubling every 5.1 years after the age of 65 (CSHA Working Group, 1994a; Jorm et al., 1987; Rockwood & Stadnyk, 1994). The costs associated with dementia in Canada are approximately $3.9 billion per year (Ostbye & Crosse, 1994). These will increase steadily as the segment of our elderly population continues to grow. It has been estimated that the prevalence of dementia in Canada will more than double by the year 2021 and triple by the year 2031 (CSHA Working Group, 1994a).
What is CIND and why is it important to study CIND individuals? CIND is a diagnostic label applied to individuals who present with cognitive impairment but do not meet formal criteria for dementia (Ebly et al., 1995). It is a cognitively heterogeneous condition, with impairment present in either one domain or in multiple domains. CIND has been estimated to characterize 16.8% of Canadians over the age of 65 (Graham et al., 1997). In the CSHA (CSHA Working Group 1994a; Hogan & Ebly, 1999), the conversion rate to dementia in CIND individuals over a 5-year period was 45.5% compared to 14.5% in individuals initially classified as Not-Cognitively-Impaired (NCI). However, the CSHA results also revealed that not all CIND individuals necessarily progress to dementia: Of the individuals initially diagnosed as CIND, 42.1% maintained their cognitive status over the same 5-year period and 12.4% even showed improvement. These results indicate that CIND is an at-risk population for dementia and support the assumption that CIND might be a prodromal phase of dementia (Rediess & Caine, 1996; Tuokko & Frerichs, 2000). More importantly these results also underscore the need for a better understanding of the mechanisms that determine diagnostic outcomes in this population.

The overall goal of the present investigation was to illuminate the heterogeneity of neuropsychological profiles and diagnostic outcomes in CIND individuals. More specifically, I wanted to determine whether the cognitive heterogeneity demonstrated by CIND individuals could account for the prognostic heterogeneity associated with this population. The first step toward this goal was to explore the cognitive heterogeneity of CIND individuals and to provide data that would facilitate their accurate identification. Clinicians must be able to identify CIND individuals in a valid and reliable manner before they can determine which of these individuals will and will not progress to dementia.

The second step toward achieving the overall goal of this investigation was to try to identify reliable subgroups of CIND individuals with distinct neuropsychological profiles, and to determine whether these subgroups differ in their risk for progressing to dementia. The
identification of both high- and low-risk subgroups would have several important implications. The clinical treatment and management of CIND individuals could be tailored according to their risk for conversion to dementia. There is currently no cure for neurodegenerative types of dementia such as Alzheimer disease, but there are pharmaceutical treatments available that are effective in stabilizing or even enhancing cognitive functioning in these conditions for periods of 6 to 12 months (for review, see Knopman, 2003). If high-risk subgroups of CIND individuals can be identified, these pharmaceutical interventions could be implemented earlier and possibly help to delay disease progression and maintain quality of life. High-risk individuals and their families could also receive psychological counselling earlier to help them prepare for the many difficult aspects of the disease and to make decisions for long-term care if necessary.

The reliable identification of low-risk subgroups of CIND individuals would also yield diverse benefits. Currently available pharmaceutical treatments for dementia are often costly and are associated with negative side effects in some individuals. For this reason, physicians may choose to forego or delay immediate treatment in individuals from low-risk subgroups. The utility of alternative treatment interventions may also be explored in individuals from low-risk subgroups. It might be the case that some of these individuals are suffering from alcohol abuse or psychiatric conditions such as depression or anxiety that are amenable to other forms of treatment.

The identification of high- and low-risk subgroups of CIND individuals would also allow for more effective channelling of clinic resources. Because there is currently no way to identify which CIND individuals will and will not develop dementia, it is common practice for many of these individuals to receive annual or even semi-annual assessments to monitor their status. However, clinical assessments are costly in terms of both time and money. If there were subgroups of CIND individuals with low probabilities of developing dementia, limited clinic resources could be allocated more efficiently to the assessment of higher-risk individuals.
Knowledge of CIND subgroups that differ with respect to diagnostic outcome would allow for cleaner and more informative research into the etiological factors and biomarkers of dementia. Putative biomarkers might be present, for example, in a high-risk subgroup of CIND individuals but not in other subgroups with a lower risk for dementia, and thus any significant subgroup effects would be washed out if one examines average data collected from a large sample comprising individuals from different subgroups.

Pharmaceutical companies are currently in the process of developing more effective symptomatic treatments and possible cures for various types of dementia such as Alzheimer disease and vascular dementia (Doraiswamy, 2002; Knopman, 2003). Inclusion of CIND individuals from high-risk subgroups only would allow for cleaner and more conclusive clinical trials. Pooling together all CIND individuals and using an overall group mean might obscure any significant effects that may be present for individual subgroups.

Overview of Dissertation

The remainder of this manuscript is divided into four chapters. Chapter 2 is a review of the literature on CIND and on the neuropsychological prediction of dementia. It is broken down into three sections. The first section focuses mostly on terminology. It begins by defining dementia and is followed by an overview of the many diagnostic labels that have been used to describe individuals believed to be at risk for dementia. The second section of Chapter 2 summarizes the two major approaches that have been adopted by contemporary neuropsychological researchers interested in predicting dementia as well as the major findings that have been reported. The third section of Chapter 2 is an overview of the present investigation.

Chapters 3 and 4 are self-contained manuscripts based on results from this dissertation that will be submitted for publication. In the manuscript presented in Chapter 3, the neuropsychological component of the Collaborative Cohort of Related Dementias (ACCORD)
study is described in detail. The ACCORD study is one of the two studies that provided data for the present investigation. Baseline results for the CIND and NCI participants are reported. Clinical comparison data for CIND individuals are also reported for a number of neuropsychological tests. The following manuscript, Chapter 4, is a detailed report of the neuropsychological subgroups that were identified by this investigation. The methodology used to identify and validate these subgroups is presented.

In Chapter 5, the main findings of the present investigation are summarized and their implications are discussed in the context of the issues raised in the literature review (Chapter 2). Chapter 5 also addresses a few of the limitations of the present investigation and lays out potential avenues for future research.
CHAPTER 2

Cognitive Impairment Without Dementia: A Literature Review

This chapter is a review of the literature on cognitively impaired individuals who do not meet diagnostic criteria for dementia. The first section focuses mostly on terminology. It begins with a general description of dementia and is followed by a summary of the various labels that have been used to characterize individuals who are at-risk for dementia. The rationale for selecting one of these labels in particular, CIND, for this investigation is then presented. The second section focuses on two different methods, the best test and subgroup approaches, that have been used to predict conversion to dementia, and it reviews the findings from studies using these two methods. The third section is a general overview of the present investigation. The rationale for choosing the subgroup approach over the best test approach is also discussed in this section.

An Overview of Diagnostic Labels

According to the revised fourth edition of the Diagnostic and Statistical Manual (American Psychiatric Association, 2000), the diagnosis of dementia requires there to be significant memory impairment and at least one of the following: aphasia (language difficulties), agnosia (perception difficulties), apraxia (motor difficulties), or disturbances of executive functioning (planning, organizing, or abstracting). Typically, a diagnosis is made after a comprehensive clinical assessment including a neuropsychological examination of the individual to determine the presence of objective cognitive deficits. The cognitive impairments must be progressive, interfere with activities of daily living, and not be attributable to delirium or any other mental disorder. Making a definitive diagnosis of dementia, as well as efforts to differentiate among the various types of dementia, is facilitated by knowledge -- obtained from a detailed medical history -- about the possible etiology of cognitive impairments, by neuroimaging data, and by the presence of abnormal changes in behaviour and personality.
Alzheimer disease is the most frequent cause of dementia, accounting for up to 75% of cases (Fratiglioni et al., 1999), and the most common types of non-Alzheimer dementias are vascular dementia, frontotemporal dementia, and dementia with Lewy bodies (Kaye, 1998; Knopman, 2001).

Considerable research focused on the early identification of dementia has involved individuals who demonstrate cognitive impairment but do not meet formal criteria for dementia (Collie & Maruff, 2002; Tuokko & Frerichs, 2000). These individuals are characterized by greater than normal rates of conversion to dementia. Previous investigations have demonstrated annualized conversion rates that range between 6.3% and 43.8%, with an average of 12 to 15% (Albert, Moss, Tanzi, & Jones, 2001; Berg et al., 1982; Bowen et al., 1997; Devanand, Folz, Gorlyn, Moeller, & Stern, 1997; Petersen et al., 1999; Tierney, Szalai, & Snow, 1996). These conversion rates are considerably higher than those for individuals without cognitive impairment, which range from 0% to 5.1% per annum (Albert et al., 2001; Elias et al., 2000; Hogan & Ebly, 1999; Jacobs et al., 1995; Masur, Sliwinski, Lipton, Blau, & Crystal, 1994; Small, Fratiglioni, Viitanen, Winblad, & Backman, 2000). The conversion rate data indicate that nondemented individuals with cognitive impairment are an at-risk population; these individuals have been described as being in a subclinical or prodromal phase for certain types of dementia (Petersen, 2000a; Rediess & Caine, 1996; Tuokko, Frerichs, Kristjansson, & McDowell, 2000).

Many diagnostic labels have been used to describe nondemented individuals with cognitive impairment. Tuokko and colleagues (Tuokko & Frerichs, 2000; Tuokko, Frerichs, & Kristjansson, 2001) have grouped existing labels into the following three categories: labels that are descriptive, labels associated with specific assessment instruments, and labels that are based upon formal sets of diagnostic criteria.

Descriptive labels for nondemented individuals. The majority of descriptive labels for nondemented individuals with cognitive impairment are based on a clinician's impression after
performing a neurological examination of an individual or after reviewing an individual’s performance on neuropsychological tests. Such labels include *age-related cognitive decline* (American Psychiatric Association, 1994), *mild dysfunction* (Johansson & Zarit, 1997), *cognitively-impaired-not demented* (CIND; Ebly et al., 1995; Graham et al., 1997), *pre-dementia* (Reifler, 1997), *at-risk for Dementia of the Alzheimer Type* (Cahn et al., 1995), and *mild dementia* (Berg, 1990). Descriptive labels have also been used to characterize individuals without cognitive impairment who eventually convert to dementia (Chen et al., 2001; Elias et al., 2000; Jacobs et al., 1995; Linn et al., 1995; Small et al., 2000) as well as asymptomatic individuals that are at risk for developing dementia given a family history of dementia or the presence of other risk factors (Almkvist et al., 1998; Almkvist, Basun, Wagner, Wahlund, & Lannfelt, 1997; Fox et al., 1996).

*Labels associated with specific assessment instruments.* In an attempt to characterize nondemented individuals with cognitive impairment in a more quantitative manner, several diagnostic labels associated with specific assessment instruments have been used. These assessment instruments are typically rating scales that are used to provide an index of the severity of an individual’s cognitive and/or functional impairment. For example, the Global Deterioration Scale (GDS; Reisberg, Ferris, de Leon, & Crook, 1982) is an instrument used by clinicians to rate the severity of cognitive impairment, ranging from 1 (no cognitive decline) to 7 (very severe cognitive decline). GDS scores of 2 and 3 designate *very mild cognitive decline* and *mild cognitive decline* respectively.

The Clinical Dementia Rating Scale (CDR) covers a wider range of dimensions than the GDS, and includes memory, orientation, judgement and problem-solving, community affairs, home and hobbies, and personal care (Hughes, Berg, & Danzinger, 1982). CDR scores range from 0 to 3, with 0.5 indicating *questionable dementia*. Interpretation of the CDR scale, however, is complicated and subjective and there is a lack of agreement over its use. For
example, a score of 0.5 can be given to a nondemented individual with mild cognitive impairment as well as someone already diagnosed with mild Alzheimer disease. In an attempt to reduce such ambiguity, Morris and colleagues (Morris et al., 2001) have recently subdivided the CDR score of 0.5 into three classifications on the basis of prognosis: 0.5/dementia of the Alzheimer type, 0.5/incipient dementia of the Alzheimer type, and 0.5/uncertain dementia.

Palmer, Wang, Backman, Winblad, and Fratiglioni (2002) have recently used the Mini Mental Status Exam (MMSE; Folstein, Folstein, & McHugh, 1975), a brief screening instrument used to assess general cognitive functioning, to define varying levels of cognitive impairment without dementia. Three levels of impairment were delineated: mild, moderate, and severe. These three levels correspond to performance on the MMSE that is 1-, 1.5-, and 2-standard deviations below normal. Other examples of labels based on specific assessment instruments include limited cognitive disturbance from the Comprehensive Assessment and Referral Evaluation (Gurland, Dean, Copeland, Gurland, & Golden, 1982), minimal dementia from the Cambridge Mental Disorders in the Elderly Examination (Roth et al., 1986), and borderline dementia from the Clifton Assessment Procedures for the Elderly (Clarke et al., 1996). The use of specific assessment instruments has resulted in more quantitative labels for nondemented individuals with cognitive impairment. However, there is currently no consensus over which instrument should be used and new labels continue to appear.

Labels based on formal sets of diagnostic criteria. A number of investigators have also described nondemented individuals with cognitive impairment using formal diagnostic criteria. Some of the more popular labels are reported in Table 2.1. There are several similarities among the labels presented in Table 2.1. Most require the following stipulations: the presence of a subjective complaint, some predetermined level of impairment on cognitive testing, and no functional impairment. The majority of labels also come with several exclusion criteria in
Table 2.1
Description of Selected Labels That Are Based on Formal Sets of Diagnostic Criteria

<table>
<thead>
<tr>
<th>Label</th>
<th>Subjective Complaint</th>
<th>Domain(s) Affected</th>
<th>Objective Cut-off</th>
<th>Functional Impairment</th>
<th>Exclusions (other conditions that may contribute to impairment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAMI⁴</td>
<td>Yes</td>
<td>Memory</td>
<td>-1 SD for young adults</td>
<td>--</td>
<td>Delirium, medical, neurological, or psychiatric conditions, alcohol or substance abuse, use of psychotropic drugs</td>
</tr>
<tr>
<td>MCI²</td>
<td>Yes</td>
<td>Memory</td>
<td>-1.5 SDs for age and education; CDR⁸ = 0.5</td>
<td>No</td>
<td>None specified</td>
</tr>
<tr>
<td>MCI³</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSM-III-R⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type I</td>
<td>--</td>
<td>Memory</td>
<td>--</td>
<td>No</td>
<td>Delirium, organic or mental disorders</td>
</tr>
<tr>
<td>Type II</td>
<td>--</td>
<td>Memory plus</td>
<td>--</td>
<td>No</td>
<td>Delirium, organic or mental disorders</td>
</tr>
<tr>
<td>ICD-10¹⁰</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type I</td>
<td>--</td>
<td>Memory</td>
<td>--</td>
<td>No</td>
<td>Delirium</td>
</tr>
<tr>
<td>Type II</td>
<td>--</td>
<td>Memory plus</td>
<td>--</td>
<td>No</td>
<td>Delirium</td>
</tr>
<tr>
<td>Type III</td>
<td>--</td>
<td>Memory plus</td>
<td>--</td>
<td>No</td>
<td>Delirium</td>
</tr>
<tr>
<td>AACD⁴</td>
<td>≥ 6 months</td>
<td>Any domain</td>
<td>-1 SD for age and education</td>
<td>--</td>
<td>Physical, neurological, or psychiatric conditions, psychoactive substance abuse</td>
</tr>
<tr>
<td>MCD⁵</td>
<td>≥ 2 weeks</td>
<td>Any domain</td>
<td>Yes, but not specified</td>
<td>--</td>
<td>Delirium, post-encephalic syndrome, postconcussional syndrome, psychoactive substance use (NOTE: other physical or neurological conditions are allowed)</td>
</tr>
<tr>
<td>MND⁶</td>
<td>≥ 2 weeks</td>
<td>Any domain</td>
<td>Yes, but not specified</td>
<td>Yes</td>
<td>Delirium, amnestic disorder, mental disorders (NOTE: other medical or neurological conditions are allowed)</td>
</tr>
<tr>
<td>CIWD⁷</td>
<td>Yes</td>
<td>Any domain</td>
<td>≤ 7th percentile for age and education</td>
<td>No</td>
<td>None specified</td>
</tr>
</tbody>
</table>
Note: Entries with double dash (--) indicate that data were not reported
1. Age associated memory impairment (Crook et al., 1986)
2. Mild cognitive impairment (Petersen et al., 1995)
3. Mild cognitive impairment (Zaudig, 1992)
4. Age associated cognitive decline (Levy, 1994)
5. Mild cognitive disorder (World Health Organization, 1993)
7. Cognitive impairment without dementia (Unverzagt et al., 2001)
8. Clinical Dementia Rating Scale (Berg et al., 1982)

✈ Memory plus at least one of the following: abstract thinking, judgement, apraxia, agnosia, aphasia, or personality change
♦ Memory plus intellectual decline
♥ Memory plus intellectual decline plus deterioration in emotional control, social behaviour, or motivation
addition to dementia. A key difference among the labels is that some focus predominately on deficits in memory while others allow impairment in multiple cognitive domains.

Summary of diagnostic labels for nondemented individuals. Many diagnostic labels have been used to describe nondemented individuals with cognitive impairment. Some of these labels are based on objective criteria, whereas others are more subjective. The fact that investigators have used such a wide variety of labels has resulted in confusion and difficulty in drawing conclusions across studies (Petersen, 2000b; Ritchie, Artero, & Touchon, 2001; Tuokko & Frerichs, 2000; Tuokko et al., 2001).

Rationale for Examining CIND

The present investigation focuses on individuals diagnosed as CIND. This label was first used in the CSHA (CSHA Working Group, 1994) to describe individuals who were cognitively impaired but did not meet formal diagnostic criteria for dementia (American Psychiatric Association, 1987; Ebly et al., 1995). In the CSHA, CIND was a consensus-based diagnosis made by a team of individuals including the principal investigator, a physician, a nurse, and a neuropsychologist. Compared to other labels, CIND is a more comprehensive label that captures individuals with cognitive impairment in one domain or in multiple domains and it comes with fewer exclusion criteria.

There were three reasons for selecting CIND over the various other labels that have been used. First, CIND individuals are cognitively heterogeneous and thus similar to patients with dementia, a condition defined by the presence of impairments in memory and at least one other cognitive domain (American Psychiatric Association, 2000). In the present investigation, CIND individuals are assumed to be representative of individuals at risk for dementia. A second reason for focusing on CIND is that individuals with this condition constitute the majority of the older adults who seek help for their cognitive difficulties. Persons who attend memory or dementia clinics tend to exhibit a wide range of cognitive deficits and comorbid physical conditions
(Devanand et al., 1997; Reifler, 1997), and thus, a better understanding of the varieties of clinical presentation and prognosis of CIND promises to yield a range of clinical benefits. A third reason for choosing CIND was that it is one of the most encompassing labels, and therefore, the relationship between patterns of cognitive impairment and the possible causes of cognitive impairment could be explored further to determine whether this relationship contributes to the prediction of dementia. Is it the case, for example, that memory impairment associated with stroke is more predictive of dementia than memory impairment associated with depression?

Neuropsychological Prediction of Dementia: Contemporary Approaches

Although there are numerous ways of predicting dementia, most investigators have attempted to use individual's neuropsychological test performance for this purpose (for reviews, see Collie & Maruff, 2000; Tuokko & Frerichs, 2000). This focus is understandable, first, because the clinical diagnosis of dementia is based on the presence of objective neuropsychological deficits, and second, because of the assumption that mild cognitive impairment is predictive of the global cognitive impairment that characterizes dementia.

Primarily two methods have been used for making predictions based on neuropsychological test performance. One method aims to find a single instrument, the best test, for making predictions. The other method focuses on the identification of subgroups that are characterized by distinct neuropsychological profiles and different diagnostic outcomes (e.g., dementia versus no dementia).

*The best test approach.* This approach involves administering a neuropsychological test battery to a sample of individuals and then monitoring their diagnostic status over time to identify which neuropsychological test is the best predictor of progression to dementia (Kluger, Ferris, Golomb, Mittelman, & Reisberg, 1999). In a world of decreasing healthcare resources, the best test approach seems like a reasonable and practical solution, but it might not be optimal for two reasons. First, dementia includes a number of conditions that are associated with
different causes and present with different symptoms. Even the diagnostic criteria for a specific type of dementia, such as Alzheimer disease, allows for diagnostic heterogeneity by stipulating that in addition to memory dysfunction, individuals must demonstrate impairment in at least one other area of cognitive functioning (American Psychiatric Association, 2000). The fact that individuals with different combinations of cognitive deficits can be diagnosed with the same disease makes it unlikely that a single test will emerge as a consistent predictor of this outcome.

The second reason why the best test approach might not be optimal for predicting dementia in CIND individuals is that, by definition, CIND itself is a cognitively heterogeneous condition. Individuals diagnosed as CIND present with a variety of diverse impairments in cognitive domains such as memory, language, visuospatial abilities, and executive functioning. These possible patterns of cognitive impairment would make it unlikely that a single test could predict who will progress to dementia. Consider the following example of two CIND individuals, one who presents with impaired memory performance and the other presents with a deficit in language abilities. It is entirely possible that both of these individuals are at risk for dementia, but a single test that focuses on one cognitive function would not identify them both.

The subgroup approach. The goal of this approach is to identify subgroups that share one or more features in common, such as a test profile, for example. With respect to predicting dementia, the first step is to identify subgroups of individuals with distinct profiles of neuropsychological test performance, and the second step is to follow these subgroups longitudinally to determine whether they differ in their risk for dementia. Subgroups are typically identified using empirically based methods such as cluster analysis (Martin et al., 1986; Ritchie, Leibovici, Ledesert, & Touchon, 1996). A detailed overview of the conceptual issues and different types of cluster analyses is provided in Appendix 2.1, and thus, only a brief overview of the different methods of forming subgroups is presented here. The two most common ways of defining subgroups are by overall profile elevation or profile shape (Hair, Anderson, Tatham, &
Black, 1998). Profile elevation refers to whether an individual's overall profile is composed of low-, intermediate-, or high-test scores. Profile shape refers to the pattern of ups and downs in an individual's profile relative to his or her mean level of performance, and therefore, provides information regarding the individual's pattern of strengths and weaknesses.

From a clinical assessment point of view, defining subgroups based on profile elevation seems to be a rationale choice: Subgroups that are characterized by overall low performance profiles are more likely to convert to dementia than subgroups with intermediate or high performance profiles. This method would be good at capturing individuals who are already on the cusp of converting to dementia, and thus have good short-term prognostic value (e.g., over a period of 6 to 12 months). To identify individuals destined to develop dementia at the earliest possible stage, however, it might be more useful to define subgroups based on profile shape rather than elevation. By focusing on profile shape, it might be possible to gain insight into which individuals are likely to develop dementia over a 5- to 10-year period.

Neuropsychological Prediction of Dementia: The Findings

Neither the best test approach nor the subgroup approach has been tried with CIND individuals with the exception of the investigation by Hogan and Ebly (1999) which found that low scores on a general cognitive screening instrument were a significant predictor of dementia in this population. Both of these approaches, however, have been used with cognitively healthy adults and with populations that are similar to CIND. A summary of the findings from these two approaches might therefore provide insight into what one would expect to find with CIND.

Best test approach. A wide variety of neuropsychological predictors of dementia have been identified, including tests of memory (Albert et al., 2001; Bozoki, Giordani, Heidebrink, Berent, & Foster, 2001; Daly et al., 2000; Devanand et al., 1997; Flicker, Ferris, & Reisberg, 1991; Kluger et al., 1999; Morris et al., 2001; Petersen et al., 1995; Tierney et al., 1996; Tuokko, Vernon-Wilkinson, Weir, & Beattie, 1991), tests of attention (Daly et al., 2000; Devanand et al.,
1997; Kluger et al., 1999; Morris et al., 2001; Tierney et al., 1996), tests of language (Bozoki et al., 2001; Devanand et al., 1997; Flicker et al., 1991; Morris et al., 2001), tests of visuospatial abilities (Devanand et al., 1997; Flicker et al., 1991; Morris et al., 2001), and tests of executive functioning (Albert et al., 2001; Daly et al., 2000; Devanand et al., 1997; Flicker et al., 1991; Morris et al., 2001). Despite the finding of multiple predictors, there has been little consistency across studies regarding which test is the best predictor of dementia (Tuokko et al., 2000). The inconsistency of predictors is more pronounced for tests that preferentially tap cognitive domains other than memory. For example, some studies have shown that performance on the Similarities subtest from the Wechsler Adult Intelligence Scale-Revised, a measure of abstract reasoning, is a significant predictor of dementia (Elias et al., 2000; Jacobs et al., 1995), whereas others have not (Devanand et al., 1997; Masur et al., 1994). The fact that no single test has been identified that consistently predicts dementia suggests that neuropsychologists will have to continue administering a wide variety of tests that cover each of the basic cognitive domains.

As pointed by Tuokko and Frerichs (2000), differences in methodology might explain some of the inconsistency of predictors of dementia in the literature. Studies in this area often differ in how participants are selected, the sample size, the neuropsychological tests administered, the duration of follow-up, and how the outcome is defined (e.g., cognitive decline versus dementia). All of these factors alone or in combination may help to explain why some neuropsychological tests are significant predictors in one study but not in others.

The subgroup approach. Several investigations have used the subgroup approach with cognitively healthy adults (Fisher, Rourke, & Bieliauskas, 1999; Larrabee, Levin, & High, 1986; Malec, Smith, Ivnik, Petersen, & Tangalos, 1996; Mitrushina, Uchiyama, & Satz, 1995; Valdois, Joanette, Poissant, Ska, & Dehaut, 1990) and with nondemented individuals with cognitive impairment (Reischies & Hellweg, 2000; Ritchie et al., 1996; Ylikoski et al., 1999). Regarding the latter category, participants in the Reischies and Hellweg study were diagnosed as having
mild cognitive disorder (World Health Organization, 1993; see Table 2.1). The sample used by Ritchie et al. comprised individuals with subclinical impairment, which included individuals who were deemed to be normal but for whom a relative had noticed some change in their cognitive abilities over the previous year. Ylikoski et al. used a mixed sample of healthy adults and individuals diagnosed with age associated cognitive decline (Levy, 1994; see Table 2.1) or mild cognitive impairment (Petersen et al., 1995; see Table 2.1).

The number of subgroups that have been identified in the literature has ranged from 2 to 16, with most studies typically reporting five or six. The majority of investigations (Larrabee et al., 1986; Malec et al., 1996; Mitrushina et al., 1995; Reischies & Hellweg, 2000; Ritchie et al., 1996; Valdois et al., 1990; Ylikoski et al., 1999) have used cluster methods that form subgroups primarily on the basis of profile elevation (i.e., hierarchical methods with squared Euclidean distance or the k-means procedure with Euclidean distance). Subsequently, subgroups have tended to differ mainly in their overall level of impairment; most researchers have identified low-, intermediate-, and high-performing subgroups. Some qualitative differences have appeared among the lower performing subgroups in these studies: subgroups have been identified with predominately impairments in learning and memory abilities (Larrabee et al., 1986; Malec et al., 1996; Mitrushina et al., 1995; Reischies & Hellweg, 2000; Ritchie et al., 1996), in visuospatial abilities (Malec et al., 1996; Mitrushina et al., 1995; Valdois et al., 1990; Ylikoski et al., 1999), and in language abilities (Mitrushina et al., 1995; Ritchie et al., 1996; Valdois et al., 1990). The only study to use cluster methods that form subgroups based on profile shape was conducted by Fisher et al. (1999), who identified two subgroups of cognitively healthy adults: one with a relatively flat profile and the other with isolated memory impairment.

Four of the eight studies in this area have followed subgroups over time to monitor changes in cognitive status. Two of these investigations (Reischies & Hellweg, 2000; Ritchie et al., 1996) reported evidence of cognitive decline or an increased risk of dementia in the lower
performing subgroups over periods of less than 5 years. The other two investigations (Larrabee et al., 1986; Malec et al., 1996) reported no evidence of cognitive decline in any of the identified subgroups. These mixed findings could be due to the fact that Reisches and Hellweg and Ritchie et al. used nondemented individuals with cognitive impairment and Larrabee et al. and Malec et al. used cognitively healthy adults.

*Overall summary of findings.* Investigations using the best test approach have met with mixed results: A wide variety of neuropsychological tests have been identified as significant predictors of dementia but no single test has been isolated that consistently predicts this outcome. The majority of investigations using the subgroup approach have identified clusters that differ primarily in overall profile elevation and the predictive ability of these subgroups has been mixed.

*Overview of the Present Investigation*

The overall goal of the present investigation was to examine whether the cognitive heterogeneity of CIND could explain the prognostic heterogeneity associated with this population. The first step toward achieving this goal was to characterize the neuropsychological test performance of a large sample of clinic-referred CIND participants. Clinical comparison data for various demographically defined groups of CIND individuals and data that describe the amount of intertest scatter and overall level of cognitive impairment are provided in Chapter 3.

The second step towards achieving the overall goal of this investigation was to use the subgroup approach in two independent samples of CIND participants. There were two reasons for choosing the subgroup approach over the best test approach. First, the best test approach would not likely be optimal for predicting dementia in CIND individuals, who by definition are cognitively heterogeneous. Second, the subgroup approach seems ideal for exploring and categorizing the variety of cognitive deficits observed in this population. The specific objectives in Chapter 4 were, first to try to identify subgroups with distinct profiles of neuropsychological
test performance in two large cohorts of CIND individuals, and second, to determine whether these subgroups are associated with different diagnostic outcomes (e.g., dementia versus no dementia).

The methods used in the present investigation to identify subgroups differed from previous studies in several important ways. First, the clustering variables used here were derived from principal component analyses of neuropsychological test scores after removing the variance due to age, education, and sex. This was done to control for the confounding effects of these demographic variables on neuropsychological test performance and to avoid the negative impact of multicollinearity among variables in the cluster analysis.

The second difference between the present investigation and previous studies is that the method used to classify individuals into subgroups was based on profile shape rather than elevation. The factor scores obtained from the principal component analysis were standardized by case to produce ipsative factor z-scores. To compute these scores, the mean and standard deviation of all the factor scores were determined for each individual. Each of the original factor scores was then subtracted from the mean factor score and this result was divided by the standard deviation to generate ipsative factor z-scores. Using this method, it was possible to identify subgroups with profiles based on individual patterns of cognitive strengths and weaknesses rather than on the overall level of performance.

The present investigation was also the first in this area of research to use two independent samples. One of the samples was a population-based sample from the CSHA (CSHA Working Group, 1994a), while the other was a clinic-based sample from the ACCORD study¹ (Feldman et al., 2002). By employing two samples, it was possible to determine whether the results obtained from one sample generalized to an independent sample. The decision to use population- and

¹ My role in the ACCORD study consisted of administering the neuropsychological test battery to participants at the Vancouver site as well as reviewing data protocols from all sites and creating a central neuropsychological database.
clinic-based samples was guided by findings from previous research showing that baseline level of impairment and incidence rates for dementia can differ markedly between these two types of samples (for reviews, see Bischkopf, Busse, & Angermeyer, 2002; Collie & Maruff, 2002; Tuokko & Frerichs, 2000). To make generalizations about the CIND population, it is therefore necessary to examine both sample types.

To assess whether subgroup membership had prognostic value, the proportion of individuals who developed dementia, who remained stable, and who improved was determined for each subgroup. Logistic regression analyses were then performed to determine whether subgroup membership at baseline was a significant predictor of diagnostic outcome (i.e., dementia versus no dementia). The relative risk of conversion for each subgroup was assessed using odds ratios and 95% confidence intervals.
Appendix 2.1

Cluster Analysis: A General Overview

Cluster analysis refers to a set of empirical techniques that form clusters of participants so that the between-cluster variance is as large as possible and the within-cluster variance is as small as possible (Hair et al., 1998). This goal can be achieved by many different methods (for good overviews, see Blashfield & Aldenderfer, 1988; Everitt, 1980; Hair et al., 1998). Several key points are discussed here to provide the reader with some basic information on cluster analysis. These points focus on some of the basic issues that need to be considered before performing a cluster analysis, including which cluster method will be employed, which measure of similarity will be used, and which variables will be entered into the cluster analysis.

Different types of cluster methods. Many of the more commonly used cluster methods can be divided into hierarchical and nonhierarchical variants (Hair et al., 1998). Hierarchical cluster methods can start out with each participant representing one cluster and then proceed to group together cases until there is only one cluster that comprises all cases. These agglomerative hierarchical methods contrast with divisive hierarchical methods, which start out with one cluster comprising all cases and then proceeds to divide the clusters until each case is a single cluster. Regardless of whether agglomerative or divisive hierarchical methods are used, the researcher usually examines the clusters that are formed for a range of solutions (e.g., two to eight clusters) and then chooses one of these solutions using some predetermined set of criteria (Hair et al., 1998).

Unlike hierarchical cluster methods, nonhierarchical methods do not produce tree-like cluster solutions ranging from one cluster for each participant to one cluster comprising all cases. Rather, these methods require the investigator to specify a priori how many clusters to form, and thus, only one cluster solution is produced. Nonhierarchical methods such as the k-means method typically assign cases to clusters on the basis of some pre-specified distance threshold (Everitt,
The end result is a cluster solution containing the prespecified number of clusters and these clusters are formed in such a way that maximizes the distance between cases in each cluster. Nonhierarchical methods tend to be iterative, meaning that the data set is continually scanned and cases are allowed to change cluster membership until the best solution is found. In some cases, the investigator specifies the initial cluster centroids to be used for each of the predetermined clusters. Cluster centroids refer to the average performance on the clustering variables for all participants in a given cluster. Using this approach, one can attempt to replicate a cluster solution obtained in a previous study. Alternatively, one can allow the computer program to randomly select the initial cluster centroids to be used for each cluster.

There is no firm evidence to suggest that hierarchical methods are better than nonhierarchical methods or vice versa; there are advantages and disadvantages for both types of cluster methods (Blashfield & Aldenderfer, 1988; Hair et al., 1998). The major advantage of hierarchical methods is that they provide a reasonably good estimate of how many clusters may be present in a dataset. These methods are therefore a good place to start. One of the disadvantages of hierarchical methods, however, is that they make only one pass through the dataset. Suboptimal solutions can therefore result since cases that are poorly assigned early in the clustering process cannot be reassigned to other clusters. Outliers therefore have a large impact on the cluster solution. One of the main advantages of nonhierarchical cluster methods is that since they are iterative, cases can be reassigned to other clusters to produce more optimal solutions. The disadvantages of nonhierarchical methods is that the investigator has to specify \textit{a priori} how many clusters are to be formed and how the cluster centroids are to be selected.

In some respects, the advantages and disadvantages of hierarchical and nonhierarchical methods compliment one another. A good approach, therefore, is to use a combination of the two types of methods (Hair et al., 1998). One can begin with a hierarchical method to get a sense of how many clusters might be present and then perform a nonhierarchical method to fine-tune the
results obtained using the hierarchical method. To perform this additional step, the researcher uses the results obtained from the hierarchical method to specify the number of clusters to generate and the cluster centroids to be used.

*Measures of similarity.* The vast majority of cluster methods assign participants to clusters on the basis of some measure of similarity or distance. Therefore, one of the necessary steps in performing a cluster analysis is to select a measure of similarity. Measures of similarity that are based on distance, such as Euclidean distance or squared Euclidean distance, are the most widely used (Hair et al., 1998). These distance measures are appropriate when the goal is to form clusters of individuals that differ from one another in overall profile elevation (e.g., low performing versus high performing clusters). However, when the goal is to classify individuals on the basis of profile shape regardless of elevation (e.g., the pattern of ups and downs relative to the individual’s mean level of performance), the correlation coefficient is often used as a measure of similarity. An alternative approach to capture profile shape is to standardize test scores by case so that one can focus on individual patterns of strengths and weaknesses in performance (Livingston et al., 1997; Morey, 1991; Moses & Pritchard, 1996). This approach requires the computation of ipsative z-scores and then using distance measures such as squared Euclidean distance.

*Assumptions of cluster analysis.* Cluster analysis is founded on two main assumptions (Hair et al., 1998): First, that the sample being analyzed is representative of the general population, and second, that there is not a high degree of multicollinearity among the clustering variables. The first of these assumptions holds for all statistical procedures and is relatively straightforward. The second assumption, however, requires some additional discussion. Each variable in cluster analysis is weighted equally, and therefore, a high degree of multicollinearity among a subset of variables may result in a solution that is biased in favour of the subset of highly intercorrelated variables (Hair et al., 1998). If several tests are highly correlated with one
another, the investigator should choose only one of the tests, or better yet, create a composite of all tests for the analysis. Using a single composite score rather than several individual test scores will result in less redundancy and random error variance, and increase the reliability and validity of the underlying construct being tapped (Cohen & Cohen, 1983).
CHAPTER 3

Neuropsychological Characterization of Cognitively-Impaired-Not-Demented (CIND) Individuals: Clinical Comparison Data

This chapter is a self-contained manuscript that will be submitted for publication. It has been prepared according to the guidelines set out by the American Psychological Association (APA; American Psychological Association, 2001). To make this manuscript part of my dissertation, I have made three deviations from APA style. First, I have omitted the page-header. Second, I have removed the reference section from the back of the manuscript; all of the references are included in the Reference Section located after Chapter 5. Third, I have numbered the Tables, Figures, and Appendices to indicate the chapter in which they are located (e.g., Table 3.1 versus Table 1).
Abstract

The primary objective of the present investigation was to characterize the neuropsychological test performance of a large sample of clinic-referred individuals diagnosed as Cognitively-Impaired-Not-Demented (CIND). Participants classified as Not-Cognitively-Impaired (NCI; \( n = 68 \)) differed from CIND individuals (\( n = 205 \)) on a number of demographic, clinical, and neuropsychological measures. Measures of learning and memory, visuoconstruction abilities, and cognitive flexibility provided the best discrimination between NCI and CIND participants.

Clinical comparison data for CIND participants were generated for various demographically defined groups. The results feature the amount of intertest scatter and the overall level of cognitive impairment in CIND individuals, supporting the impression that CIND is a cognitively heterogeneous condition.
Neuropsychological Characterization of Cognitively-Impaired-Not-Demented (CIND)

Individuals: Clinical Comparison Data

*Cognitively-Impaired-Not-Demented (CIND)* is a diagnosis given to individuals who present with cognitive impairment but do not meet formal criteria for dementia according to the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2000; Ebly et al., 1995). A number of other labels have been used for individuals with cognitive impairment in the absence of dementia, the more common ones being *age-associated memory impairment* (Crook et al., 1986), *late-life forgetfulness* (Blackford & La Rue, 1989), and *mild cognitive impairment* (Petersen et al., 1995; Petersen et al., 1999). Compared to these alternatives, CIND is a more comprehensive label that captures individuals with cognitive impairment in one domain or in multiple domains, and it has fewer exclusion criteria than other labels such as *aging-associated cognitive impairment* (Levy, 1994), or *aging-related cognitive decline* (American Psychiatric Association, 2000). The population of CIND may be as heterogeneous as the population of individuals diagnosed with dementia, and we¹ have opted to focus on CIND individuals because our goal was to advance understanding of CIND as an at risk population for dementia. Because CIND allows impairment to be present in any cognitive domain we believe it may capture a more representative sample of individuals who are at risk for dementia.

CIND individuals are not only heterogeneous with respect to their cognitive impairment; the diagnostic outcome of these individuals is also varied, although still poorly understood. At least three diagnostic outcomes are possible for CIND individuals: their cognitive impairments may become worse, they may remain stable, or they may improve over time. In the Canadian Study of Health and Aging (CSHA Working Group, 1994a; Hogan & Ebly, 1999; Tuokko et al.,

¹ Although I was the primary author of this manuscript, the version to-be-submitted for publication will have the following three additional authors: Peter Graf, Sherri Hayden, and Howard Feldman.
2000), for example, 45.5% of individuals initially classified as CIND progressed to dementia over a 5-year period (the conversion rate in CIND is considerably higher than the rate of 14.5% that is observed in individuals initially classified as Not-Cognitively-Impaired or NCI). By contrast, over the same period, 42.1% remained stable and 12.4% actually improved and were later classified as NCI. These results do provide some support for the assumption that CIND is a prodromal phase of dementia as some have suggested (Rediess & Caine, 1996; Tuokko & Frerichs, 2000). More importantly, however, these results demonstrate that not all CIND individuals develop dementia, and thus, a better understanding of the mechanisms that determine diagnostic outcomes in this population is needed. Elucidation of the mechanisms underlying the prognosis of CIND individuals will ultimately enable one to identify, prospectively, which individuals will and will not convert to dementia.

The long-term goal of our research program is to illuminate the heterogeneity of cognitive profiles and diagnostic outcomes in CIND individuals. More specifically, we would like to determine whether the cognitive heterogeneity demonstrated by CIND individuals is able to account for the prognostic heterogeneity associated with this population. The ability to use baseline cognitive profiles to identify those CIND individuals at high risk for developing dementia would be invaluable. Treatment interventions that may stabilize or even enhance cognitive functioning and quality of life could be implemented earlier, prior to progressing to a full-blown dementia. Limited research funding and health care resources could be channelled more efficiently to those CIND individuals most at risk for conversion to dementia. In general, the early identification of high-risk CIND individuals would lessen the personal and social costs associated with dementia.

As a first step toward our long-term goal we must be able to identify CIND individuals in a valid and reliable manner, and the present investigation was designed to achieve this objective. We have adopted two different, complementary approaches for this purpose. First, we generated
the first comprehensive collection of clinical comparison data for CIND individuals. In contrast to normative data that are used to describe the performance of cognitively healthy individuals on a given test or test battery, clinical comparison data describe how a specified clinical population performs on a given test or test battery (Mitrushina, Boone, & D'Elia, 1999). The use of normative and clinical comparison data allow clinicians to determine whether an individual is cognitively impaired relative to persons of similar age, level of education, and sex.

The second approach that we used to enhance the identification of CIND individuals compared the performance of NCI and CIND individuals on various neuropsychological tests. The goal here was to identify which neuropsychological tests best differentiate NCI and CIND individuals. Steenhuis and Ostbye (1995) showed that for a population-based sample, tests of memory, judgement, verbal fluency, and visuoconstruction were able to discriminate NCI and CIND individuals. By contrast, our investigation used a clinic-based sample because this is the group of individuals faced by clinicians who are forced to differentiate between NCI and CIND. By focusing on a different population, our study can be viewed as a test of whether the findings of Steenhuis and Ostbye can be replicated and extended to a clinic-based sample of NCI and CIND individuals.

Method

Participants

The participants in the present investigation were part of the Canadian Collaborative Cohort of Related Dementias (ACCORD) study (Feldman et al., in press). The ACCORD is a longitudinal study of individuals referred to one of eight specialized dementia research clinics across Canada for cognitive assessment. At baseline, there were 124 NCI and 342 CIND participants recruited into the ACCORD study. The average age in years for NCI and CIND participants was 61.26 years ($SD = 11.73$; range = 30-84) and 66.35 years ($SD = 10.90$; range = 37-91) respectively. The average number of years of education for these two groups was 13.01
and 12.42 (SD = 4.13) respectively. The percentage of females was slightly higher in the NCI group (57.26%) than in the CIND group (50.00%). Informed consent was obtained from all participants before entry into the study.

Every effort was made to ensure that all NCI and CIND participants completed the neuropsychological test battery. However, several of the test sites experienced periods when access to neuropsychological services was limited or not available, and thus, a total of only 287 (61.59%) NCI and CIND participants received the baseline neuropsychological battery. Reasons for not administering the neuropsychological tests included: participant refusal (11.82%), testing not ordered (3.43%), testing performed prior to entering study or in another centre (1.93%), withdrew from study before testing could be done (1.07%), no show (0.64%), language barrier (0.21%), and reasons not stated (19.31%). Compared to those who received the neuropsychological tests (n = 287), participants not tested (n = 179) were significantly older (M = 67.45 years, SD = 12.05 vs. M = 63.46 years, SD = 10.60), t(464) = 3.75, p < .001, less educated (M = 11.87 years, SD = 4.14 vs. M = 13.02 years, SD = 3.96), t(463) = 3.01, p = .003, and scored significantly lower on the Mini Mental Status Exam (M = 26.73, SD = 3.25 vs. M = 27.77, SD = 2.20), t(455) = 4.07, p < .001 (MMSE; Folstein et al., 1975). However, on the Functional Rating Scale (Crockett, Tuokko, Koch, & Parks, 1989), participants not tested did not differ from those who received the battery (M = 13.17, SD = 4.15 vs. M = 12.95, SD = 3.07), t(464) = 4.07, p = .502. In view of these obtained differences, the findings reported here may be biased in favour of slightly higher functioning persons.

Of the 287 participants who received neuropsychological tests, a further 14 were excluded because they did not provide sufficient data for the purposes of this study (see the Procedure section below for details). The final sample of NCI and CIND participants that provided data reported here consisted of 68 and 205 respectively. Demographic and clinical information for these two groups are given later in this report (see Table 3.1).
**Diagnostic evaluation.** A neurologist or geriatric physician made a preliminary diagnosis of dementia or no dementia using modified criteria from the Diagnostic and Statistical Manual of Mental Disorders, Third Edition-Revised (DSM-III-R; American Psychiatric Association, 1987). The modification allowed for impairment in either short- or long-term memory rather than requiring impairment in both types of memory. Participants designated as not demented were further classified as being either NCI if there were no positive DSM-III-R items, or CIND if there was at least one positive DSM-III-R item. After receiving a preliminary diagnosis of NCI, CIND, or dementia, participants underwent additional testing, including neuropsychological testing, laboratory work-up, and neuroimaging where indicated. Based on the results of this additional testing, the physician then made a final diagnosis.

**Test Instruments**

A comprehensive neuropsychological test battery was administered to NCI and CIND participants. The test battery comprised a number of standardized neuropsychological tests and took approximately 2 hours to complete. The tests used in this investigation, which are described below, were selected to tap the DSM-III-R criteria for dementia (American Psychiatric Association, 1987), which requires impairment in memory and at least one of the following areas: abstract thinking, judgement, or higher cortical functions (e.g., aphasia, agnosia, apraxia). All tests were administered according to published methods and instructions.

**Memory.** Four different tests were used to tap this criterion, including the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987), the Logical Memory subtest from the Wechsler Memory Scale-Revised (WMS-R; Wechsler, 1987), the Rey Complex Figure Test (RCF; Meyers & Meyers, 1995), and the Wechsler Adult Intelligence Scale-Revised

---

2 Although all participants were referred for clinical evaluation, some were found to be NCI after clinical examination. Individuals classified as NCI in this investigation therefore, demonstrated no clinically significant cognitive deficits. Some of these participants may have been overly concerned about experiencing normal age-related cognitive declines. Others have referred to similar individuals as the “worried well” (Del Guercio, 1971; Petersen, 2000b).
(WAIS-R; Wechsler, 1981) Digit Span subtest. A subset of variables from each of these tests was selected for this investigation. These variables included the total words correctly recalled during Trials 1 to 5 from List A (Total) and the long delay free recall of List A (LDFR) from the CVLT, the immediate recall (IR) and delayed decall (DR) scores from the WMS-R Logical Memory, the RCF DR, and the total score (forwards and backwards) from the WAIS-R Digit Span subtest.

Abstract thinking. This criterion was assessed using the total score from the WAIS-R Similarities subtest.

Judgment. The total score from the WAIS-R Comprehension subtest provided an index for this diagnostic criterion.

Disturbances in higher cortical functions. Several test instruments were used to tap this diagnostic criterion. Language abilities were assessed by the number of items correctly named either spontaneously or with stimulus cues from the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983) and the total number of words generated for the letters F, A, and S from the Controlled Oral Word Association Test (COWAT; Spreen & Benton, 1969, 1977). Visuoconstruction abilities were tested using the total scores of the Copy trial for the RCF test and the WAIS-R Block Design subtest. The time to complete Parts A and B from the Trail Making Test (TMT; Army Individual Test Battery, 1944) provided indices of psychomotor speed and cognitive flexibility, respectively.

Procedure

A detailed description of the methods and procedures of the ACCORD study is provided elsewhere (Feldman et al., in press), and thus, only a brief overview is given here. Eight specialized dementia research clinics across Canada took part in the study: Dalhousie University, Laval University, McGill University, University of Ottawa, University of Toronto, University of Western Ontario, University of Calgary, and the University of British Columbia. All newly
referred patients to each clinic were approached for recruitment into the study. The recruitment phase took place between September 1997 and May 1999. After obtaining informed consent, participants received a comprehensive baseline clinical assessment including: The Cumulative Illness Rating Scale (Conwell, Forbes, Cox, & Caine, 1993), the Functional Rating Scale, the Global Deterioration Scale (Reisberg et al., 1982), a vascular dementia checklist (Rockwood et al., 2000), the Mini Mental Status Exam (Folstein et al., 1975), the Disability Assessment for Dementia (Gelinas, Gauthier, McIntyre, & Gauthier, 1999), the Caregiver Burden Scale (Gerritsen & Van Der Ende, 1994), a short form of the Geriatric Depression Scale (Burke, Roccaforte, & Wengel, 1991; Yesavage, Brink, Rose, & Adey, 1986), the Neuropsychiatric Inventory (Cummings et al., 1994), the Quality of Life Short Form 36 (Hobson, Bhowmick, & Meara, 1996), and a modified version of the Canadian Assessment of Utilization of Services Tracking (Ostbye & Crosse, 1994).

Participants were followed for two subsequent annual visits. However, this report describes the neuropsychological data for the baseline assessment only. Data from the subsequent visits are presented elsewhere (Peters, Graf, Hayden, & Feldman, in preparation-b).

**Neuropsychological testing.** Psychologists with a mean of 12.50 years of practical experience (SD = 5.59 years) in neuropsychological assessment were responsible for training psychometrists in the administration of the test battery, as well as for interpretation of the test results. Each psychometrist had previous testing experience (M = 10.80 years, SD = 7.64 years), and had at least post-secondary education with 69.23% having the equivalent of a Master’s degree or higher.

Participants with French as their first language had the option of being tested in their native language. We used either a published French version or an appropriate French translation of each neuropsychological test for this purpose. There were also several other minor differences in test administration across sites. The IR trial of the RCF was not administered before the DR
trial in 30 (11.54%) of the participants. The WAIS-III was given to 17 (6.18%) participants instead of the WAIS-R. Finally, 81 (29.45%) participants completed the WMS-R Digit Span rather than the WAIS-R Digit Span.

**Procedure for handling missing data.** Missing data is a frequent problem encountered in multicentre longitudinal studies, especially in studies involving elderly participants with cognitive impairments. One common solution is to include only those participants with complete data, but this approach has several limitations (Hair et al., 1998). Most importantly, it reduces the sample size and thus statistical power, often to an unacceptable level. We felt that it would be especially unfortunate to exclude participants with missing data for only a small number of variables. A second limitation of the complete case approach is that the subsample with complete data is often no longer representative of the population (Hair et al., 1998); it may consist of higher functioning individuals, or data may be missing for individuals who wished to discontinue certain tests they perceived as being too difficult. The failure to include lower functioning individuals in the study may bias the data away from those individuals most at risk for converting to dementia.

We used the following methods to accommodate a large portion of cases with incomplete data. First, we excluded the IR trial of the RCF because the extent of missing data for this variable was fairly high (19.86%) relative to the remaining 14 measures (see below), partly because this measure was often omitted at one of the test sites. Second, we excluded a total of 14 (4.88%) participants with missing data on 7 (50%) or more of the 14 remaining variables. The total sample size available after implementing the above steps was 273 NCI and CIND participants, 213 (78.02%) of whom provided complete data. For the remaining 60 participants, we substituted each missing data point with the diagnostic group (i.e., NCI or CIND) mean for the corresponding variable. Regarding the extent of imputed data, 22 (8.06%) participants were missing data on one variable only, 12 (4.40%) were missing two variables, 15 (5.49%) were
missing three variables, 5 (1.83%) were missing four variables, 4 (1.47%) were missing five variables, and 2 (0.73%) were missing six variables. The overall amount of missing data was 3.74%. The extent of missing data did not differ significantly between NCI (2.42%) and CIND (4.18%) participants, \( \chi^2(6) = 8.33, p = .215 \). The extent of missing data for each of the 14 variables is as follows: COWAT (0.37%), Digit Span (0.37%), Block Design (0.73%), Similarities (1.84%), TMT Part A (1.84%), RCF Copy (2.20%), Logical Memory IR (2.20%), Logical Memory DR (2.20%), Comprehension (5.49%), BNT (5.86%), RCF DR (5.86%), TMT Part B (6.96%), CVLT Total (8.06%), and CVLT LDFR (8.42%).

**Overview of Data Analyses**

All data analyses were performed using the Statistical Package for the Social Sciences, Version 10.0 (*SPSS for Windows*, 2002). Given that participants in the present investigation were assessed in either English \( n = 179 \) or French \( n = 94 \), preliminary data analyses were performed to determine whether data from these two groups could be pooled together for subsequent analyses. Participants assessed in English and French were therefore compared to one another on a number of demographic variables and neuropsychological measures.

Participants diagnosed as NCI \( n = 68 \) and CIND \( n = 205 \) were also compared to one another on a number of demographic and neuropsychological measures. With respect to neuropsychological performance, the first step was to compare these two diagnostic groups on each of the neuropsychological measures separately using a series of univariate logistic regression analyses. Alpha levels for all univariate logistic regression analyses were adjusted for multiple comparisons using Bonferroni correction to \( p < .004 \). A backward stepwise logistic regression analysis was then performed to determine which tests best discriminated between NCI and CIND individuals.

Clinical comparison data for CIND individuals were then generated separately for different levels of age, education, and sex. We also provide two indices of intertest performance.
variability in CIND individuals. First, we computed intertest scatter scores, defined as the difference between an individual's highest and lowest test score, and report the frequency distribution of these scatter scores in our sample of CIND individuals. Next, we determined the level of impairment for each CIND individual by tallying the number of test scores falling below 1 standard deviation from the mean in the CIND sample. The percentage of CIND individuals who are impaired on 0 through 7 or more tests is reported.

Results and Discussion

Preliminary Analyses: Language of Test Administration

Participants who completed testing in English \((n = 179)\) and in French \((n = 94)\) were compared on demographic and neuropsychological variables to determine the feasibility of pooling them together for subsequent analyses.

Demographic information. There was no significant difference in age between participants assessed in English \((M = 62.89\text{ years}, SD = 10.85)\) or French \((M = 64.00\text{ years}, SD = 10.37)\), \(t(271) = .833, p = .406\). There was also no significant difference in the percentage of females in each group: English = 46.37\% and French = 56.38\%, \(\chi^2(1) = 2.47, p = .116\). However, participants who were tested in English reported significantly more years of education \((M = 13.57, SD = 3.62)\) than participants tested in French \((M = 12.02, SD = 4.43), t(271) = 3.10, p = .002\). Accordingly, education was entered as a covariate in the subsequent analyses comparing the neuropsychological test performance of participants assessed in English and French.

Neuropsychological test performance. We conducted a series of univariate logistic regression analyses to compare the neuropsychological test performance of participants assessed in English and in French while controlling for education. The alpha level was adjusted for multiple comparisons by the Bonferroni method to \(p < .004\). Participants assessed in English and in French differed on 2 out of the 14 neuropsychological variables only. French participants scored significantly higher \((M = 31.05, SD = 6.08)\) than English participants \((M = 29.74, SD =\)
7.47) on the Copy trial from the RCF, Wald z-ratio = 8.17, \( p = .004 \). In contrast, English participants scored significantly higher (\( M = 51.73, SD = 7.20 \)) than French participants (\( M = 44.81, SD = 10.39 \)) on the BNT, Wald z-ratio = 23.71, \( p < .001 \). There was no significant difference in the proportion of NCI and CIND diagnoses for participants tested in English (CIND = 73.74%, NCI = 26.26%) and French (CIND = 77.66%, NCI = 22.34%), \( \chi^2(1) = .505, p = .477 \), and thus, it appears that any differences in neuropsychological test performance between participants tested in English and in French were not substantial enough to have an effect diagnostically. Participants assessed in English and in French were therefore combined to increase sample size, and thus power, for all subsequent analyses.

The Relation of Diagnostic Group to Demographic and Neuropsychological Variables

Demographic information. Table 3.1 summarizes demographic and screening test data for NCI and CIND participants. CIND participants were significantly older than NCI participants, \( t(271) = 3.62, p < .001 \), but they did not differ significantly in terms of years of education, sex, first language, or handedness, \( p > .05 \) for all tests. Age was entered as a covariate in the subsequent logistic regression analyses that compared the neuropsychological test performance of the NCI and CIND participants.

Neuropsychological test performance. Univariate logistic regression analyses were performed to compare the neuropsychological test performance of participants diagnosed as NCI (\( n = 68 \)) and CIND (\( n = 205 \)) while controlling for age. The alpha level was adjusted for multiple comparisons using the Bonferroni correction with \( p \) set at < .004. Descriptive data on the neuropsychological test performance of NCI and CIND participants and the results of the univariate logistic regression analyses are shown in Table 3.2. As expected, CIND participants performed worse than NCI participants on all 14 neuropsychological measures. After controlling for multiple comparisons, the differences between NCI and CIND participants were no longer significant for the RCF Copy and Part B of the TMT (see Table 3.2 for statistics). Table 3.2
### Table 3.1

**Demographic and Clinical Data for NCI and CIND Participants**

<table>
<thead>
<tr>
<th>Diagnostic Group</th>
<th>NCI ( (n = 68) )</th>
<th>CIND ( (n = 205) )</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic Information</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>59.28 (11.25)</td>
<td>64.58 (10.17)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Education (years)</td>
<td>13.69 (3.66)</td>
<td>12.82 (4.06)</td>
<td>.118</td>
</tr>
<tr>
<td>Sex (% Female)</td>
<td>55.88%</td>
<td>47.80%</td>
<td>.248</td>
</tr>
<tr>
<td>First Language</td>
<td></td>
<td></td>
<td>.598</td>
</tr>
<tr>
<td>English</td>
<td>52.94%</td>
<td>54.15%</td>
<td></td>
</tr>
<tr>
<td>French</td>
<td>35.29%</td>
<td>38.05%</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>11.76%</td>
<td>7.80%</td>
<td></td>
</tr>
<tr>
<td>Handedness (% Right)</td>
<td>95.59%</td>
<td>92.20%</td>
<td>.547</td>
</tr>
<tr>
<td><strong>Clinical Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMSE(^1) Total Score</td>
<td>28.94 (1.21)</td>
<td>27.53 (2.22)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>FRS(^2) Total Score</td>
<td>11.06 (1.98)</td>
<td>13.40 (3.01)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*Note.* Data entries are means with standard deviations in parentheses unless noted otherwise

1. Mini Mental State Exam (maximum score = 30)
2. Functional Rating Scale (maximum score = 40)

* For this variable, data are missing for 3 NCI and 1 CIND participants
Table 3.2
*Mean (and Standard Deviation) of Performance on Neuropsychological Tests with Univariate Logistic Regression Analyses Results for NCI (n = 68) and CIND (n = 205) Participants*

<table>
<thead>
<tr>
<th>Test/Measure</th>
<th>Diagnostic Group</th>
<th>Effect Size</th>
<th>Wald*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NCI</td>
<td>CIND</td>
<td>d</td>
</tr>
<tr>
<td>CVLT1^1 Total2</td>
<td>51.71 (10.19)</td>
<td>40.24 (12.86)</td>
<td>0.94</td>
</tr>
<tr>
<td>CVLT1^1 LDFR3</td>
<td>11.20 (3.22)</td>
<td>7.53 (4.00)</td>
<td>0.96</td>
</tr>
<tr>
<td>WMS-R^4 Logical Memory IR^5</td>
<td>26.42 (7.82)</td>
<td>18.69 (8.01)</td>
<td>0.97</td>
</tr>
<tr>
<td>WMS-R^4 Logical Memory DR^6</td>
<td>22.19 (9.63)</td>
<td>13.37 (9.06)</td>
<td>0.96</td>
</tr>
<tr>
<td>RCF^7 Copy</td>
<td>32.06 (3.53)</td>
<td>29.57 (6.34)</td>
<td>0.43</td>
</tr>
<tr>
<td>RCF^7 DR^6</td>
<td>16.52 (7.26)</td>
<td>10.77 (7.40)</td>
<td>0.78</td>
</tr>
<tr>
<td>WAIS-R^8 Digit Span</td>
<td>15.53 (3.76)</td>
<td>13.59 (3.98)</td>
<td>0.49</td>
</tr>
<tr>
<td>WAIS-R^8 Similarities</td>
<td>20.90 (3.52)</td>
<td>17.69 (5.70)</td>
<td>0.61</td>
</tr>
<tr>
<td>WAIS-R^8 Comprehension</td>
<td>23.61 (3.86)</td>
<td>20.53 (5.85)</td>
<td>0.57</td>
</tr>
<tr>
<td>WAIS-R^8 Block Design</td>
<td>29.15 (8.56)</td>
<td>21.68 (9.79)</td>
<td>0.79</td>
</tr>
<tr>
<td>BNT^9</td>
<td>52.68 (8.73)</td>
<td>48.24 (8.89)</td>
<td>0.50</td>
</tr>
<tr>
<td>COWAT10 ^</td>
<td>39.30 (12.34)</td>
<td>32.05 (12.86)</td>
<td>0.57</td>
</tr>
<tr>
<td>TMT11 Part A</td>
<td>38.78 (16.66)</td>
<td>51.04 (23.46)</td>
<td>0.56</td>
</tr>
<tr>
<td>TMT11 Part B</td>
<td>99.66 (58.80)</td>
<td>132.69 (64.43)</td>
<td>0.52</td>
</tr>
</tbody>
</table>

*Age was entered as a covariate in all regression analyses
^1California Verbal Learning Test
^2Total words correctly recalled during trials 1 to 5 from List A
^3Long delay free recall
^4Wechsler Memory Scale-Revised
^5Immediate recall
^6Delayed recall
^7Rey Complex Figure
^8Wechsler Adult Intelligence Scale-Revised
^9Boston Naming Test
^10Controlled Oral Word Association Test
^11Trail Making Test
*Adjusted p < .004 (following Bonferroni correction)
shows Cohen effect sizes ranging from 0.43 to 0.97, revealing that the performance of CIND individuals was approximately one-half to a full standard deviation below the performance of NCI individuals. Consistent with previous research, the largest differences between NCI and CIND individuals were found for tests of learning and memory (Collie & Maruff, 2000; Petersen et al., 2001). More importantly, these results reveal that compared to NCI persons, the deficits of CIND individuals go beyond memory; the CIND condition affects performance on the majority of commonly used neuropsychological tests.

A backward stepwise logistic regression analysis, using the Wald statistic with $p$-to-remove = .10, was performed to determine which neuropsychological tests best discriminated between NCI and CIND individuals. To assess the stability of the regression analysis results, the entire sample was randomly divided into two subsamples: a Base sample ($n = 223$) and a Holdout sample ($n = 50$). A cross-validation procedure was carried out in which the regression analysis was performed using data from the Base sample (56 NCI and 167 CIND) to identify those measures that best differentiated CIND from NCI. The regression equation generated from this Base sample was then used to classify the remaining 50 participants (12 NCI and 38 CIND) in the Holdout sample. The sensitivity, specificity, and overall accuracy of these analyses are reported. Sensitivity here refers to the proportion of CIND participants correctly identified as CIND by the regression model. Specificity refers to the proportion of NCI participants correctly identified as NCI by the regression model. Overall accuracy refers to the overall proportion of participants correctly identified (i.e., NCI and CIND) by the regression model.

The results of the backward stepwise logistic regression analysis performed on the Base sample are reported in Table 3.3. Of the initial 14 neuropsychological measures entered into the logistic regression analysis, 4 remained in the model, including the CVLT-LDFR, the WAIS-R Block Design, the WMS-R Logical Memory IR, and Part B from the TMT (see Table 3.3 for statistics). Each of these four neuropsychological measures accounted for a significant amount of
Table 3.3
Results of a Backward Stepwise Logistic Regression Analysis for the Base Sample of NCI and CIND Participants (n = 223)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Regression Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>5.77</td>
<td>.001</td>
</tr>
<tr>
<td>Age</td>
<td>.016</td>
<td>.414</td>
</tr>
<tr>
<td>California Verbal Learning Test: Long delay free recall</td>
<td>-.191</td>
<td>.002</td>
</tr>
<tr>
<td>Wechsler Adult Intelligence Scale-Revised: Block Design</td>
<td>-.066</td>
<td>.005</td>
</tr>
<tr>
<td>Wechsler Memory Scale-Revised Logical Memory: Immediate recall</td>
<td>-.062</td>
<td>.020</td>
</tr>
<tr>
<td>Trail Making Test: Part B</td>
<td>-.007</td>
<td>.083</td>
</tr>
</tbody>
</table>

unique variance in differentiating between CIND and NCI participants. Based on these findings, the best discrimination between NCI and CIND individuals is achieved by using measures of immediate recall, delayed recall, visuoconstruction, and cognitive flexibility. These results are consistent with findings reported by Steenhuis and Ostbye (1995) for a population-based sample of NCI and CIND individuals. Unlike the results reported by Steenhuis and Ostbye, however, measures of verbal fluency and judgement were not included in the reduced logistic regression model for discriminating between NCI and CIND in the present investigation. These discrepancies could be due to the fact that the sample examined by Steenhuis and Ostbye was larger and drawn from a different population of CIND individuals -- from the community rather than from dementia clinics.

With the default cut-point of .50 used by SPSS our model was very successful in accurately identifying CIND participants, achieving sensitivity values of 89.82% in the Base sample and of 86.84% in the Holdout sample. However, the specificity values we achieved were quite low in the Base (30.36%) and Holdout (41.67%) samples, underscoring the fact that using the model’s cut-off value of .50 produces an unacceptably high number of false positives. Following the recommendations of Hosmer and Lemeshow (2000), we determined an alternative cut-off value in order to improve the classification ability of the regression model. A receiver-
operating curve (ROC) was first generated to determine the sensitivity and specificity values for each level of the predicted probability values (i.e., the probability of NCI vs. CIND). The sensitivity and specificity values were then plotted against the range of predicted probabilities, and the new cut-off value of .70 was chosen as the point at which the two functions crossed. With this new cut-off, the sensitivity for the Base and Holdout samples were still quite good at 78.44% and 78.95% respectively. The specificity values were a little lower than the sensitivity values for the Base and Holdout samples at 64.28% and 75.00% respectively, but they were much higher than the values achieved with a cut-off of .50. Overall accuracy values were relatively satisfactory at 74.89% and 78.00% in the Base and Holdout samples respectively. (Appendix 3.1 contains information for using our logistic regression analysis results to identify CIND individuals.)

**Clinical Comparison Data for CIND Participants**

In order to produce clinical comparison data that are optimally suited for decision-making in clinical settings, our first step was to assess the relation of three global factors -- age, years of education, and sex -- to neuropsychological test performance. Previous research has revealed that these three factors are important predictors of neuropsychological test performance in a variety of populations (e.g., Alzheimer disease, Parkinson disease, Multiple Sclerosis, head trauma, and stroke) as well as in healthy normal adults (Lezak, 1995; Spreen & Strauss, 1998). If these factors also relate to CIND participants' performance, a useful set of clinical comparison data must capture this relationship and show performance across different levels of each factor. For this purpose, we first used an overall MANOVA, with Wilk's criterion, to examine the relationship of age (3 levels: 38-59 years, 60-69 years, 70-91 years), education (2 levels: 3 to 12 years, 13 to 24 years), and sex (2 levels: male, female) to CIND participants' performance on the neuropsychological test battery. The results of the MANOVA showed significant main effects for age, $F(28, 360) = 3.86, p < .001$, partial $\eta^2 = .231$, education, $F(14, 180) = 5.32, p < .001$,
partial $\eta^2 = .293$, and sex, $F(14, 180) = 3.10, p < .001$, partial $\eta^2 = .194$. No interactions among these factors were significant, $p > .05$ for all. The main effects for age, education, and sex were followed up with univariate ANOVAs or independent $t$-tests with an adjusted alpha value of .004. The results of these follow-up analyses are described below.

Age-grouped data. The clinical comparison data and follow-up univariate ANOVA results for the three age groups of CIND participants are reported in Table 3.4. The results show that, after correction for multiple comparisons, significant declines in performance were present in only 8 of the 14 neuropsychological variables: the CVLT (both Total and LDFR), the WMS-R Logical Memory (both IR and DR), the RCF-DR, the WAIS-R Block Design subtest, and the TMT (both Parts A and B; see Table 3.4 for statistics). In contrast, the relation of age on the remaining six measures was minimal. In terms of effect size, the largest partial $\eta^2$ values were observed for tests of learning and memory, whereas, the smallest partial $\eta^2$ values were present for the COWAT, the WAIS-R Digit Span and Comprehension subtests. This pattern of findings is consistent with previous investigations that have examined age effects on neuropsychological test performance (Lezak, 1995; Mitrushina et al., 1999; Spreen & Strauss, 1998). These data suggest that the cognitive domains examined here are not equally affected by the mechanisms that cause age-related cognitive decline.

Education-grouped data. The clinical comparison data and follow-up independent $t$-test results for the two education groups of CIND participants are reported in Table 3.5. Performance on all 14 neuropsychological measures was nominally higher in the group with at least 13 years of education. After correction for multiple comparisons, the effect of education was present for the following nine measures: the WMS-R Logical Memory IR, the RCF Copy, all of the WAIS-R subtests (Digit Span, Similarities, Comprehension, and Block Design), the BNT, the COWAT, and Part B from the TMT (see Table 3.5 for statistics). Across all 14 measures, Cohen effect sizes ranged from 0.13 to 1.02. The largest effect sizes were observed for the WAIS-R
Table 3.4
Mean (and Standard Deviation) Neuropsychological Test Performance for Various Age Groups with ANOVA Results

<table>
<thead>
<tr>
<th>Test/Measure</th>
<th>38 – 59 yrs</th>
<th>60 – 69 yrs</th>
<th>70 – 91 yrs</th>
<th>F(2,202)</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVLT¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total²</td>
<td>46.61 (12.34)</td>
<td>40.88 (12.36)</td>
<td>33.10 (10.15)</td>
<td>22.74*</td>
<td>.18</td>
</tr>
<tr>
<td>LDFR³</td>
<td>9.49 (3.43)</td>
<td>7.95 (3.59)</td>
<td>5.11 (3.74)</td>
<td>25.92*</td>
<td>.20</td>
</tr>
<tr>
<td>WMS-R LM⁴</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR⁵</td>
<td>20.61 (8.59)</td>
<td>19.83 (8.08)</td>
<td>15.54 (6.32)</td>
<td>8.41*</td>
<td>.08</td>
</tr>
<tr>
<td>DR⁶</td>
<td>16.82 (9.47)</td>
<td>14.44 (8.96)</td>
<td>8.75 (6.59)</td>
<td>16.28*</td>
<td>.14</td>
</tr>
<tr>
<td>RCF⁷</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copy</td>
<td>30.99 (5.54)</td>
<td>28.74 (6.83)</td>
<td>28.99 (6.41)</td>
<td>2.62</td>
<td>.03</td>
</tr>
<tr>
<td>DR⁶</td>
<td>14.49 (7.87)</td>
<td>10.28 (6.84)</td>
<td>7.51 (5.67)</td>
<td>17.77*</td>
<td>.15</td>
</tr>
<tr>
<td>WAIS-R⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Span</td>
<td>13.61 (4.22)</td>
<td>13.80 (4.02)</td>
<td>13.35 (3.74)</td>
<td>0.22</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Similarities</td>
<td>18.43 (5.50)</td>
<td>18.37 (5.42)</td>
<td>16.21 (5.97)</td>
<td>3.42</td>
<td>.03</td>
</tr>
<tr>
<td>Comprehension</td>
<td>20.50 (5.69)</td>
<td>20.84 (5.63)</td>
<td>20.23 (6.30)</td>
<td>.191</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Block Design</td>
<td>24.44 (10.45)</td>
<td>22.13 (10.20)</td>
<td>18.39 (7.57)</td>
<td>6.94*</td>
<td>.06</td>
</tr>
<tr>
<td>BNT⁹</td>
<td>48.81 (8.49)</td>
<td>49.33 (8.21)</td>
<td>46.53 (9.80)</td>
<td>1.92</td>
<td>.02</td>
</tr>
<tr>
<td>COWAT¹⁰</td>
<td>32.12 (11.27)</td>
<td>32.50 (13.41)</td>
<td>31.51 (13.93)</td>
<td>0.10</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>TMT¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part A</td>
<td>42.94 (21.35)</td>
<td>51.36 (21.30)</td>
<td>58.91 (25.18)</td>
<td>8.40*</td>
<td>.08</td>
</tr>
<tr>
<td>Part B</td>
<td>108.27 (47.68)</td>
<td>127.40 (60.20)</td>
<td>163.00 (71.83)</td>
<td>14.15*</td>
<td>.12</td>
</tr>
</tbody>
</table>

¹California Verbal Learning Test
²Total words correctly recalled during trials 1 to 5 from List A
³Long delay free recall
⁴Wechsler Memory Scale-Revised Logical Memory
⁵Immediate recall
⁶Delayed recall
⁷Rey Complex Figure
⁸Wechsler Adult Intelligence Scale-Revised
⁹Boston Naming Test
¹⁰Controlled Oral Word Association Test
¹¹Trail Making Test
*Adjusted p < .004 (following Bonferroni correction)
Table 3.5
Mean (and Standard Deviation) Neuropsychological Test Performance Arranged by Education Group with Independent t-test Results

<table>
<thead>
<tr>
<th>Test/Measure</th>
<th>Education Group</th>
<th>Effect Size</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3 – 12 yrs (n = 105)</td>
<td>13 – 24 yrs (n = 100)</td>
<td>d</td>
</tr>
<tr>
<td>CVLT¹</td>
<td></td>
<td></td>
<td>d</td>
</tr>
<tr>
<td>Total²</td>
<td>38.43 (12.01)</td>
<td>42.14 (13.50)</td>
<td>0.29</td>
</tr>
<tr>
<td>LDFR³</td>
<td>7.27 (3.46)</td>
<td>7.81 (4.50)</td>
<td>0.13</td>
</tr>
<tr>
<td>WMS-R⁴ Logical Memory</td>
<td></td>
<td></td>
<td>d</td>
</tr>
<tr>
<td>IR⁵</td>
<td>16.96 (7.98)</td>
<td>20.50 (7.67)</td>
<td>0.45</td>
</tr>
<tr>
<td>DR⁶</td>
<td>11.82 (8.19)</td>
<td>15.00 (9.66)</td>
<td>0.36</td>
</tr>
<tr>
<td>RCF⁷</td>
<td></td>
<td></td>
<td>d</td>
</tr>
<tr>
<td>Copy</td>
<td>27.90 (6.71)</td>
<td>31.32 (5.43)</td>
<td>0.56</td>
</tr>
<tr>
<td>DR⁶</td>
<td>9.74 (7.21)</td>
<td>11.86 (7.48)</td>
<td>0.29</td>
</tr>
<tr>
<td>WAIS-R⁸</td>
<td></td>
<td></td>
<td>d</td>
</tr>
<tr>
<td>Digit Span</td>
<td>12.36 (3.56)</td>
<td>14.88 (4.01)</td>
<td>0.67</td>
</tr>
<tr>
<td>Similarities</td>
<td>15.15 (6.23)</td>
<td>20.34 (3.50)</td>
<td>1.02</td>
</tr>
<tr>
<td>Comprehension</td>
<td>18.20 (5.96)</td>
<td>22.97 (4.64)</td>
<td>0.89</td>
</tr>
<tr>
<td>Block Design</td>
<td>18.68 (9.95)</td>
<td>24.82 (8.60)</td>
<td>0.67</td>
</tr>
<tr>
<td>BNT⁹</td>
<td>46.09 (8.95)</td>
<td>50.50 (8.28)</td>
<td>0.51</td>
</tr>
<tr>
<td>COWAT¹⁰</td>
<td>27.59 (11.32)</td>
<td>36.73 (12.77)</td>
<td>0.76</td>
</tr>
<tr>
<td>TMT¹¹</td>
<td></td>
<td></td>
<td>d</td>
</tr>
<tr>
<td>Part A</td>
<td>54.79 (25.76)</td>
<td>47.09 (20.16)</td>
<td>0.33</td>
</tr>
<tr>
<td>Part B</td>
<td>151.65 (72.14)</td>
<td>112.78 (48.02)</td>
<td>0.63</td>
</tr>
</tbody>
</table>

¹California Verbal Learning Test
²Total words correctly recalled during trials 1 to 5 from List A
³Long delay free recall
⁴Wechsler Memory Scale-Revised
⁵Immediate recall
⁶Delayed recall
⁷Rey Complex Figure
⁸Wechsler Adult Intelligence Scale-Revised
⁹Boston Naming Test
¹⁰Controlled Oral Word Association Test
¹¹Trail Making Test

*Adjusted p < .004 (following Bonferroni correction)
Similarities and Comprehension subtests, whereas, the smallest effect sizes were present for the CVLT. The effect of education on neuropsychological test performance observed in the present investigation is consistent with previous findings (Lezak, 1995; Mitrushina et al., 1999; Spreen & Strauss, 1998). Our data show that although the relationships to education varies across tests, they are more pronounced than the relationships to age that were reported in Table 3.4.

Sex-grouped data. The clinical comparison data and follow-up independent t-test results for males and females are reported in Table 3.6. There were no significant group differences after correcting for multiple comparisons (see Table 3.6 for statistics). Cohen’s effect sizes ranged from 0.04 to 0.39. Accordingly, the relation of sex to neuropsychological test performance in CIND participants was not as pronounced as it was for age and education, a finding that is consistent with previous research (Lezak, 1995; Mitrushina et al., 1999; Spreen & Strauss, 1998), indicating that sex is not as important to consider as other demographic variables.

Intertest scatter among CIND individuals. The clinical comparison data reported for age, education, and sex are useful for describing how various demographically defined subsets of CIND individuals perform on individual neuropsychological tests. However, there is another important aspect of CIND individuals’ neuropsychological test performance that is not described by these data: the variability or spread of performance across tests for each individual and how this variability may differ between individuals. To capture some aspect of individual cognitive variability, we computed z-scores for each of the neuropsychological variables using the clinical comparison data for different age groups of CIND individuals (see Table 3.4). The difference between each individual’s lowest and highest z-scores was determined to provide an index of within-subject cognitive variability or intertest scatter.

Figure 3.1 is a histogram showing the distribution of resulting scatter scores (highest-lowest z-scores) for CIND individuals. The data reported in Figure 3.1 provide an estimation of how much intertest scatter is typically observed in CIND individuals. Figure 3.1 shows that the
Table 3.6
Mean (and Standard Deviation) Neuropsychological Test Performance for Males and Females with Independent t-test Results

<table>
<thead>
<tr>
<th>Test/Measure</th>
<th>Males (n = 107)</th>
<th>Females (n = 98)</th>
<th>Effect Size</th>
<th>t(203)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVLT¹</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Total²</td>
<td>37.87 (11.05)</td>
<td>42.83 (14.19)</td>
<td>0.39</td>
<td>2.80</td>
</tr>
<tr>
<td>LDFR³</td>
<td>6.98 (6.65)</td>
<td>8.14 (4.30)</td>
<td>0.21</td>
<td>2.10</td>
</tr>
<tr>
<td>WMS-R⁴ Logical Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR⁵</td>
<td>18.94 (7.89)</td>
<td>18.40 (8.17)</td>
<td>0.07</td>
<td>0.48</td>
</tr>
<tr>
<td>DR⁶</td>
<td>13.95 (8.65)</td>
<td>12.74 (9.48)</td>
<td>0.13</td>
<td>0.95</td>
</tr>
<tr>
<td>RCF⁷</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copy</td>
<td>29.70 (6.53)</td>
<td>29.42 (6.15)</td>
<td>0.04</td>
<td>0.32</td>
</tr>
<tr>
<td>DR⁶</td>
<td>11.64 (7.80)</td>
<td>9.83 (6.85)</td>
<td>0.25</td>
<td>1.76</td>
</tr>
<tr>
<td>WAIS-R⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Span</td>
<td>14.10 (4.04)</td>
<td>13.03 (3.87)</td>
<td>0.27</td>
<td>1.95</td>
</tr>
<tr>
<td>Similarities</td>
<td>17.83 (5.72)</td>
<td>17.53 (5.70)</td>
<td>0.05</td>
<td>0.38</td>
</tr>
<tr>
<td>Comprehension</td>
<td>21.25 (6.01)</td>
<td>19.74 (5.60)</td>
<td>0.26</td>
<td>1.86</td>
</tr>
<tr>
<td>Block Design</td>
<td>23.24 (10.22)</td>
<td>19.96 (9.03)</td>
<td>0.34</td>
<td>2.43</td>
</tr>
<tr>
<td>BNT⁹</td>
<td>47.99 (10.18)</td>
<td>48.51 (7.26)</td>
<td>0.06</td>
<td>0.42</td>
</tr>
<tr>
<td>COWAT¹⁰</td>
<td>31.74 (12.52)</td>
<td>32.39 (13.28)</td>
<td>0.05</td>
<td>0.36</td>
</tr>
<tr>
<td>TMT¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part A</td>
<td>50.23 (22.24)</td>
<td>51.92 (24.81)</td>
<td>0.07</td>
<td>0.52</td>
</tr>
<tr>
<td>Part B</td>
<td>130.73 (61.29)</td>
<td>134.83 (67.95)</td>
<td>0.06</td>
<td>0.45</td>
</tr>
</tbody>
</table>

¹California Verbal Learning Test
²Total words correctly recalled during trials 1 to 5 from List A
³Long delay free recall
⁴Wechsler Memory Scale-Revised
⁵Immediate recall
⁶Delayed recall
⁷Rey Complex Figure
⁸Wechsler Adult Intelligence Scale-Revised
⁹Boston Naming Test
¹⁰Controlled Oral Word Association Test
¹¹Trail Making Test
Figure 3.1. Frequency histogram of intertest scatter scores in CIND individuals. To obtain these scatter scores, z-scores were calculated using the data reported in Table 3.4 for various age groups. A scatter score was then computed for each individual by subtracting their lowest z-score from their highest z-score.
majority of CIND individuals (81.46%) had a scatter score of at least 2.00, which would correspond to a difference of 2 standard deviations between an individual's lowest and highest z-scores. These data indicate that the interrelations among test scores in CIND individuals are highly variable and that performance on any one test might not capture overall neuropsychological functioning.

Figure 3.2 illustrates the patterns of neuropsychological test performance that would produce high and low scatter scores respectively. Participant A, with the smallest scatter score, shows very little variability in her neuropsychological profile with most scores being above 0. In contrast, Participant B, with the largest scatter score, shows considerable variability in his neuropsychological profile. In view of the scatter scores summarized in Figure 3.1, it appears that the vast majority of CIND individuals have neuropsychological test profiles that are more scattered than Participant A in Figure 3.2 but less scattered than Participant B. Examination of the performance profiles of all CIND participants in the present study confirms the clinical impression that the CIND label captures a cognitively heterogeneous population. These profiles further underscore the importance of examining performance across multiple tests to assess neuropsychological functioning in CIND individuals. Performance on a single test may reflect overall performance for some but not all CIND individuals.

Level of impairment in CIND. In an attempt to capture further the differences among CIND individuals, we tallied the total number of tests on which each individual scored at least 1 standard deviation below the group mean. For this purpose we used the means and standard deviations for the different age groups reported in Table 3.4. Figure 3.3 shows the cumulative percentage of CIND individuals who are impaired as a function of the number of neuropsychological variables. The figure illustrates that approximately half (49.76%) of CIND individuals are either not impaired on any measure or on one measure only and that 1.50% of CIND individuals are impaired on 10 or more measures. These data demonstrate that rather than
Figure 3.2. Neuropsychological profiles for CIND individuals with the smallest (squares with dashed line) and largest (diamonds with solid line) intertest scatter scores.

Note. CVLT = California Verbal Learning test; Total = Total words correctly recalled during trials 1 to 5 of List A; LDFR = Long delay free recall; LM = Logical Memory; IR = Immediate recall; DR = Delayed recall; RCF = Rey Complex Figure; Digits = Digit Span; Sim = Similarities; Comp = Comprehension; BNT = Boston Naming Test; COWAT = Controlled Oral Word Association Test; Blocks = Block Design; TMT = Trail Making Test.
Figure 3.3. Plot showing the cumulative percentage of CIND individuals for each level of cognitive impairment ranging from 0 to 14 measures. Impairment scores for each individual were computed by tallying the number of test scores that fell below one standard deviation for his or her respective age group (see Table 3.4).
there being a single group of CIND individuals who are characterized by global cognitive impairment, there are groups of CIND individuals who are impaired on only a few select measures and that these groups are not impaired on the same measures. From a practical point of view, these data are useful to clinicians assessing individual patients. If a patient demonstrates impairment on four of the measures used in this investigation, for example, our data indicate that only 20% of CIND individuals would have impairment that surpasses this level.

The intertest scatter and impairment scores reported in this investigation provide useful information for characterizing the neuropsychological performance in CIND individuals. Taken together, these data support the clinical impression that CIND is a cognitively heterogeneous condition. Using a combination of scatter and impairment scores, however, might not even fully capture the degree of cognitive heterogeneity in CIND. Figure 3.4 contains performance profiles for two individuals who have similar scatter and impairment scores. Participant C has a scatter score of 4.03 and an impairment score of 2; participant D has a scatter score of 3.86 and an impairment score of 3. Despite these similarities, the qualitative aspects, or the shape, of these two profiles are quite different. Just as performance on a single test cannot fully capture neuropsychological functioning in CIND individuals, these profiles underscore the fact that no single neuropsychological profile can be used to describe this population either.

General Discussion

The primary objective of the present investigation was to characterize the neuropsychological test performance of a clinic-based sample of CIND individuals in order to facilitate their accurate identification. Two approaches were utilized to accomplish this objective: One was to identify which neuropsychological measures best discriminated between NCI and CIND individuals. The second approach was to provide clinical comparison data for CIND individuals in a way that takes into account the relationship of important demographic variables such as age, education, and sex. The clinical comparison data reported here for CIND are the
Participant C: Scatter score = 4.03; Impairment score = 2

Participant D: Scatter score = 3.86; Impairment score = 3

Figure 3.4. Neuropsychological profiles for CIND individuals with similar intertest scatter scores and impairment scores. Participant C (diamonds with solid line) has a scatter score 4.03 and an impairment score of 2. Participant D (squares with dashed line) has a scatter score of 3.86 and an impairment score of 3. Despite the similarities between these scatter and impairment scores, the shapes of these two profiles are qualitatively different.

Note. CVLT = California Verbal Learning Test; Total = Total words correctly recalled during trials 1 to 5 of List A; LDFR = Long delay free recall; LM = Logical Memory; IR = Immediate recall; DR = Delayed recall; RCF = Rey Complex Figure; Digits = Digit Span; Sim = Similarities; Comp = Comprehension; BNT = Boston Naming Test; COWAT = Controlled Oral Word Association Test; Blocks = Block Design; TMT = Trail Making Test.
first of their kind; they will assist clinicians to identify CIND individuals. Only by accurately identifying CIND individuals is it possible to isolate those who are at risk for developing dementia.

A consistent theme that emerges throughout the present investigation is that CIND is a cognitively heterogeneous condition. The results reported here indicate that there is considerable intertest scatter or within-subject cognitive variability in CIND individuals. The majority of CIND individuals (over 80%) showed a difference of at least 2 standard deviations between their lowest and highest test scores. Regarding overall level of impairment, performance that is greater than 1 standard deviation below the mean is present on only four measures or less in the majority of CIND individuals (80%). These findings indicate that there is not a single homogeneous group of CIND individuals who show global cognitive impairment, but rather, impairment appears to be limited to a few select measures in most CIND individuals.

The cognitive heterogeneity of CIND has important implications for the identification of these individuals and for predicting which individuals will and will not progress to dementia. Many investigators have adopted a single test approach for identifying cognitive impairment (Lezak, 1995). The results of the present investigation however, suggest that the single test approach might be limited for identifying CIND. Although measures of learning and memory, visuoconstruction, and cognitive flexibility were the best at discriminating between NCI and CIND individuals, it is clear from the results of this investigation that poor performance on these measures would not capture all CIND individuals. None of these measures alone or in combination were able to differentiate between NCI and CIND participants with 100% accuracy. In addition, Participant D in Figure 4, for example, performed well on measures of learning and memory, visuoconstruction, and cognitive flexibility but very poorly on other measures such as the Similarities subtest and the COWAT.
The cognitive heterogeneity of CIND individuals might also hinder research efforts focused on identifying neuropsychological predictors of dementia. A wide variety of neuropsychological tests have been found to predict conversion to dementia, including tests of memory (Albert et al., 2001; Bozoki et al., 2001; Daly et al., 2000; Devanand et al., 1997; Flicker et al., 1991; Kluger et al., 1999; Morris et al., 2001; Petersen et al., 1995; Tierney et al., 1996; Tuokko et al., 1991), attention (Daly et al., 2000; Devanand et al., 1997; Kluger et al., 1999; Morris et al., 2001; Tierney et al., 1996), language abilities (Bozoki et al., 2001; Devanand et al., 1997; Flicker et al., 1991; Morris et al., 2001), visuospatial abilities (Devanand et al., 1997; Flicker et al., 1991; Morris et al., 2001), and executive functioning (Albert et al., 2001; Daly et al., 2000; Devanand et al., 1997; Flicker et al., 1991; Morris et al., 2001). With the exception of tests of learning and memory, there has been little consistency in terms of which neuropsychological tests are the best predictors of dementia (Tuokko et al., 2000). One example of this inconsistency is the predictive ability of performance on the WAIS-R Similarities subtest, a measure of abstract reasoning: some investigators have found this test to be a significant predictor of dementia (Elia et al., 2000; Jacobs et al., 1995), whereas other have not (Devanand et al., 1997; Masur et al., 1994). Without doubt, methodological differences such as differences in sample size and the tests used may explain some of this inconsistency. The heterogeneity of cognitive abilities, however, might account for the observed inconsistency in neuropsychological test predictors as well.

The NCI participants that were used in the present investigation were an interesting group that deserves some discussion. These participants represent a comparison group rather than a true control group. These cases were referred to a dementia clinic with cognitive complaints, but were found to have no clinically significant cognitive deficits upon assessment. The results of this investigation are therefore very useful because they reflect the challenge faced by clinicians: assessing individuals who are referred for cognitive assessment and then making a diagnosis.
One important implication of the fact that the NCI participants in the present investigation were not a true control group is that the differences between CIND and cognitively healthy adults may be more pronounced than those reported here.

There were two major limitations of the present investigation. The first limitation pertains to the selection bias of the samples. Both the NCI and CIND samples comprised individuals who were referred to a dementia clinic. Consequently, although these individuals may be representative of the clientele that are referred for evaluation, they are not representative of the larger population of individuals who experience mild cognitive difficulties. In addition, only those NCI and CIND cases that received neuropsychological testing were included in the present investigation. Despite the myriad difficulties associated with multicentre studies, it is imperative for future investigators to ensure that as many participants as possible complete the study protocols.

The second limitation of this study refers to the neuropsychological test battery that was used. Although this battery was relatively comprehensive in its coverage of cognitive domains, it was lacking in some respects, especially in the area of executive functioning. Clearly, there is a need to balance the comprehensiveness and practicality of any study battery; however, future research ought to expand the range of neuropsychological measures used to assess interdomain (e.g., attention, memory, language, visuoconstruction, and executive functioning) and intra-domain (e.g., expressive and receptive language) functioning.

To summarize, the present investigation characterized the neuropsychological test performance of CIND individuals in two very useful ways. First, we provided the first set of clinical comparison data for CIND individuals. These data support the clinical impression that CIND is a cognitively heterogeneous population. Second, we identified the measures that best differentiated between individuals classified as NCI and CIND. The data yielded from both of
these approaches will facilitate the accurate detection of cognitive impairment in older adults and subsequently enhance the identification of CIND individuals.
Appendix 3.1

*Example of How to Use Logistic Regression Results To Facilitate Diagnosing CIND*

The data reported in Table 3 can be used to help facilitate the identification of CIND individuals. Consider the following example to illustrate how one would go about implementing these data for clinical purposes. A 77-year old individual has the following test score results: 6 on the CVLT-LDFR, 21 on the WAIS-R Block Design, 8 on the WMS-R Logical Memory IR, and 132 on Part B of the TMT. The question of interest is whether or not this individual should be classified as NCI or CIND, a decision that can be facilitated by calculating the probability that this individual is CIND. The equation used to estimate the probability of being CIND is

\[
\text{prob}(\text{CIND}) = \frac{e^y}{1 + e^y}
\]

All of the necessary information to plug into the equation is provided in Table 3.3. Multiplying each regression coefficient by its corresponding value and then adding the sum of these products to the constant obtain the value of \(x\) for this equation. In this example, the value of \(x\) would be

\[
5.77 + (.016)(\text{Age}) + (-.191)(\text{CVLT-LDFR score}) + (-.066)(\text{Block Design score}) + (-.062)(\text{Logical Memory IR score}) + (-.007)(\text{Time to complete Part B of TMT})
\]

which works out to be

\[
5.77 + (.016)(77) + (-.191)(6) + (-.066)(21) + (-.062)(8) + (-.007)(132) = 3.05
\]

To obtain the estimated probability of being CIND, one then plugs the value of \(x\) into the above equation. Using a calculator with an \([e^y]\) button on it, one simply pushes the \([e^y]\) button
followed by the value of x, which is 3.05, and then divides this amount by \((1 + e^{x})\). The steps to perform this equation are

Step 1. Press \([e^x]\) button, then the value of x (3.05), which results in 21.11534442

Step 2. Add 1 to \(e^x\), which results in 22.11534442

Step 3. Divide the value in Step 1 (21.11534442) by the value in Step 2 (22.11534442), which works out to be .9548

The estimated probability of the individual in this example being diagnosed as CIND would be .95, which is substantially higher than the classification cut-off of .70. Note that probabilities below the cut-off of .70 would be classified as NCI according to this model.
Acknowledgements

We would like to express our sincerest gratitude to the patients and their families for their commitment to this study. We would also like to acknowledge the ACCORD Study Group and all site coordinators for their hard work in making this study possible. This research is supported by an MRC PMAC program grant # PA14197 to H.F. as well as a Doctoral Training Award given to the first author (K.R.P.) jointly funded by the Alzheimer Society of Canada and the Canadian Institutes of Health Research.

Correspondence concerning this article should be addressed to Kevin Peters at the Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, B.C., Canada, V6T 1Z4 (email: kpeters@interchange.ubc.ca).
CHAPTER 4

Subgroups of Cognitively-Impaired-Not-Demented (CIND) Individuals: Delineation, Reliability, and Predictive Validity

This chapter is a self-contained manuscript that will be submitted for publication. It has been prepared according to the guidelines set out by the American Psychological Association (APA; American Psychological Association, 2001). To make this manuscript part of my dissertation, I have made three deviations from APA style. First, I have omitted the page-header. Second, I have removed the reference section from the back of the manuscript; all of the references are included in the Reference Section located after Chapter 5. Third, I have numbered the Tables, Figures, and Appendices to indicate the chapter in which they are located (e.g., Table 4.1 versus Table 1).
The objectives of the present investigation were to determine whether subgroups of Cognitively-Impaired-Not-Demented (CIND) individuals with distinct neuropsychological profiles exist in two independent samples, and whether subgroup membership is related to diagnostic outcome over periods of 2 to 5 years. A series of cluster analyses was performed on ipsative factor z-scores derived from principal component analyses. Five subgroups were identified in the Base Sample (n = 461): Verbal Dysfunction, Verbal/Visuospatial Dysfunction, Memory/Verbal Dysfunction, Memory Dysfunction, and Visuospatial Dysfunction. This five-cluster solution was replicated in an independent sample of CIND individuals (n = 166). The highest rates of conversion to dementia were observed in the Memory Dysfunction and Memory/Verbal Dysfunction subgroups. The Verbal Dysfunction subgroup was most likely to show improvement in cognitive status. The cognitive heterogeneity of the CIND condition must be taken into account in future research focusing on the early identification of dementia.
Subgroups of Cognitively-Impaired-Not-Demented (CIND) Individuals: Delineation, Reliability, and Predictive Validity

*Cognitively-Impaired-Not-Demented (CIND)* is a diagnostic label applied to individuals who present with cognitive impairment but do not meet formal criteria for dementia according to the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2000; Ebly et al., 1995). A number of other labels have been used for individuals with cognitive impairment in the absence of dementia, the more common ones being *age-associated memory impairment* (AAMI; Crook et al., 1986), *late-life forgetfulness* (LLF; Blackford & La Rue, 1989), *aging-associated cognitive impairment* (AACI; Levy, 1994), *aging-related cognitive decline* (ARCD; American Psychiatric Association, 2000), and *mild cognitive impairment* (MCI; Petersen et al., 1995; Petersen et al., 1999). Compared to these alternatives, CIND is a more comprehensive label that captures individuals with cognitive impairment in one domain or in multiple domains and comes with fewer exclusion criteria than the aforementioned conditions, such as AACI and ARCD.

Our research efforts are focused on CIND, in part, because persons with this condition are cognitively heterogeneous and thus similar to patients with dementia, a condition defined by the presence of impairments in memory and at least one other cognitive domain (American Psychiatric Association, 2000). For the present project, we assume that CIND persons are representative of individuals at risk for dementia. A second reason for our focus on CIND is that individuals with this condition constitute the majority of the older adults who seek help for their cognitive difficulties. Persons who attend memory or dementia clinics tend to exhibit a wide range of cognitive deficits and comorbid physical conditions (Devanand et al., 1997; Reifler,

1 Although I was the primary author of this manuscript, the version to-be-submitted for publication will have the following three additional authors: Peter Graf, Sherri Hayden, and Howard Feldman.
1997), and thus, a better understanding of the varieties of clinical presentation and prognosis of CIND promises to yield a range of clinical benefits.

Data from the Canadian Study of Health and Aging (CSHA) have provided support for the assumption that CIND individuals are at risk for dementia (CSHA Working Group, 1994a; Hogan & Ebly, 1999). Over a five-year period, 45.5% of CIND individuals converted to dementia compared to only 14.5% of individuals initially classified as Not-Cognitively-Impaired (NCI). The CSHA results also revealed that not all CIND individuals develop dementia however. Of the individuals initially diagnosed as CIND, 42.1% maintained their cognitive status over the same 5-year period and 12.4% even showed improvements (Tuokko et al., 2000). This combination of findings underscores the heterogeneity of the CIND condition and thereby highlights the need for valid and reliable prognostic methods.

Two different methods have been used for predicting conversion to dementia in a variety of clinical populations that are similar to CIND. One method aims to find a single instrument, the best test, for making predictions. The other method focuses on the identification of patient subgroups that are distinct in terms of their developmental pathway toward dementia.

The best test approach is reasonable and straightforward, and it has been tried by many investigators (for reviews, see Collie & Maruff, 2000; Tuokko & Frerichs, 2000). This approach has met with mixed results however. A wide variety of neuropsychological predictors of dementia have been identified but no single test has been isolated that consistently predicts this outcome (Tuokko et al., 2000; Tuokko & Frerichs, 2000). The best test approach might not be optimal for CIND because patients present with diverse impairments in higher-level skills (e.g., memory, language, visuoconstruction, and executive functioning). For example, in a recent study we found that CIND individuals were characterized by considerable cognitive variability, the difference between an individual’s highest and lowest test scores was at least 2 standard deviations in the majority of our sample (Peters, Graf, Hayden, & Feldman, in preparation-a).
More importantly, however, was the fact that we identified individuals with the same degree of intertest scatter but with qualitatively different neuropsychological profiles. These results indicated that CIND is characterized by cognitive heterogeneity within and between individuals. Given that a fairly substantial proportion of CIND individuals are expected to progress to dementia and the fact that these individuals may demonstrate a variety of cognitive deficits, it is unlikely that a single test would be able to predict outcome in this population.

We believe the subgroup approach to prognostic decision-making is more likely to succeed with CIND patients. This approach begins with the identification of distinct patient subgroups, and is followed by longitudinal research that aims to document whether the subgroups differ in their risk to develop dementia. Subgroups can be defined in a number of ways. For example, several investigations have classified individuals into subgroups primarily on the basis of overall profile elevation, yielding low-, intermediate-, and high-performing subgroups (Larrabee et al., 1986; Malec et al., 1996; Reischies & Hellweg, 2000; Ritchie et al., 1996). In terms of prognosis, the results of these investigations have been mixed. Some studies have found no differential risk for subsequent cognitive decline among subgroups (Larrabee et al., 1986; Malec et al., 1996); in contrast, other investigators have reported that subgroups initially characterized by low or intermediate overall neuropsychological test performance had the highest rates of conversion to dementia and/or cognitive decline (Reischies & Hellweg, 2000; Ritchie et al., 1996).

The finding that individuals from the lowest performing subgroups were the most likely to convert to dementia is not surprising. It is reasonable to assume that many of the individuals from these subgroups were already on the cusp of conversion at baseline. This method may therefore have short-term prognostic value (e.g., over a period of 6 to 12 months), but is not likely to give insight into who might convert from CIND to dementia over a 5- or 10-year period. This kind of predictive power might be achieved with a version of the subgroup approach that
focuses not on the overall level of performance, but on the pattern or shape of performance. This is the goal of the present project, to investigate whether there are subgroups of CIND individuals with distinct patterns of neuropsychological test performance and whether these subgroups have predictive power.

Method

Participants

*Base sample.* The Base sample comprised 461 CIND individuals from the CSHA. This study has been described in detail elsewhere (CSHA Working Group, 1994a) and therefore will be summarized only briefly here. The CSHA is a population-based study of elderly individuals over the age of 65 years that was initiated in 1991 with 5-year follow-up waves in 1996 and 2001. For the present investigation, we had access to data from only the baseline and first follow-up visit.

The Base sample included individuals across the spectrum of cognitive functioning, including individuals classified as Not-Cognitively-Impaired (NCI), CIND, and demented. Consensus based diagnoses were made by a team of individuals including the principal investigator, a physician, a nurse, and a neuropsychologist. Dementia was diagnosed according to the criteria set out in the revised third edition of the Diagnostic and Statistical Manual (DSM-III-R; American Psychiatric Association, 1987). The diagnostic category of no dementia was further broken down into NCI and CIND. An NCI diagnosis was given to individuals who did not demonstrate any degree of cognitive impairment and a CIND diagnosis was given to participants who demonstrated cognitive impairment that was insufficient to meet criteria for dementia. Cognitive impairment was based on clinical judgement rather than using specified psychometric criteria. The present investigation focuses on only the 461 CIND individuals with complete neuropsychological data from the baseline visit. The mean age of the CIND participants at baseline was 79.37 years ($SD = 7.07$; range = 65-97), and their mean number of
years of education was 8.48 ($SD = 3.74$). There were more female participants (60.09%) than male participants (39.92%).

**Replication sample.** The Replication sample comprised 166 CIND individuals from the Canadian Collaborative Cohort of Related Dementias (ACCORD) study. This study, which has been described in detail elsewhere (Feldman et al., in press), involved eight dementia clinics across Canada and was initiated in 1997 with two consecutive annual follow-up visits. For the present investigation, we were able to access data from the baseline visit and from the second follow-up visit that occurred 2 years later.

Following referral to one of the clinics, participants underwent clinical examinations that included physical and neurological assessments. The DSM-III-R (American Psychiatric Association, 1987) criteria for dementia were used to classify participants as demented or nondemented. Nondemented participants were further classified as NCI or CIND. Participants who were negative on all DSM-III-R criteria were diagnosed as NCI, and a CIND diagnosis was given to participants who were positive on at least one but not all of the necessary DSM-III-R criteria for dementia. Individuals classified as either NCI or CIND were administered additional testing including a neuropsychological battery. The present investigation focuses on only the 166 CIND individuals with complete neuropsychological data from the baseline visit. The mean age of the CIND participants at baseline was 65.10 years ($SD = 10.00$; range = 38-91), and their mean number of years of education was 12.90 ($SD = 3.99$). Approximately half of the participants (47.59%) were female.

**Sample comparison.** The two samples differ in several respects. The Base sample was part of a population-based study, whereas the Replication sample came from a clinic-based study. Participants from the Base sample were significantly older, $t(625) = 19.84$, $p < .001$, and less educated, $t(625) = 12.44$, $p < .001$, than participants in the Replication sample. The
percentage of females was significantly higher in the Base sample than in the Replication sample, \( \chi^2(1) = 7.77, p < .01 \).

In the CSHA and the ACCORD study, an attempt was made to identify the underlying cause of the cognitive impairment. The etiological classification of participants from the Base and Replication samples is reported in Table 4.1. The largest subcategory in both samples was the “other” subcategory, which was composed largely of individuals for whom no underlying cause could be determined. The percentage of individuals with subclassifications of vascular disease and substance abuse was significantly greater in the Base sample than in the Replication sample (see Table 4.1 for statistics). In contrast, there were higher percentages of individuals with the subclassifications of neurological disease and other in the Replication sample than in the Base sample. Unlike the CSHA (Base sample), investigators in the ACCORD study (Replication sample) had the option of indicating more than one cause for each participant’s cognitive impairments; a total of 19 (11.45%) participants in the Replication sample were classified as having mixed etiology.

**Test Instruments**

Participants in both samples completed a similar battery of standard neuropsychological tests, as reported in Table 4.2. The specific test instruments used in each case were selected to tap the DSM-III-R criteria for diagnosing cognitive impairment in dementia. Although the tests completed by the two samples were not identical, they were similar, well matched, and equally suited for applying the DSM-III-R criteria for dementia to CIND. The neuropsychological tests and administration procedures of the CSHA and ACCORD studies have been described in detail elsewhere (Peters et al., in preparation-a; Steenhuis & Ostbye, 1995; Tuokko, Kristjansson, & Miller, 1995).
Table 4.1

*Number (and Percentage) of Participants in Each CIND Subcategory in the Base (n = 461) and Replication (n = 166) Samples*

<table>
<thead>
<tr>
<th>Subcategory</th>
<th>Sample</th>
<th></th>
<th></th>
<th>(\chi^2(1))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base</td>
<td>Replication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Circumscribed Memory Impairment</td>
<td>107 (23.21%)</td>
<td>27 (16.27%)</td>
<td>3.50</td>
<td></td>
</tr>
<tr>
<td>Vascular Disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>34 (7.38%)</td>
<td>11 (6.63%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebral/Stroke</td>
<td>51 (11.06%)</td>
<td>7 (4.22%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>85 (18.44%)</td>
<td>18 (10.85%)</td>
<td>5.13*</td>
<td></td>
</tr>
<tr>
<td>Psychiatric Disorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>38 (8.24%)</td>
<td>18 (10.84%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>29 (6.29%)</td>
<td>3 (1.81%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>67 (14.53%)</td>
<td>21 (12.65%)</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>Substance Abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol</td>
<td>27 (5.86%)</td>
<td>2 (1.20%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug</td>
<td>21 (4.55%)</td>
<td>1 (0.61%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>48 (10.41%)</td>
<td>3 (1.81%)</td>
<td>12.09*</td>
<td></td>
</tr>
<tr>
<td>Neurological Disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parkinson</td>
<td>5 (1.08%)</td>
<td>2 (1.20%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple Sclerosis</td>
<td>1 (0.22%)</td>
<td>1 (0.61%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epilepsy</td>
<td>1 (0.22%)</td>
<td>0 (0.00%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain Injury</td>
<td>0 (0.00%)</td>
<td>4 (2.41%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>7 (1.52%)</td>
<td>7 (4.22%)</td>
<td>4.07*</td>
<td></td>
</tr>
<tr>
<td>Mixed*</td>
<td>--</td>
<td>19 (11.45%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other/Unknown</td>
<td>125 (27.11%)</td>
<td>70 (42.17%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental Retardation</td>
<td>7 (1.52%)</td>
<td>0 (0.00%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socio-Cultural Factors</td>
<td>7 (1.52%)</td>
<td>0 (0.00%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blind/Deaf</td>
<td>8 (1.74%)</td>
<td>1 (0.61%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>147 (31.89%)</td>
<td>71 (42.77%)</td>
<td>6.38*</td>
<td></td>
</tr>
</tbody>
</table>

*Only participants in the Replication sample could be classified as mixed; participants in the Base sample were assigned one cause only.

*\(p < .05\)*
Table 4.2
Neuropsychological Tests Grouped by DSM-III-R Criteria for Dementia in the Base and Replication Samples

<table>
<thead>
<tr>
<th>DSM-III-R Criterion*</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base</strong></td>
<td></td>
</tr>
<tr>
<td>A1. Memory</td>
<td></td>
</tr>
<tr>
<td>Buschke CRT(^1)</td>
<td>WMS-R(^2) LM</td>
</tr>
<tr>
<td>RAVLT(^3)</td>
<td>CVLT(^4)</td>
</tr>
<tr>
<td>WAIS-R(^5) Digit Span (Forward)</td>
<td>WAIS-R Digit Span (Forward)</td>
</tr>
<tr>
<td>B1. Abstract Thinking</td>
<td></td>
</tr>
<tr>
<td>WAIS-R Similarities*</td>
<td>WAIS-R Similarities</td>
</tr>
<tr>
<td>B2. Judgment</td>
<td></td>
</tr>
<tr>
<td>WAIS-R Comprehension*</td>
<td>WAIS-R Comprehension</td>
</tr>
<tr>
<td>B3. Other disturbances of higher cortical functioning</td>
<td></td>
</tr>
<tr>
<td>Token Test (11-item)(^6)</td>
<td>BNT(^7)</td>
</tr>
<tr>
<td>COWAT(^8)</td>
<td>COWAT</td>
</tr>
<tr>
<td>WAIS-R Digit Symbol</td>
<td>TMT(^9) Part A</td>
</tr>
<tr>
<td>3-MS(^10) Pentagon Copy</td>
<td>RCF(^11)Copy</td>
</tr>
<tr>
<td>WAIS-R Block Design*</td>
<td>WAIS-R Block Design</td>
</tr>
</tbody>
</table>

*One of the requirements to diagnose dementia is to have impairment in memory (A1) and in at least one of the other areas of cognitive functioning listed in B1, B2, and B3.

* A short form of this test was used

\(^1\) Modified version of Buschke’s Cued Recall Test (Tuokko et al., 1991)
\(^2\) Wechsler Memory Scale-Revised: Logical Memory (Wechsler, 1987)
\(^3\) Rey Auditory Verbal Learning Test (Spreen & Strauss, 1998)
\(^4\) California Verbal Learning Test (Delis et al., 1987)
\(^5\) Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981)
\(^6\) Abbreviated Token Test (Benton & Hamsher, 1989)
\(^7\) Boston Naming Test: The dependent variable we used was the number of correct spontaneous responses and the number of correct responses with stimulus cues (Kaplan et al., 1983)
\(^8\) Controlled Oral Word Association Test (Spreen & Benton, 1969, 1977)
\(^9\) Trail Making Test (Army Individual Test Battery, 1944)
\(^10\) Modified Mini-Mental Status Exam (Teng & Chui, 1987)
\(^11\) Rey Complex Figure Test (Meyers & Meyers, 1995)
Data Preparation

Neuropsychological test performance is known to be confounded by demographic variables such as age, education, and sex (Lezak, 1995). Multiple regression analyses were therefore performed to remove the variance associated with age, years of education, and sex from each of the dependent variables obtained with the instruments listed in Table 4.2. The residual scores were then submitted to a principal component analysis, with varimax rotation, to reduce the number of variables and to avoid the problem of multicollinearity in the subsequent cluster analysis (Hair et al., 1998). The factor scores of the principal component analyses were then standardized by case to produce ipsative factor z-scores (Livingston et al., 1997; Morey, 1991; Moses & Pritchard, 1996). Ipsative factor z-scores were calculated for each individual by subtracting each factor score from the overall mean of all factor scores and then dividing this amount by the standard deviation of all factor scores. We used ipsative factor z-scores because we were interested in examining profile shape, that is, each individual’s pattern of cognitive strengths and weaknesses, while simultaneously controlling for profile elevation.

Data Analysis: Overview

Cluster analyses. A series of cluster analyses were performed on the ipsative factor z-scores using a variety of different cluster methods. To assess clustering reliability, the solution generated by one procedure was designated as the criterion solution and the results of this solution were then compared to those generated by other cluster methods. To produce the criterion solution, we opted for a combination of hierarchical (Ward’s method) and nonhierarchical (k-means) cluster methods to help overcome the limitations of using either method alone (Hair et al., 1998). This strategy consisted of performing the hierarchical procedure first to get a sense of how many clusters were present. The nonhierarchical analysis was then performed to fine-tune the results obtained from the hierarchical analysis. This additional step is desirable because hierarchical methods make only one pass through the data, and thus, a
suboptimal solution may be generated because cases that are poorly assigned early in the process cannot be reassigned to other clusters (Aldenderfer & Blashfield, 1984). Nonhierarchical methods such as k-means are iterative and therefore allow cases to be reassigned to create a more optimal solution if possible.

To assess between-sample clustering reliability, we determined whether the criterion solution generated in the Base sample could be replicated in the Replication sample by using the k-means procedure with the cluster centroids from the criterion solution. To determine the within-sample clustering reliability in the Base sample, the results of the criterion solution were compared to the results produced by other hierarchical (average-linkage between subjects) and nonhierarchical (k-means with randomly generated cluster centroids) methods, and the Two Step cluster procedure from SPSS, which is a model-based approach to clustering (SPSS for Windows, 2002). Assessing intermethod consistency is important as different methods can sometimes produce different solutions using the same data (Blashfield & Aldenderfer, 1988; Hair et al., 1998).

Profiling clusters on demographic, clinical, and neuropsychological measures. After delineation of the clusters, the identified subgroups were profiled on a number of important variables in the Base and Replication samples. One-way ANOVAs were performed to determine whether there were subgroup differences for age, years or education, overall cognitive status, and individual neuropsychological test scores. Subgroup differences for sex and the various etiological subclassifications reported in Table 4.1 were assessed using chi-square analyses.

Predictive validity of cluster membership. The proportion of CIND individuals who were diagnosed as NCI, CIND, and demented at follow-up was determined for each subgroup in the Base and Replication samples. Logistic regression analyses were performed to determine whether cluster membership at baseline predicted diagnostic outcome (i.e., dementia versus no
 Results

Principal Component Analyses

The decision over the number of factors to extract was based on inclusion of all factors with Eigenvalues greater than 1 and an examination of the scree plot (Hair et al., 1998). Only factor loadings greater than .50 were interpreted. The results of the principal component analyses were very similar in both samples (see Tables 4.3 and 4.4). Three factors each were extracted in the Base and Replication samples accounting for 57.13% and 60.04% of the total variance respectively. These three factors were labelled: Memory Ability, Verbal Ability, and Visuospatial Ability.

In the Base sample, the following four variables loaded on the Memory Ability factor: the total words correctly recalled during Trials 1 to 3 (Total) and the delayed free recall (DR) from the Buschke Cued Recall Test (CRT) and the total words correctly recalled during Trials 1 to 5 (Total) and the delayed free recall (DR) from the Rey Auditory Verbal Learning Test (RAVLT). This factor accounted for the greatest proportion of variance: 23.21%. The Verbal Ability factor received loadings from five instruments: the WAIS-R Digit Span (Forward), Similarities, and Comprehension subtests, the Token Test, and the Controlled Oral Word Association Test (COWAT). This factor accounted for the second highest proportion of variance: 18.72%. The Visuospatial Ability factor accounted for 15.20% of the total variance and received loadings from the WAIS-R Block Design and Digit Symbol Substitution subtests, and the Pentagon Copy trial of the Modified Mini-Mental Status Exam (3-MS).

In the Replication sample, the Memory Ability factor also accounted for the highest proportion of total variance (24.34%), with loadings from the immediate recall (IR) and delayed recall (DR) trials from the Wechsler Memory Scale-Revised Logical Memory (LM) subtest and
Table 4.3

Results of Principal Component Analysis with Varimax Rotation for the Base Sample (n = 461)

<table>
<thead>
<tr>
<th>Test/Measure</th>
<th>Factor Ability</th>
<th>Memory Ability</th>
<th>Verbal Ability</th>
<th>Visuospatial Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRT-Total(^1)</td>
<td></td>
<td>.87</td>
<td>-.07</td>
<td>.11</td>
</tr>
<tr>
<td>CRT-DR(^2)</td>
<td>.85</td>
<td>-.12</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>RAVLT-Total(^3)</td>
<td>.79</td>
<td>.18</td>
<td>-.02</td>
<td></td>
</tr>
<tr>
<td>RAVLT-DR(^4)</td>
<td>.79</td>
<td>.09</td>
<td>-.09</td>
<td></td>
</tr>
<tr>
<td>WAIS-R(^5) Digit Span (Forward)</td>
<td>-.04</td>
<td>.58</td>
<td>-.01</td>
<td></td>
</tr>
<tr>
<td>WAIS-R Similarities</td>
<td>-.09</td>
<td>.61</td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td>WAIS-R Comprehension</td>
<td>.03</td>
<td>.74</td>
<td>-.47</td>
<td></td>
</tr>
<tr>
<td>Token Test</td>
<td>.02</td>
<td>.63</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td>COWAT(^6)</td>
<td>.15</td>
<td>.64</td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td>WAIS-R Block Design</td>
<td>.03</td>
<td>.10</td>
<td>.79</td>
<td></td>
</tr>
<tr>
<td>WAIS-R Digit Symbol Substitution</td>
<td>.17</td>
<td>.35</td>
<td>.68</td>
<td></td>
</tr>
<tr>
<td>3-MS(^7) Pentagon Copy</td>
<td>-.08</td>
<td>.07</td>
<td>.73</td>
<td></td>
</tr>
<tr>
<td>Total Variance Explained</td>
<td>23.21%</td>
<td>18.72%</td>
<td>15.20%</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Only factor loadings greater than .50 are interpreted; these entries are in bold.

\(^1\)Buschke's Cued Recall Test: Total words correctly recalled during trials to 3

\(^2\)Buschke's Cued Recall Test: Delayed free recall

\(^3\)Rey Auditory Verbal Learning Test: Total words correctly recalled during trials 1 to 5

\(^4\)Rey Auditory Verbal Learning Test: Delayed free recall

\(^5\)Wechsler Adult Intelligence Test-Revised

\(^6\)Controlled Oral Word Association Test

\(^7\)Modified Mini-Mental State Exam
Table 4.4

*Results of Principal Component Analysis with Varimax Rotation for the Replication Sample (n = 166)*

<table>
<thead>
<tr>
<th>Test/Measure</th>
<th>Memory Ability</th>
<th>Verbal Ability</th>
<th>Visuospatial Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMS-R LM-IR(^1)</td>
<td>.83</td>
<td>.29</td>
<td>-.15</td>
</tr>
<tr>
<td>WMS-R LM-DR(^2)</td>
<td>.89</td>
<td>.20</td>
<td>-.03</td>
</tr>
<tr>
<td>CVLT-Total(^3)</td>
<td>.79</td>
<td>.09</td>
<td>.31</td>
</tr>
<tr>
<td>CVLT-LDFR(^4)</td>
<td>.82</td>
<td>.03</td>
<td>.27</td>
</tr>
<tr>
<td>WAIS-R(^5) Digit Span (Forward)</td>
<td>.07</td>
<td>.40</td>
<td>.24</td>
</tr>
<tr>
<td>WAIS-R Similarities</td>
<td>.18</td>
<td>.65</td>
<td>.28</td>
</tr>
<tr>
<td>WAIS-R Comprehension</td>
<td>.18</td>
<td>.80</td>
<td>.04</td>
</tr>
<tr>
<td>BNT(^6)</td>
<td>.18</td>
<td>.67</td>
<td>.06</td>
</tr>
<tr>
<td>COWAT(^7)</td>
<td>-.09</td>
<td>.53</td>
<td>.43</td>
</tr>
<tr>
<td>WAIS-R Block Design</td>
<td>.09</td>
<td>.12</td>
<td>.76</td>
</tr>
<tr>
<td>TMT(^8)-Part A</td>
<td>.07</td>
<td>.21</td>
<td>.70</td>
</tr>
<tr>
<td>RCF(^9)-Copy</td>
<td>.11</td>
<td>.13</td>
<td>.74</td>
</tr>
</tbody>
</table>

Total Variance Explained 24.34% 18.04% 17.67%

*Note. Only factor loadings greater than .50 are interpreted; these entries are in bold.
\(^1\)Wechsler Memory Scale-Revised Logical Memory: Immediate recall
\(^2\)Wechsler Memory Scale-Revised Logical Memory: Delayed recall
\(^3\)California Verbal Learning Test: Total words correctly recalled during trials 1 to 5
\(^4\)California Verbal Learning Test: Long delay free recall
\(^5\)Wechsler Adult Intelligence Test-Revised
\(^6\)Boston Naming Test
\(^7\)Controlled Oral Word Association Test
\(^8\)Trail Making Test
\(^9\)Rey Complex Figure Test*
total words correctly recalled during Trials 1 to 5 (Total) and the long delay free recall (LDFR) trial from the California Verbal Learning Test (CVLT). The Similarities and Comprehension subtests, the Boston Naming Test (BNT), and the COWAT loaded highly on the Verbal Ability factor, which accounted for 18.04% of the total variance. Unlike in the Base sample, the factor loading for the Digit Span (Forward) subtest was not interpretable, but the highest loading for this subtest was on the Verbal Ability factor. The Visuospatial Ability factor accounted for 17.67% of the total variance, and received loadings from the Block Design subtest, Part A of the Trail Making Test (TMT), and the Copy trial of the Rey Complex Figure (RCF) test.

Cluster Analyses

Between-sample reliability. A criterion solution was generated using data from the Base sample, according to the following steps. First, we entered the ipsative factor z-scores into a hierarchical cluster procedure to determine the number of possible clusters. For this purpose, we used Ward’s method with squared Euclidean distance. A decision about the number of clusters to select was made by examining the agglomeration coefficients starting at the 10-cluster solution and proceeding down to the 2-cluster solution. A commonly used stopping rule is to select the cluster solution that immediately proceeds the first large increase in the agglomeration coefficient, indicating that two clusters are about to be merged that are quite dissimilar from one another (Hair et al., 1998). Other criteria used to help determine the number of clusters focus on the dendogram and clinical interpretability of the clusters. Based on these criteria, we chose a 5-cluster solution for further analysis. The percent change in agglomeration coefficient going from five to four clusters was 79.93%, almost double the magnitude of the largest preceding changes in agglomeration coefficients. In the second step toward generating the criterion solution, we ran a nonhierarchical (k-means) cluster procedure to fine-tune the results obtained by Ward’s method. Following the recommendations of Hair et al. (1998), a k-means analysis was performed
using the cluster centroids (cluster means of each of the clustering variables, which in this case are the ipsative factor z-scores) from the five-cluster solution produced by Ward’s method.

Each of the five clusters or subgroups identified in the Base sample solution was named according to the factor(s) on which they were most impaired. The factor score profiles for these five subgroups are shown as the solid line profiles in Figures 4.1 to 4.5. The subgroups ordered by cluster size are: *Verbal Dysfunction* \( (n = 98; 21.26\%)\), *Verbal/Visuospatial Dysfunction* \( (n = 98; 21.26\%)\), *Memory/Verbal Dysfunction* \( (n = 91; 19.73\%)\), *Memory Dysfunction* \( (n = 88; 19.09\%)\), and *Visuospatial Dysfunction* \( (n = 86; 18.66\%)\).

Also shown in Figures 4.1 to 4.5 are the results of performing a k-means cluster procedure on the ipsative factor z-scores from the Replication sample using the cluster centroids from the criterion solution in the Base sample. The five subgroups obtained in the Replication sample were highly similar to the results of the criterion solution in the Base sample and included: *Verbal Dysfunction* \( (n = 32; 19.28\%)\), *Verbal/Visuospatial Dysfunction* \( (n = 39; 23.49\%)\), *Memory/Verbal Dysfunction* \( (n = 27; 16.27\%)\), *Memory Dysfunction* \( (n = 34; 20.48\%)\), and *Visuospatial Dysfunction* \( (n = 34; 20.48\%)\). The factor score profiles of the five subgroups from the Replication sample are the dashed line profiles in Figures 4.1 to 4.5. Examination of the factor score profiles in Figures 4.1 to 4.5 reveals that the profile shapes and percentage of participants assigned to each subgroup were very similar across the two samples.

*Within-sample reliability.* An important caveat with many cluster methods is that the solutions produced are sensitive to the ordering of cases in the data file (Aldenderfer & Blashfield, 1984). To address this issue, we re-ordered the cases and performed the same analyses again with the criterion solution achieved for the Base sample. The solution produced after re-ordering of the cases was identical to the solution generated using the original ordering of cases.
Figure 4.1. Factor score profile for Verbal Dysfunction subgroup. Factors scores of 0 indicate average performance for the entire sample; standard deviation for factor scores = 1. Negative factor scores indicate worse performance. Solid line with diamonds represents profile for Base sample ($n = 461$); dotted line with triangles represents profile for Replication sample ($n = 166$). Values in parentheses in legend correspond to the percent of participants in this particular subgroup for each sample.
Figure 4.2. Factor score profile for Verbal/Visuospatial Dysfunction subgroup. Factors scores of 0 indicate average performance for the entire sample; standard deviation for factor scores = 1. Negative factor scores indicate worse performance. Solid line with diamonds represents profile for Base sample (n = 461); dotted line with triangles represents profile for Replication sample (n = 166). Values in parentheses in legend correspond to the percent of participants in this particular subgroup for each sample.
Figure 4.3. Factor score profile for Memory/Verbal Dysfunction subgroup. Factors scores of 0 indicate average performance for the entire sample; standard deviation for factor scores = 1. Negative factor scores indicate worse performance. Solid line with diamonds represents profile for Base sample (n = 461); dotted line with triangles represents profile for Replication sample (n = 166). Values in parentheses in legend correspond to the percent of participants in this particular subgroup for each sample.
Figure 4.4. Factor score profile for Memory Dysfunction subgroup. Factors scores of 0 indicate average performance for the entire sample; standard deviation for factor scores = 1. Negative factor scores indicate worse performance. Solid line with diamonds represents profile for Base sample ($n = 461$); dotted line with triangles represents profile for Replication sample ($n = 166$). Values in parentheses in legend correspond to the percent of participants in this particular subgroup for each sample.
Figure 4.5. Factor score profile for Visuospatial Dysfunction subgroup. Factors scores of 0 indicate average performance for the entire sample; standard deviation for factor scores = 1. Negative factor scores indicate worse performance. Solid line with diamonds represents profile for Base sample (n = 461); dotted line with triangles represents profile for Replication sample (n = 166). Values in parentheses in legend correspond to the percent of participants in this particular subgroup for each sample.
Three additional cluster analyses were performed to assess clustering reliability across different methods in the Base sample. The first of these additional cluster methods was the average linkage (between-groups) hierarchical method with squared Euclidean distance as the measure of similarity. The number of clusters to select was determined by the criteria described above. The SPSS Two Step procedure is a modelling-based approach to clustering that automatically determines the number of clusters to select according to the Bayesian Information Criterion (BIC), an index of model parsimony (*SPSS for Windows*, 2002). The optimal number of clusters selected in the average-linkage and Two Step cluster analyses was five. A k-means (nonhierarchical) analysis was also performed which involved pre-specifying a five-cluster solution to be produced using randomly generated cluster centroids. The cluster solutions produced by these three alternative methods were consistent with the five-cluster criterion solution. Compared to the criterion solution, the total number of cases classified into different subgroups for each of these additional methods was 0 (0%), 6 (1.30%), and 40 (8.67%) for the k-means, Two Step, and average-linkage methods respectively. The corresponding kappa values for these three methods were 1.00, .984, and .892, indicating excellent agreement among the various cluster methods in the assignment of cases to subgroups. Taken together these results indicate that the original five-cluster criterion solution replicates across several different types of cluster methods.

*Characterizing the Five-Cluster Solution: Demographic, Clinical, and Neuropsychological Variables*

Demographic and clinical information for each of the subgroups is reported in Table 4.5. In the Base sample, the subgroups did not differ significantly in terms of age, $F(4, 456) = .521, MSE = 50.24$, $p = .720$, education, $F(4, 456) = .543, MSE = 14.02$, $p = .704$, sex, $\chi^2(4) = 8.42, p = .077$, or on the 3-MS, a screening measure of general cognitive functioning, $F(4, 456) = .724, MSE = 89.74, p = .576$. The subgroups from the Replication sample also did not differ
<table>
<thead>
<tr>
<th>Variable</th>
<th>Subgroup/Clusters</th>
<th>Base</th>
<th>Verbal Dysfunction</th>
<th>Cluster Size</th>
<th>Verbal Dysfunction</th>
<th>Cluster Size</th>
<th>Memory Dysfunction</th>
<th>Cluster Size</th>
<th>Visuospatial Dysfunction</th>
<th>Cluster Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Cluster Size</td>
<td></td>
<td>Base</td>
<td>98</td>
<td>(21.26%)</td>
<td>98</td>
<td>(21.26%)</td>
<td>91</td>
<td>(19.74%)</td>
<td>88</td>
<td>(19.09%)</td>
</tr>
<tr>
<td></td>
<td>Replication</td>
<td>n</td>
<td>32</td>
<td>(19.28%)</td>
<td>39</td>
<td>(23.49%)</td>
<td>27</td>
<td>(16.27%)</td>
<td>34</td>
<td>(20.48%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td>Base</td>
<td>79.82</td>
<td>(6.58)</td>
<td>79.71</td>
<td>(6.95)</td>
<td>79.08</td>
<td>(7.41)</td>
<td>79.63</td>
<td>(7.20)</td>
</tr>
<tr>
<td></td>
<td>Replication</td>
<td>M</td>
<td>67.28</td>
<td>(10.71)</td>
<td>63.95</td>
<td>(10.17)</td>
<td>61.63</td>
<td>(11.20)</td>
<td>66.47</td>
<td>(9.00)</td>
</tr>
<tr>
<td>Education (years)</td>
<td></td>
<td>Base</td>
<td>8.32</td>
<td>(3.70)</td>
<td>8.74</td>
<td>(3.84)</td>
<td>8.69</td>
<td>(3.38)</td>
<td>8.57</td>
<td>(3.75)</td>
</tr>
<tr>
<td></td>
<td>Replication</td>
<td>M</td>
<td>13.50</td>
<td>(4.13)</td>
<td>12.74</td>
<td>(4.54)</td>
<td>12.82</td>
<td>(4.22)</td>
<td>12.44</td>
<td>(3.82)</td>
</tr>
<tr>
<td>Sex (% female)</td>
<td></td>
<td>Base</td>
<td>54.08%</td>
<td></td>
<td>68.37%</td>
<td></td>
<td>65.93%</td>
<td></td>
<td>60.23%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Replication</td>
<td>50.00%</td>
<td></td>
<td>48.72%</td>
<td></td>
<td>48.15%</td>
<td></td>
<td>50.00%</td>
<td></td>
<td>41.18%</td>
</tr>
<tr>
<td>3-MS ^1 Total</td>
<td></td>
<td>Base</td>
<td>77.56</td>
<td>(8.59)</td>
<td>76.39</td>
<td>(10.91)</td>
<td>78.16</td>
<td>(7.58)</td>
<td>76.18</td>
<td>(9.68)</td>
</tr>
<tr>
<td></td>
<td>Replication</td>
<td>M</td>
<td>90.30^</td>
<td>(6.86)</td>
<td>92.00^</td>
<td>(6.21)</td>
<td>87.50^</td>
<td>(14.60)</td>
<td>88.88^</td>
<td>(8.78)</td>
</tr>
</tbody>
</table>

^1Modified Mini-Mental State Exam

^ n = 27  ^ n = 27  ^ n = 20  ^ n = 24  ^ n = 25
significantly with respect to age, $F(4, 161) = 1.54$, $MSE = 98.71$, $p = .194$, education, $F(4, 161) = .319$, $MSE = 16.22$, $p = .865$, sex, $\chi^2(4) = .738$, $p = .947$, or on the 3-MS, $F(4, 118) = 1.27$, $MSE = 76.55$, $p = .285$. These results show that the differences in neuropsychological profiles among subgroups cannot be explained by differences on relevant demographic or clinical variables.

Subgroups from each sample did not differ in terms of the proposed etiological subclassifications reported in Table 4.1 with the single exception of circumscribed memory impairment (CMI) in the Base sample, which just reached statistical significance, $\chi^2(4) = 9.65$, $p = .047$. The Memory/Verbal Dysfunction subgroup in the Base sample had a significantly greater percentage of participants assigned to the CMI subcategory (31.87%) than the Verbal/Visuospatial Dysfunction (15.31%), $\chi^2(1) = 7.25$, $p = .007$, and the Visuospatial Dysfunction (18.60%), $\chi^2(1) = 4.10$, $p = .043$ subgroups. There was also a significantly greater percentage of participants subcategorized as CMI in the Memory Dysfunction subgroup (28.40%) than in the Verbal/Visuospatial Dysfunction subgroup (15.31%), $\chi^2(1) = 4.72$, $p = .030$. These findings indicate that the causes of cognitive impairment reported in Table 4.1 cannot account for the differences among the subgroups identified in this investigation.

Each of the cluster analysis in this investigation was performed using ipsative factor z-scores that were initially derived from principal component analyses. To facilitate interpretation of the five-cluster solutions, we also examined the performance of the subgroups on each of the original 12 neuropsychological variables. Although we expected that subgroups would differ on many of the original neuropsychological variables in view of the cluster methods used, we wanted to determine which of these 12 variables showed the largest and smallest differences among the subgroups. Obviously, these additional analyses are entirely descriptive and should not to be interpreted as testing differences among the subgroups. To address this question, an ANOVA was performed for each of the original 12 neuropsychological variables with subgroup
membership as the independent variable. Significant results were followed-up with Bonferroni adjusted multiple comparisons.

Tables 4.6 and 4.7 contain the descriptive data for each of the five subgroups on the 12 original neuropsychological measures and the ANOVA results for the Base and Replication samples respectively. In the Base sample, there were significant subgroup differences for each of the original 12 neuropsychological measures with effect sizes ranging from .06 (COWAT) to .33 (CRT-DR; see Table 4.6 for statistics). Subgroup differences in the Replication sample were present for all of the neuropsychological variables with the exception of the Digit Span and Similarities subtests (see Table 4.7 for statistics). In terms of effect size, partial $\eta^2$ values ranged from .02 (Digit Span) to .33 (LM-IR). Despite the differences in sample size between the Base and Replication samples, the magnitudes of effect sizes were similar across tests within each battery. These data indicate that the matching of tests in the two samples by DSM-III-R criteria was largely successful; cluster membership explains similar amounts of variance for each of the matched tests.

The pattern of effect size magnitudes across the test battery was consistent with the pattern of factor loadings from the principal component analyses in each of the samples (see Tables 4.3 and 4.4). For each of the three factors in the Base sample, the variables with the highest factor loadings are the same as those with the largest partial $\eta^2$ values: the two CRT variables (Memory Ability), Comprehension (Verbal Ability), Block Design and 3-MS Pentagon Copy (Visuospatial Ability). Similarly, for each of the three factors in the Replication sample, the variables with the highest factor loadings were also those with the largest partial $\eta^2$ values: LM-DR (Memory Ability), Comprehension (Verbal Ability), and Block Design and RCF Copy (Visuospatial Ability). These results are particularly useful for clinicians who wish to use the descriptive data provided in Table 4.6 or 4.7 to classify clients into one of the five subgroups identified by this investigation but, for whatever reason, cannot administer all of the tests listed
Table 4.6  
_Mean (and Standard Deviation) Neuropsychological Test Scores for Each Subgroup in the Base Sample (n = 461)_

<table>
<thead>
<tr>
<th>Subgroup/Clusters</th>
<th>Verbal Dysfunction</th>
<th>Visuospatial Dysfunction</th>
<th>Memory Dysfunction</th>
<th>Verbal Dysfunction</th>
<th>Visuospatial Dysfunction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test/Variable</td>
<td>(n = 98)</td>
<td>(n = 98)</td>
<td>(n = 91)</td>
<td>(n = 88)</td>
<td>(n = 86)</td>
</tr>
<tr>
<td>CRT-Total&lt;sup&gt;1&lt;/sup&gt;</td>
<td>22.45&lt;sub&gt;a,d&lt;/sub&gt;</td>
<td>24.26&lt;sub&gt;c&lt;/sub&gt;</td>
<td>17.20&lt;sub&gt;b&lt;/sub&gt;</td>
<td>14.38&lt;sub&gt;c&lt;/sub&gt;</td>
<td>21.51&lt;sub&gt;d&lt;/sub&gt;</td>
</tr>
<tr>
<td>CRT-DR&lt;sup&gt;2&lt;/sup&gt;</td>
<td>8.67&lt;sub&gt;a,d&lt;/sub&gt;</td>
<td>9.45&lt;sub&gt;a&lt;/sub&gt;</td>
<td>5.97&lt;sub&gt;b&lt;/sub&gt;</td>
<td>4.93&lt;sub&gt;c&lt;/sub&gt;</td>
<td>8.01&lt;sub&gt;d&lt;/sub&gt;</td>
</tr>
<tr>
<td>RAVLT-Total&lt;sup&gt;3&lt;/sup&gt;</td>
<td>26.01&lt;sub&gt;a&lt;/sub&gt;</td>
<td>33.01&lt;sub&gt;b&lt;/sub&gt;</td>
<td>21.60&lt;sub&gt;c&lt;/sub&gt;</td>
<td>22.17&lt;sub&gt;c&lt;/sub&gt;</td>
<td>28.67&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>RAVLT-DR&lt;sup&gt;4&lt;/sup&gt;</td>
<td>3.55&lt;sub&gt;a&lt;/sub&gt;</td>
<td>6.10&lt;sub&gt;b&lt;/sub&gt;</td>
<td>1.79&lt;sub&gt;c&lt;/sub&gt;</td>
<td>2.01&lt;sub&gt;c&lt;/sub&gt;</td>
<td>4.23&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>Digit Span</td>
<td>4.83&lt;sub&gt;a&lt;/sub&gt;</td>
<td>4.94&lt;sub&gt;a&lt;/sub&gt;</td>
<td>5.19&lt;sub&gt;a,b&lt;/sub&gt;</td>
<td>5.85&lt;sub&gt;c&lt;/sub&gt;</td>
<td>5.60&lt;sub&gt;b,c&lt;/sub&gt;</td>
</tr>
<tr>
<td>Similarities</td>
<td>2.72&lt;sub&gt;a&lt;/sub&gt;</td>
<td>3.03&lt;sub&gt;a&lt;/sub&gt;</td>
<td>4.33&lt;sub&gt;c&lt;/sub&gt;</td>
<td>5.91&lt;sub&gt;c&lt;/sub&gt;</td>
<td>4.55&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
<tr>
<td>Comprehension</td>
<td>4.61&lt;sub&gt;a&lt;/sub&gt;</td>
<td>6.06&lt;sub&gt;b&lt;/sub&gt;</td>
<td>6.74&lt;sub&gt;b&lt;/sub&gt;</td>
<td>8.75&lt;sub&gt;c&lt;/sub&gt;</td>
<td>8.15&lt;sub&gt;c&lt;/sub&gt;</td>
</tr>
<tr>
<td>Token Test</td>
<td>29.47&lt;sub&gt;a&lt;/sub&gt;</td>
<td>31.56&lt;sub&gt;a,d&lt;/sub&gt;</td>
<td>34.98&lt;sub&gt;b,d&lt;/sub&gt;</td>
<td>38.45&lt;sub&gt;c&lt;/sub&gt;</td>
<td>34.01&lt;sub&gt;d&lt;/sub&gt;</td>
</tr>
<tr>
<td>COWAT&lt;sup&gt;5&lt;/sup&gt;</td>
<td>15.06&lt;sub&gt;a&lt;/sub&gt;</td>
<td>15.52&lt;sub&gt;a,c&lt;/sub&gt;</td>
<td>17.20&lt;sub&gt;a,c&lt;/sub&gt;</td>
<td>21.59&lt;sub&gt;b,c&lt;/sub&gt;</td>
<td>19.24&lt;sub&gt;c&lt;/sub&gt;</td>
</tr>
<tr>
<td>Block Design</td>
<td>7.96&lt;sub&gt;a&lt;/sub&gt;</td>
<td>4.47&lt;sub&gt;b,d&lt;/sub&gt;</td>
<td>9.10&lt;sub&gt;a&lt;/sub&gt;</td>
<td>6.30&lt;sub&gt;c&lt;/sub&gt;</td>
<td>3.98&lt;sub&gt;d&lt;/sub&gt;</td>
</tr>
<tr>
<td>Digit Symbol</td>
<td>18.09&lt;sub&gt;a&lt;/sub&gt;</td>
<td>14.12&lt;sub&gt;b,c&lt;/sub&gt;</td>
<td>20.35&lt;sub&gt;a&lt;/sub&gt;</td>
<td>18.25&lt;sub&gt;a&lt;/sub&gt;</td>
<td>14.05&lt;sub&gt;c&lt;/sub&gt;</td>
</tr>
<tr>
<td>Copy&lt;sup&gt;6&lt;/sup&gt;</td>
<td>8.94&lt;sub&gt;a,c&lt;/sub&gt;</td>
<td>7.12&lt;sub&gt;b&lt;/sub&gt;</td>
<td>9.40&lt;sub&gt;a&lt;/sub&gt;</td>
<td>8.26&lt;sub&gt;c&lt;/sub&gt;</td>
<td>6.24&lt;sub&gt;d&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

_Note._ Means within each row that do not share the same subscript differ at p < .05 (adjusted by Bonferroni correction)

<sup>1</sup>Buschke's Cued Recall Test: Total words correctly recalled during trials 1 to 3

<sup>2</sup>Buschke's Cued Recall Test: Delayed free recall

<sup>3</sup>Rey Auditory Verbal Learning Test: Total words correctly recalled during trials 1 to 5

<sup>4</sup>Rey Auditory Verbal Learning Test: Delayed free recall

<sup>5</sup>Controlled Oral Word Association Test

<sup>6</sup>Modified Mini-Mental State Exam: Pentagon Copy

**** p < .001
Table 4.7
Mean (and Standard Deviation) Neuropsychological Test Scores for Each Subgroup in the Replication Sample (n = 166)

<table>
<thead>
<tr>
<th>Test/Variable</th>
<th>Subgroup/Clusters</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>LM-IR(^1)</td>
<td></td>
<td>17.31(_{a,c})</td>
<td>(6.75)</td>
<td>25.33(_{b})</td>
<td>(8.40)</td>
<td>14.22(_{a,c})</td>
<td>(6.07)</td>
<td>13.71(_{c})</td>
<td>(5.68)</td>
<td>22.82(_{b})</td>
<td>(6.26)</td>
</tr>
<tr>
<td>LM-DR(^2)</td>
<td></td>
<td>12.78(_{a,c})</td>
<td>(8.46)</td>
<td>21.13(_{b})</td>
<td>(8.72)</td>
<td>9.11(_{c,d})</td>
<td>(7.86)</td>
<td>6.82(_{d})</td>
<td>(5.74)</td>
<td>16.24(_{a,b})</td>
<td>(7.67)</td>
</tr>
<tr>
<td>CVLT-Total(^3)</td>
<td></td>
<td>42.59(_{a,b})</td>
<td>(13.70)</td>
<td>48.87(_{a})</td>
<td>(15.51)</td>
<td>36.93(_{b})</td>
<td>(10.60)</td>
<td>34.59(_{b})</td>
<td>(10.45)</td>
<td>37.24(_{b})</td>
<td>(9.28)</td>
</tr>
<tr>
<td>CVLT-LDFR(^4)</td>
<td></td>
<td>8.88(_{a,b})</td>
<td>(4.20)</td>
<td>10.26(_{a})</td>
<td>(4.00)</td>
<td>6.33(_{b,c})</td>
<td>(3.74)</td>
<td>5.24(_{c})</td>
<td>(3.62)</td>
<td>6.50(_{b,c})</td>
<td>(3.51)</td>
</tr>
<tr>
<td>Digit Span</td>
<td></td>
<td>7.09</td>
<td>(2.56)</td>
<td>7.28</td>
<td>(2.74)</td>
<td>7.48</td>
<td>(2.24)</td>
<td>8.12</td>
<td>(1.77)</td>
<td>7.68</td>
<td>(2.06)</td>
</tr>
<tr>
<td>Similarities</td>
<td></td>
<td>16.78</td>
<td>(5.28)</td>
<td>16.62</td>
<td>(7.07)</td>
<td>17.59</td>
<td>(4.83)</td>
<td>19.24</td>
<td>(4.31)</td>
<td>19.53</td>
<td>(4.93)</td>
</tr>
<tr>
<td>Comprehension</td>
<td></td>
<td>17.19(_{a})</td>
<td>(6.78)</td>
<td>20.31(_{a,b,c})</td>
<td>(6.30)</td>
<td>19.44(_{a,b})</td>
<td>(4.92)</td>
<td>22.82(_{b,c})</td>
<td>(4.39)</td>
<td>23.77(_{c})</td>
<td>(4.15)</td>
</tr>
<tr>
<td>BNT(^5)</td>
<td></td>
<td>42.41(_{a})</td>
<td>(10.84)</td>
<td>46.26(_{a,b})</td>
<td>(11.08)</td>
<td>45.41(_{a,b})</td>
<td>(8.98)</td>
<td>50.29(_{b,c})</td>
<td>(5.97)</td>
<td>52.41(_{c})</td>
<td>(6.13)</td>
</tr>
<tr>
<td>COWAT(^6)</td>
<td></td>
<td>30.88(_{a})</td>
<td>(11.83)</td>
<td>27.67(_{a})</td>
<td>(12.44)</td>
<td>32.89(_{a,b})</td>
<td>(12.47)</td>
<td>40.41(_{b})</td>
<td>(11.77)</td>
<td>31.71(_{a})</td>
<td>(13.37)</td>
</tr>
<tr>
<td>Block Design</td>
<td></td>
<td>23.66(_{a,b})</td>
<td>(7.15)</td>
<td>19.36(_{a})</td>
<td>(9.57)</td>
<td>28.4(_{b})</td>
<td>(9.05)</td>
<td>22.09(_{a})</td>
<td>(7.17)</td>
<td>18.44(_{a})</td>
<td>(9.32)</td>
</tr>
<tr>
<td>TMT-Part A(^7)</td>
<td></td>
<td>48.88(_{a,b,c})</td>
<td>(19.71)</td>
<td>56.36(_{a,c})</td>
<td>(25.15)</td>
<td>40.37(_{b})</td>
<td>(15.89)</td>
<td>45.97(_{a,b})</td>
<td>(20.67)</td>
<td>61.53(_{c})</td>
<td>(27.10)</td>
</tr>
<tr>
<td>RCF(^8)-Copy</td>
<td></td>
<td>30.72(_{a})</td>
<td>(4.67)</td>
<td>29.19(_{a})</td>
<td>(7.16)</td>
<td>31.28(_{a})</td>
<td>(5.04)</td>
<td>31.27(_{a})</td>
<td>(4.02)</td>
<td>24.32(_{b})</td>
<td>(7.83)</td>
</tr>
</tbody>
</table>

Note. Means within each row that do not share the same subscript differ at p < .05 (adjusted by Bonferroni correction)

1 WMS-R Logical Memory: Immediate recall  
2 WMS-R Logical Memory: Delayed recall  
3 California Verbal Learning Test: Total words correctly recall during trials 1 to 5  
4 California Verbal Learning Test: Long delay free recall  
5 Boston Naming Test  
6 Controlled Oral Word Association Test  
7 Trail Making Test  
8 Rey Complex Figure  
* Higher scores indicate worse performance  
** p < .01  **** p < .001
in Table 4.2. If needed, clinicians can choose to administer only those tests that have the highest factor loadings for each of the three factors, which are also the tests that show the largest differences among the five subgroups.

Each of the significant ANOVAs reported in Tables 4.6 and 4.7 were followed up with multiple comparisons. In Tables 4.6 and 4.7, the results of the follow-up comparisons are indicated with subscripts: Means within each row that do not share the same subscript are significantly different from one another at \( p < .05 \) (\( p \)-values adjusted using Bonferroni correction). Rather than describing all of the significant subgroup differences, a summary of the cognitive strengths and weaknesses for each subgroup relative to the other four subgroups is provided. In the Base sample, the Verbal Dysfunction subgroup was characterized by low performance across each of the Verbal Ability tests with the most pronounced deficits on the Comprehension and Similarities subtests, and on the Token Test. The performance of this subgroup seemed intact on each of the Memory and Visuospatial Ability tests. Compared to other the subgroups, this subgroup performed best on the Block Design subtest. In the Replication sample, the Verbal Dysfunction subgroup showed the most pronounced impairment on the Comprehension subtest, the BNT, and the LM-IR subtest. This subgroup tended to perform best on measures of delayed recall. Overall, the tests that best capture the impairment of the Verbal Dysfunction subgroup were those of judgement (Comprehension subtest) and language abilities (Token Test and BNT).

The Verbal/Visuospatial Dysfunction subgroup performed well on Memory Ability tests in both samples. In the Base sample, the impairment was more pronounced on the three Visuospatial Ability tests (Block Design, Digit Symbol, and 3-MS Pentagon Copy) than on the Verbal Ability tests. In the Replication sample, however, this subgroup demonstrated mixed deficits on Verbal Ability (BNT and COWAT) and Visuospatial Ability tests (Block Design and TMT-Part A). Overall, tests of visuoconstruction (Block Design) and attention (Digit Symbol
and TMT-Part A) appear to best capture the impairment of the Verbal/Visuospatial Dysfunction subgroup.

The Memory/Verbal Dysfunction subgroup was more impaired on the Memory Ability tests than on the Verbal Ability tests. In the Base sample, however, performance of this subgroup was worse on the RAVLT than on the CRT; test scores were actually significantly higher in this subgroup on the CRT than the Memory Dysfunction subgroup. In the Replication sample, performance of this subgroup was worse on the LM subtest than on the CVLT. Of the Verbal Ability tests, however, performance was impaired on the Comprehension subtest in both samples. Within each sample, performance on each of the three tests of Visuospatial Ability was largely intact in this subgroup, with test scores on the Block Design subtest being the highest. Overall, the tests that best captured the impairment of this subgroup were the RAVLT or LM and the Comprehension subtest.

The memory impairment exhibited by the Memory Dysfunction subgroup was most pronounced on the CRT (Base sample) and LM subtest (Replication sample). Better performance was observed in this subgroup on the Verbal Ability tests than the Visuospatial Ability tests in both samples. Of the Verbal Ability tests, the highest test scores demonstrated by this subgroup were on the Similarities subtest and the Token Test in the Base sample and COWAT and Comprehension subtest for the Replication sample.

The Visuospatial Dysfunction subgroup was characterized by deficits on most of the Visuospatial Ability tests with the most impairment being displayed on the Copy measures (3-MS Pentagon Copy and RCF) in both samples. Performance was relatively intact across the Verbal Ability and Memory Ability tests in each sample. In the Base sample, the highest test scores were observed on the Comprehension subtest; in the Replication sample, the highest test scores were observed on the BNT and LM-IR subtest.
Summary of additional cluster interpretation. In each of the two samples, the five subgroups identified did not differ with respect to potentially confounding demographic and clinical variables such as age, years of education, sex, and overall cognitive status. As expected, the pattern of subgroup differences on the original 12 neuropsychological variables resembled the performance patterns of these subgroups on the factor scores. Of particular interest was the fact that although the subgroups were originally formed using ipsative factor z-scores, which standardize the test scores of each individual to their own level of performance, the subgroups differed significantly from one another in the predicted direction on many of the original neuropsychological variables. These results indicate that many of the impairments observed for each subgroup were (a) the most pronounced deficits for each individual in that subgroup, and (b) these impairments were statistically significant when compared to the other subgroups.

Predictive Validity of the Five-Cluster Solution

A total of 225 participants in the Base sample received clinical assessments at the 5-year follow-up. In the Replication sample, a total of 94 participants were reassessed at the 2-year follow-up. To assess the overall predictive validity of the five-cluster solution in each sample, logistic regression analyses were performed to determine whether subgroup membership at baseline was a significant predictor of conversion to dementia. Follow-up analyses were conducted to determine the relative risk of converting to dementia for each subgroup compared to the remaining sample. In other investigations, demographic variables such as age, years of education, and sex are often covaried when examining predictors of dementia because these variables have been associated with the development of dementia (Bowwirrat, Friedland, Farrer, Baldwin, & Korczyn, 2002; Lindsay et al., 2002; Ravaglia et al., 2002). We examined the predictive ability of subgroup membership with and without covarying age, years of education, and sex, and found that predictive ability was improved after covarying these demographic variables. Even though the subgroups did not differ significantly with respect to these variables,
the differences among the subgroups may have been enough to have an influence on predicting dementia. Age, years of education, and sex were therefore covaried in both the Base and Replication samples.

Subgroup membership was a significant predictor of dementia in the Base sample, Wald z-ratio (4) = 25.78, \( p < .001 \), but its predictive ability failed to achieve significance in the Replication sample, Wald z-ratio (4) = 6.35, \( p = .174 \). We decided to conduct follow-up analyses for both samples given that the nonsignificant finding in the Replication sample may have been the result of the smaller size of this sample at Time 2 (\( n = 94 \)) and because we wanted to compare the relative risk associated with each subgroup between samples. The diagnostic status at follow-up for the Base and Replication samples is provided in Table 4.8. Also reported are the odds ratios, and 95% confidence intervals, that correspond to the relative risk of each subgroup for converting to dementia compared to the remaining sample. Odds ratios with 95% confidence intervals that do not cross 1.00 are considered significant.

As shown in Table 4.8, the overall rate of conversion to dementia was 40.44% in the Base sample and 29.79% in the Replication sample. The percentage of participants who improved, reverting to NCI, was 16.00% in both samples. The diagnostic status remained stable in 43.56% and 54.26% of participants in the Base and Replication samples respectively. The highest conversion rate to dementia in the Base (65.11%) and Replication (47.06%) samples was observed in the Memory Dysfunction subgroup. The relative risk for dementia was significantly elevated in this subgroup compared to the remaining subgroups in each sample: The odds of developing dementia were approximately 3.5 times greater for subjects in the Memory Dysfunction subgroup. The rate of conversion to dementia was significantly elevated in the Memory/Verbal Dysfunction subgroup (54.55%; OR = 2.27) in the Base sample but not in the Replication sample (23.08%; OR = 0.90).
Table 4.8  
*Diagnostic Status at Follow-up in Base (n = 225) and Replication (n = 94) Samples*

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Total</th>
<th>NCI</th>
<th>CIND</th>
<th>Dementia</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base Sample (5-year Follow-up)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Dysfunction</td>
<td>54</td>
<td>14 (25.93%)</td>
<td>27 (50.00%)</td>
<td>13 (24.07%)</td>
<td>0.30</td>
<td>0.15 to 0.64</td>
</tr>
<tr>
<td>Visuospatial/Verbal Dysfunction</td>
<td>43</td>
<td>8 (18.60%)</td>
<td>18 (41.86%)</td>
<td>17 (39.54%)</td>
<td>1.03</td>
<td>0.51 to 2.09</td>
</tr>
<tr>
<td>Memory/Verbal Dysfunction</td>
<td>44</td>
<td>4 (9.09%)</td>
<td>16 (36.36%)</td>
<td>24 (54.55%)</td>
<td>2.27</td>
<td>1.11 to 4.64</td>
</tr>
<tr>
<td>Memory Dysfunction</td>
<td>43</td>
<td>3 (6.98%)</td>
<td>12 (27.91%)</td>
<td>28 (65.11%)</td>
<td>3.82</td>
<td>1.81 to 8.03</td>
</tr>
<tr>
<td>Visuospatial Dysfunction</td>
<td>41</td>
<td>7 (17.07%)</td>
<td>25 (60.98%)</td>
<td>9 (21.95%)</td>
<td>0.35</td>
<td>0.15 to 0.81</td>
</tr>
<tr>
<td>Total</td>
<td>225</td>
<td>36 (16.00%)</td>
<td>98 (43.56%)</td>
<td>91 (40.44%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Replication Sample (2-year Follow-up)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Dysfunction</td>
<td>20</td>
<td>6 (30.00%)</td>
<td>7 (35.00%)</td>
<td>7 (35.00%)</td>
<td>0.87</td>
<td>0.26 to 2.90</td>
</tr>
<tr>
<td>Verbal/Visuospatial Dysfunction</td>
<td>23</td>
<td>5 (21.74%)</td>
<td>15 (65.23%)</td>
<td>3 (13.04%)</td>
<td>0.24</td>
<td>0.56 to 0.99</td>
</tr>
<tr>
<td>Memory/Verbal Dysfunction</td>
<td>13</td>
<td>2 (15.38%)</td>
<td>8 (61.54%)</td>
<td>3 (23.08%)</td>
<td>0.90</td>
<td>0.21 to 3.95</td>
</tr>
<tr>
<td>Memory Dysfunction</td>
<td>17</td>
<td>0 (0.00%)</td>
<td>9 (52.94%)</td>
<td>8 (47.06%)</td>
<td>3.48</td>
<td>1.02 to 11.84</td>
</tr>
<tr>
<td>Visuospatial Dysfunction</td>
<td>21</td>
<td>2 (9.52%)</td>
<td>12 (57.14%)</td>
<td>7 (33.33%)</td>
<td>1.33</td>
<td>0.43 to 4.08</td>
</tr>
<tr>
<td>Total</td>
<td>94</td>
<td>15 (16.00%)</td>
<td>51 (54.26%)</td>
<td>28 (29.79%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Odds ratios are reported after covarying age, years of education, and sex*
In the Base sample, the Visuospatial Dysfunction and Verbal Dysfunction subgroups demonstrated the lowest conversion rates to dementia at 21.95% and 24.07% respectively. Participants in each of these two subgroups were significantly less likely to develop dementia than the remaining sample: The odds of conversion for the Visuospatial Dysfunction and Verbal Dysfunction subgroups were 0.35 and 0.30 respectively. The lowest rate of converting to dementia in the Replication sample was observed in the Verbal/Visuospatial Dysfunction subgroup (13.04%); individuals from this subgroup were significantly less likely to progress to dementia (OR = 0.24). In terms of reversion to NCI, the Verbal Dysfunction subgroup demonstrated the highest rates in the Base (25.93%) and Replication (30.00%) samples.

Discussion

There were two main findings in the present investigation. First, we identified subgroups of CIND individuals with distinct neuropsychological profiles in two independent samples and with a variety of different clustering methods. The subgroups were labelled according to the cognitive domain(s) that were most impaired, and included the Verbal Dysfunction subgroup, the Verbal/Visuospatial Dysfunction subgroup, the Memory/Verbal Dysfunction subgroup, the Memory Dysfunction subgroup, and the Visuospatial Dysfunction subgroup. The second main finding was that subgroup membership at baseline was associated with diagnostic outcome. Conversion to dementia was highest in the Memory Dysfunction subgroup (in the Base and Replication samples) and Memory/Verbal Dysfunction subgroup (Base sample only). Reversion to NCI was highest in the Verbal Dysfunction subgroups in the Base and Replication samples.

The methodology and results of the present investigation differ in several important ways from previous studies that have examined neuropsychological subgroups of cognitively healthy adults (Fisher et al., 1999; Larrabee et al., 1986; Malec et al., 1996; Mitrushina et al., 1995; Valdois et al., 1990) and individuals with subclinical or mild cognitive deficits (Reischies & Hellweg, 2000; Ritchie et al., 1996; Ylikoski et al., 1999). Most importantly, this investigation
was the first to use cluster methods to identify subgroups based only on profile shape in individuals at-risk for dementia. The predictive ability of the subgroups identified here, therefore, could not be explained by differences in overall profile elevation. The subgroup with the highest rate of conversion to dementia, the Memory Dysfunction subgroup, was characterized by a distinct pattern of impaired and intact test scores rather than an overall low performance profile. It is therefore unlikely that we were simply capturing individuals who were on the cusp of converting to dementia.

Unlike many of the previous studies in the cognitive aging literature, the subgroups identified in this investigation did not differ with respect to variables such as age, years of education, sex, and general cognitive status. This finding is important because these variables have been identified as potential risk factors for dementia (Bowwirrat et al., 2002; Lindsay et al., 2002; Ravaglia et al., 2002). Subsequently, the elevated risk of conversion to dementia in the Memory Dysfunction subgroup could not be attributed to the fact that this subgroup comprised participants who were simply older or less educated.

The present investigation was also the first in this area of research to use two independent samples of participants. We were therefore able to show that the results produced in the Base sample generalized to the Replication sample, which increases our confidence that the subgroups identified in this investigation are reliable. The use of population- and clinic-based samples in this investigation was important because individuals from these two sample types have been shown to differ with respect to overall level of cognitive impairment and rate of conversion to dementia (Bischkopf et al., 2002; Petersen, 2003; Tuokko & Frerichs, 2000).

The results of the present investigation have important practical implications. We have identified subgroups of CIND individuals that differ with respect to diagnostic outcome. This information is useful in several ways. First, there are a number of the pharmaceutical treatments currently available for Alzheimer Disease, which may temporarily stabilize or even enhance
cognitive functioning and quality of life for periods of 6 to 12 months (Knopman, 2003). The results of this investigation suggest that individuals from the Memory Dysfunction subgroup, and to some extent the Memory/Verbal Dysfunction subgroup, may be ideal candidates for the early implementation of these treatment interventions. Earlier implementation of treatment interventions might help reduce the personal and societal costs associated with dementia.

The elevated risk of conversion to dementia observed in the Memory Dysfunction subgroup is consistent with previous research that has targeted the amnestic form of mild cognitive impairment (MCI) described by Petersen and colleagues (Petersen et al., 2001; Petersen et al., 1995). The fact that our empirically based approach identified a Memory Dysfunction subgroup from within a larger, more heterogeneous, sample of CIND individuals provides validation for the amnestic form of MCI, which is a clinically defined condition. Our results, therefore, support the focus on amnestic MCI in previous studies aimed at identifying biomarkers of Alzheimer disease (Graff-Radford, 2003) and in several clinical trials currently underway (Knopman, 2003). Inclusion of individuals from the other subgroups identified here, those with lower risks for converting to dementia, might decrease the power to detect significant effects due to treatment.

The identification of lower-risk subgroups in the present investigation, those involving verbal and/or visuospatial impairment without memory deficits, also has important implications. Some of the individuals from these subgroups may represent what others have referred to as the “worried well” (Petersen, 2000b). Rather than bringing these individuals back for annual or even semi-annual clinical assessments, clinicians may choose to reassess these individuals after longer time intervals, thus freeing up limited time and resources for higher-risk individuals. It is important to point out, however, that the concerns of individuals in these lower-risk subgroups should not be simply dismissed as the rates of converting to dementia in these subgroups was still higher than normal (Hogan & Ebly, 1999).
There are two limitations on the generalizability of our findings. First, the neuropsychological test batteries in the CSHA and the ACCORD study were designed to tap the DSM-III-R criteria for dementia. Given the importance of memory impairment in the DSM-III-R criteria, the test batteries were well designed to probe memory functioning, but they were not as comprehensive in their coverage of other cognitive domains (e.g., executive functioning). We realize that the results of the principal component and subsequent cluster analyses depended entirely on the variables that were included in these analyses. It is possible, therefore, that by increasing the breadth of the test battery, we would find additional factors and perhaps even additional subgroups (e.g., a subgroup with impaired executive functioning). We intend to address this issue in the future by using a more comprehensive test battery that will allow us to explore further the heterogeneity within cognitive domains (e.g., receptive versus expressive language abilities) as well as between cognitive domains (e.g., memory, language abilities, and visuospatial abilities).

The second major limitation of the present investigation pertains to the samples sizes at follow-up: 225 and 94 in the Base and Replication samples respectively. This difference may explain the inconsistencies in conversion rates between the two samples for some of the subgroups and why cluster membership was not a significant predictor of dementia in the Replication sample. In addition, the length of follow-up was only 2 years in the Replication sample and only 5 years in the Base sample. To gain a better understanding of the progression of these neuropsychological subgroups, it will be necessary to monitor their diagnostic status over longer periods of time (e.g., 10 to 15 years).

In conclusion, the results of the present investigation suggest that there are at least five distinct and reliable neuropsychological subgroups of CIND individuals. These five subgroups include the Verbal Dysfunction subgroup, the Verbal/Visuospatial Dysfunction subgroup, the Memory/Verbal Dysfunction subgroup, the Memory Dysfunction subgroup, and the Visuospatial
Dysfunction subgroup. The highest rate of conversion to dementia was observed in the Memory Dysfunction subgroup and in the Memory/Verbal Dysfunction subgroup to a lesser extent. The subgroup most likely to show improvement in cognitive status was the Verbal Dysfunction subgroup. The cognitive heterogeneity of this population must be taken into account in future research focusing on the early identification of dementia.
Acknowledgements

We would like to express our sincerest gratitude to the participants and their families for their commitment to both the CSHA and ACCORD studies. We would also like to acknowledge the hard work performed by all of the individuals who were involved in each of these two studies. The CSHA was supported through the National Health Research and Development program grant # 6606-3954-MC[S]. The ACCORD study was supported through the MRC PMAC program grant # PA14197. The first author (K.R.P.) received support from a Doctoral Training Award jointly funded by the Alzheimer Society of Canada and the Canadian Institutes of Health Research.

Correspondence concerning this article should be addressed to Kevin Peters at the Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, B.C., Canada, V6T 1Z4 (email: kpeters@interchange.ubc.ca).
CHAPTER 5

General Discussion

The overall goal of this dissertation was to determine whether the cognitive heterogeneity of CIND individuals could account for differences in diagnostic outcomes. I used two approaches to achieve this goal. The first approach was to characterize the neuropsychological test performance of a clinic-based sample of CIND individuals and to provide data that would facilitate their identification. The second approach was to identify reliable subgroups of CIND individuals with distinct neuropsychological profiles, and to determine whether these subgroups differ in their risk for conversion to dementia. The main findings of these two approaches were discussed in the context of previous research in Chapters 3 and 4, and thus, only a summary of these findings and contributions is presented here. The theoretical and practical implications of this research are then discussed followed by limitations and avenues for future research.

Summary of Findings and Contributions of This Research

The first approach toward achieving the overall goal of this dissertation was to explore the cognitive heterogeneity of the CIND condition by characterizing the neuropsychological test performance of a clinic-referred sample of CIND individuals. This approach also involved generating the first set of clinical comparison data for CIND individuals. As with normative data for cognitively healthy persons, these comparison data were generated separately for various age-, education-, and sex-defined groups. These data can be used to enhance the identification of CIND individuals in clinical settings. The rationale behind this approach was that clinicians must be able to identify CIND individuals in a reliable and valid manner before they can determine which individuals will and will not progress to dementia. Participants classified as CIND were found to differ from NCI individuals on a number of demographic, clinical, and neuropsychological variables. The best discrimination between NCI and CIND individuals was provided by measures of learning and memory, visuoconstruction abilities, and cognitive
flexibility. Additional analyses revealed that CIND individuals were characterized by considerable intertest scatter (the difference between an individual's highest and lowest test scores) and approximately half of the sample was either not impaired on any variable or impaired on only one variable. Impairment was defined as performance that was one standard deviation or more below the mean of an individual's age-defined group. Overall, the data reported in Chapter 3 support the clinical impression that CIND is a cognitively heterogeneous population.

The results of the investigation in Chapter 3 were followed up in Chapter 4 by using the subgroup approach in two independent samples of CIND individuals. There were two main findings from this second investigation. First, five subgroups of CIND individuals were reliably identified using a variety of cluster methods and in both samples. A distinct neuropsychological profile characterized each of these five subgroups, which included the Verbal Dysfunction subgroup, the Verbal/Visuospatial Dysfunction subgroup, the Memory/Verbal Dysfunction subgroup, the Memory Dysfunction subgroup, and the Visuospatial Dysfunction subgroup. The second major finding of this investigation was that subgroup membership at baseline was related to diagnostic outcome over periods of 2 and 5 years. The Memory Dysfunction subgroup was characterized by the highest rate of conversion to dementia in the Base (65.11%) and Replication (47.06%) samples. The Verbal Dysfunction subgroup had the highest rate of reversion to NCI in the Base (25.93%) and Replication (30.00%) samples.

This dissertation was the first to use cluster methods that form subgroups based on profile shape in individuals at risk for dementia. Unlike previous investigations in this field (Larrabee et al., 1986; Malec et al., 1996; Mitrushina et al., 1995; Reischies & Hellweg, 2000; Ritchie et al., 1996; Valdois et al., 1990; Ylikoski et al., 1999), the subgroups identified here did not differ with respect to overall profile elevation. It is therefore unlikely that the subgroup with the highest risk for dementia, the Memory Dysfunction subgroup, was composed of individuals who were already on the cusp of converting to dementia at baseline.
Another strength of this dissertation is that it is the first to use two independent samples of participants. By using two samples, I was able to document the generalizability of the results. The fact that these two samples were taken from population- and clinic-based studies of CIND individuals further increases the generalizability of the findings.

Unlike previous studies, the investigation reported in Chapter 4 involved performing principal component analyses on age-, education-, and sex-residualized neuropsychological test scores. There were two advantages of this data preparation procedure. First, removing the variance due to age, education, and sex from each of the neuropsychological variables simplified interpretation of the resulting clusters: The differences observed among subgroups could not be accounted for by these confounding variables. The distinct subgroup profiles observed were more likely due to different patterns of underlying cognitive strengths and weaknesses. The second advantage of the data preparation procedure was that the principal components analyses reduced the number of clustering variables from 12 neuropsychological measures to 3 orthogonal factors. Accordingly, this procedure increased the reliability and the construct validity of the clustering variables and avoided the problem of multicollinearity in the cluster analysis.

**Theoretical and Practical Implications of This Research**

From a theoretical point of view, the results of this dissertation suggest that it might be more useful to conceptualize CIND as a syndrome that comprises a variety of conditions rather than a single population of individuals at-risk for dementia. There are several features of the subgroups reported in Chapter 4 that support this hypothesis. The five subgroups were characterized by qualitatively distinct neuropsychological profiles and each of the subgroup profiles was characterized by a combination of high- and low-test scores. Subgroups differed in their risk for converting to dementia, but they did not differ with respect to potential risk factors for dementia such as age, education, sex, or performance on a general cognitive screening
instrument. These results suggest that the subgroups do not simply reflect different stages of a single developmental pathway from CIND to dementia.

A similar proposal has been made recently by Petersen and colleagues (Petersen, 2003; Petersen et al., 2001), who have described three different types of mild cognitive impairment (MCI). *Amnestic MCI* refers to individuals who present with isolated memory impairment and seems to be similar to the Memory Dysfunction subgroup identified here. Compared to amnestic MCI, the other two types of MCI described by Petersen and colleagues have received very little attention and are based more on hypothetical possibilities. *Multiple domains slightly impaired MCI* is meant to describe individuals with mild impairment in more than one cognitive domain and would be consistent with the Memory/Verbal Dysfunction and Verbal/Visuospatial Dysfunction subgroups. *Single nonmemory domain MCI* is reserved for individuals who present with impairment in a single cognitive domain other than memory and would correspond to the Verbal Dysfunction and Visuospatial Dysfunction subgroups. The advantage of the subgroups reported in this dissertation is that they were reliably identified in two independent samples using empirically based methods. The nature of these subgroup profiles were also characterized in more detail, and would therefore be easier to identify, than the types of MCI proposed by Petersen and colleagues.

Recognition that CIND is a syndrome rather than a single population of individuals at-risk for dementia will have implications for identifying CIND individuals in clinical settings. It is highly unlikely that a single cognitive test could ever be used to diagnose and classify CIND. The results reported in Chapter 3 indicated that even a combination of several diverse and standardized neuropsychological tests couldn’t discriminate between NCI and CIND individuals with 100% accuracy. In situations where there are sufficient clinical resources, it would be ideal to use the neuropsychological test data provided for each of the five subgroups identified in Chapter 4 to diagnose and classify CIND individuals. However, in situations where clinicians are
able to administer only a short test battery, the clinical comparison data provided in Chapter 3 will assist clinicians in making a general diagnosis of CIND.

The results of this dissertation also suggest that it would be very difficult to identify a single cognitive test that consistently predicts conversion to dementia in CIND individuals. It might be the case, however, that the subgroup and best test approaches actually complement one another, such that different subsets of predictors emerge for each subgroup. It is possible, for example, that the tests that predict dementia in the Visuospatial Dysfunction subgroup are different from those that predict this outcome in the Memory Dysfunction subgroup. Future research with larger subgroup sample sizes would be able to address this issue.

The cognitive heterogeneity of CIND individuals will also have implications for researchers investigating various biomarkers (e.g., apolipoprotein E genotype, cerebral spinal fluid levels of tau protein) for Alzheimer disease and other types of dementia (Graff-Radford, 2003). Putative biomarkers might be present, for example, in a higher-risk subgroup such as the Memory Dysfunction subgroup but not in the other subgroups that are characterized by a lower-risk for dementia. The presence of any significant subgroup effect would be washed out if one were to examine the average data from a large sample comprising CIND individuals from different subgroups.

In 1997, donepezil (a cholinesterase inhibitor) became the first drug approved for use in Canada for Alzheimer disease. Since that time, a number of other pharmaceutical treatments are currently available or being developed for Alzheimer disease and other types of dementias. Some of the more common agents include other cholinesterase inhibitors, nonsteroidal anti-inflammatory drugs, hormones (e.g., estradiol), antioxidants (e.g., vitamin E), NMDA receptor antagonists, nerve growth factors, herbs (e.g., ginkgo biloba), anti-amyloid strategies, lipid-lowering agents, and antihypertensives (Doraiswamy, 2002). Although these treatments are not cures, some have been shown to be effective in temporarily stabilizing or even enhancing
cognitive functioning and quality of life in dementia patients over periods of 6 to 12 months (for reviews, see Doraiswamy, 2002; Knopman, 2003). Pharmaceutical companies are continuing their efforts to develop more effective treatments and even possible cures (Knopman, 2003). The results of this dissertation indicate that individuals from the Memory Dysfunction and Memory/Verbal Dysfunction subgroups would be ideal candidates for future clinical trials research. Inclusion of individuals from lower-risk subgroups into clinical trials might decrease the power to detect significant effects due to treatment.

The results of this dissertation have a number of practical implications for how CIND individuals are treated and managed in clinical settings. There has been considerable discussion recently over implementing treatments for Alzheimer disease and other types of dementia in individuals believed to be at risk for dementia (Knopman, 2003). The dilemma faced by physicians, however, is which individuals should and should not receive such treatments. Individuals from the Memory Dysfunction and Memory/Verbal Dysfunction subgroups are obvious candidates for the early implementation of treatment interventions. Implementing treatment early in the course of the disease might help to preserve cognitive abilities and overall functioning for an extended period of time. On the other hand, currently available pharmaceutical treatments are often expensive and have unpleasant side effects in some individuals. Accordingly, physicians may choose to forego or delay prescribing these treatments to individuals from the lower-risk subgroups identified here, such as the Verbal Dysfunction and Visuospatial Dysfunction subgroups.

Dementing conditions such as Alzheimer disease are progressive and associated with many costs to individuals and to their families (CSHA Working Group, 1994b). The results of this dissertation suggest that it might be useful to provide psychological counselling to individuals from the Memory Dysfunction and Memory/Verbal Dysfunction subgroups and to their families at an early stage. Implementing such interventions might help to prepare
individuals and their families for the many difficult aspects associated with a dementing condition, and assist them to plan ahead for making decisions regarding long-term care if necessary.

This research has provided useful information that will allow for depleted healthcare resources to be channelled more efficiently to those individuals who are at higher-risk for converting to dementia. Rather than reassessing all CIND individuals every 6 to 12 months, clinicians may choose to re-prioritize the scheduling of patients so that higher-risk individuals can be reassessed more frequently than lower-risk individuals. It must be kept in mind, however, that the complaints of lower-risk individuals should not be simply dismissed as the rates of conversion to dementia in these subgroups are still higher than normal (Hogan & Ebly, 1999).

Limitations of This Research

This research has two important limitations. The first has to do with a bias in selecting the NCI and CIND participants. Not all participants in the CSHA and ACCORD studies received neuropsychological testing, and of those who agreed to have testing done, there was an additional selection bias in terms of participants who completed all of the testing. In Chapter 3, it was pointed out that the ACCORD study participants who underwent neuropsychological testing were younger, less educated, and scored lower on a general cognitive screening instrument than participants who did not receive testing. These selection biases do limit the generalizability of the findings of this investigation to some extent. It must be kept in mind, however, that problems with missing data are common in multicentre studies.

To overcome some of the selection biases present in the ACCORD study, an attempt was made in Chapter 3 to make use of all available data by imputing group means for missing data points. Imputation of missing data was not performed in Chapter 4 because the percentage of CIND participants with incomplete data was considerably higher in the CSHA sample (53.54%)
than in the ACCORD sample (28.70%). This decision was based on the desire to treat both samples equally even if it meant that there would be more selection biases present.

An additional selection bias was present for the analyses of participants at follow-up 5 years later in the CSHA and 2 years later in the ACCORD study. Attrition is a common problem in longitudinal studies and the resulting selection biases also limit the generalizability of findings. In this investigation, for example, a considerable portion of CIND participants in the CSHA (40.26%) died during the 5-year period between the baseline and first follow-up visits. Although not reported in Chapter 4, there were no significant differences among the five subgroups in terms of the percentage participants who were deceased at Time 2, $\chi^2(4) = 2.33, p = .675$. Thus, it is likely that the results of this investigation would have been similar had the follow-up sample comprised all of the original participants. One cannot make this claim for certain however.

The second limitation of this research arises from the neuropsychological tests used in the CSHA and ACCORD studies. These test batteries were intended primarily for diagnostic decision making, that is, they were designed to tap the DSM-III-R diagnostic criteria for dementia (American Psychiatric Association, 1987) which requires evidence of a deficit in memory and in at least one other domain of cognition. In view of this purpose, both batteries focused on memory but did not include specific instruments for indexing other important aspects of cognition (e.g., attention, executive functions). Consequently, because the findings reported in this dissertation are a direct product of the dataset obtained by the test batteries, it is conceivable that additional or different factors and subgroups might have emerged if a different or more comprehensive test battery had been used.

One factor that argues against the possibility that different subgroups would have emerged if participants had been assessed by means of a different test battery stems from the fact that no neuropsychological test is "process or skill pure". Performance on a memory test, for
example, is not determined exclusively by an individual's memory skills; it also reflects their understanding of the instructions given for the test (i.e., their language skills), their ability to focus and maintain attention, their level of motivation, etc. Consequently, the test batteries that were used were likely to tap many aspects of cognition other than those identified by specific test labels. Nevertheless, future research should attempt to validate this assumption by using a battery that focuses more equally on all aspects of cognition.

Future research should also look more closely inside each cognitive domain to determine whether all or only some aspects (ability, skill, function, or process) of it have predictive power. Consider for example that although the present research emphasized memory, it targeted only one of its components -- episodic verbal memory. A more comprehensive test battery might also assess components such as spatial memory or prospective memory. By a closer examination of these and other aspects of each cognitive domain, it may be possible to identify diagnostically distinct subgroups within the patient-clusters that emerged from the research reported here.

The two test batteries used in this investigation were similar but not identical. An effort was made to match them as closely as possible but some of the pairings were less ideal than others. For example, at a conceptual level, the RAVLT and CVLT are more similar to one another than the WAIS-R Digit Symbol Substitution subtest and Part A from the Trail Making Test. Having said this, the results described in Chapter 4 indicated that the matching procedure was relatively successful: The patterns of factor loadings were consistent between the two samples, and the performance of the five subgroups on the matched variables was also similar between the two samples.

Avenues for Future Research

A number of interesting questions arose from this investigation. The following is a discussion of several avenues for future research. The first avenue for future research concerns the relationship between memory impairment and the diagnosis of dementia. The Memory
Dysfunction subgroup identified in this investigation had the highest rate of conversion to dementia. This subgroup is similar to the amnestic form of MCI that has been described by Petersen and colleagues (Petersen et al., 2001; Petersen et al., 1995), which is also characterized by an elevated risk for dementia. The finding that low memory performance is associated with conversion to dementia is useful but not very surprising given that a diagnosis dementia requires memory impairment (American Psychiatric Association, 2000). The circularity of impaired performance on memory tests being predictive for dementia has been pointed out by Tuokko and Frerichs (2000). There are two points, however, that deserve mention here. First, the finding that the Memory Dysfunction subgroup exhibited the highest rate of conversion to dementia was reassuring and served to validate the cluster solutions obtained in this investigation. In fact, it would have been more surprising if the Memory Dysfunction subgroup did not have the highest conversion rate. The problem of circularity may have to do more with the diagnostic criteria for dementia themselves than for the results obtained here.

The second point regarding the role of memory impairment in predicting dementia has to do with the nature of the memory deficit. Although impaired memory performance is associated with progression to dementia, the processes that mediate these deficits in individuals at-risk for dementia are still unknown. As pointed out by Almkvist et al. (1998), the memory deficits observed in at-risk individuals could be due to impaired encoding, storage, retrieval, or some combination of the these three processes. This is an area of research that needs to be explored if we are to understand the developmental pathways of individuals at-risk for dementia.

An interesting approach to investigate the nature of the memory deficits in individuals from the Memory Dysfunction and Memory/Verbal Dysfunction subgroups would be to examine the hemispheric encoding-retrieval asymmetry (HERA) model that has been put forth by Tulving and colleagues (Tulving, Kapur, Craik, Moscovitch, & Houle, 1994) in these subgroups. According to the HERA model, encoding is associated with increased activation in the left
prefrontal cortex and retrieval is associated with increased activation in the right prefrontal cortex. Knowledge of the processes underlying the memory deficits observed in higher-risk CIND individuals would provide additional predictive power and would yield useful information for rehabilitating memory deficits in this population. Different strategies are likely to be optimal for rehabilitating impairments in encoding, storage, and retrieval (Sohlberg & Mateer, 2001).

The generalized slowing hypothesis of cognitive aging is that normal age-related changes in cognition are mediated by reductions in processing speed and capacity (Birren, 1974; Graf & Uttl, 1995; Salthouse, 1985). This hypothesis has received considerable support in investigations involving cognitively healthy adults, and a number of biologically plausible explanations for reduced processing speed have been suggested (for reviews, see Salthouse, 1985; Salthouse, 1996a). The generalized slowing hypothesis has also been examined in CIND individuals (Peters, 1999). The results of this investigation indicated that the generalized slowing hypothesis could not fully explain the age-related cognitive declines in CIND individuals: The attenuation in the ability of age to predict memory test performance after partialling out measures of processing speed was considerably less than what has been reported in the literature on normal cognitive aging. It might be the case, however, that processing speed explains age-related cognitive deficits in some of the CIND subgroups but not in others. Examination of the generalized slowing hypothesis in each of the five subgroups might provide insight into whether this hypothesis holds across different cognitive domains and it might also illuminate some of the mechanisms of cognitive decline in CIND individuals.

The present investigation examined the cognitive heterogeneity of CIND and found evidence of five subgroups that were characterized by distinct neuropsychological test profiles. An interesting approach to adopt in future research would be to investigate whether the five subgroups differ with respect to various non-cognitive factors. Changes in behaviour and personality are known to occur in different types of dementia (Cummings, 2003). The DSM-IV-
TR (American Psychiatric Association, 2000), for example, allows for two clinically defined subtypes of Alzheimer disease: one with behavioural disturbances and one without. If the CIND subgroups were to differ with respect to changes in behaviour and personality, this information might add to the predictive ability of subgroup membership.

Another avenue for future research would be to determine whether differences among the five CIND subgroups could be validated using biological measures such as structural and functional brain imaging. Different patterns of brain atrophy and neuropathological processes might account for some of the differences in conversion rates observed among the five subgroups. For example, individuals from the Memory Dysfunction subgroup appear to be quite similar to the amnestic form of MCI and other similar clinical populations, which have been shown to have smaller than normal hippocampal and entorhinal cortex volumes (Convit et al., 1997; Wolf et al., 2001; Xu et al., 2000). The area of the medial temporal lobe comprising the hippocampus and adjacent entorhinal, perirhinal, and parahippocampal cortices have been shown to be involved in normal memory functioning (Broadbent, Clarke, Zola, & Squire, 2002; Duva, Kornecook, & Pinel, 1999). The entorhinal cortex is also believed to be the initial site of neuropathology in Alzheimer disease, with damage then spreading out into other limbic and neocortical areas (Braak & Braak, 1991, 1995). It is possible that individuals from the Memory Dysfunction and Memory/Verbal Dysfunction subgroups are also characterized by atrophy of the medial temporal lobe region, which could account for their higher conversion rates to dementia.

In addition to examining structural measures of brain atrophy, it would also be useful to investigate functional measures of brain activity in individuals from the five CIND subgroups. Using positron emission tomography, Martin and colleagues (Martin, Brouwers, Lalonde, Cox, & Fedio, 1987) have demonstrated that neuropsychological subgroups of patients with mild Alzheimer disease were characterized by differential patterns of cerebral glucose metabolism. These investigators reported that patients with word finding difficulties had significantly lower
than normal cerebral glucose metabolism in their left temporal region; in contrast, cerebral glucose metabolism was significantly reduced in the right temporal and parietal lobes of patients with difficulties on tasks involving visuoconstruction abilities. It might be the case that individuals from the Verbal Dysfunction and Visuospatial Dysfunction subgroups would exhibit similar patterns of cerebral glucose activation. The presence of differential patterns of cerebral glucose metabolism in these subgroups would provide additional validation that they are in fact different from one another, and it would lend further support for hypothesis that CIND is a syndrome of conditions rather than a single population at-risk for dementia.

The NCI individuals who participated in the ACCORD study were an interesting group that deserves some discussion. These participants represent a comparison group rather than a true control group. These cases were referred to a dementia clinic with cognitive complaints, but were found to have no clinically significant cognitive deficits upon assessment. Because the NCI participants in the present investigation were not a true control group, the results reported here are very useful as they reflect the challenge faced by clinicians: assessing individuals who are referred for cognitive assessment and then making a diagnosis.

There are several potential reasons for why the NCI individuals in the present investigation were referred for clinical evaluation but were not classified as CIND. Perhaps the easiest explanation is that some of these cases were false negatives; that is, they were actually cognitively impaired but this impairment went unnoticed by the physicians involved in this study. A second, and more likely, reason is that the physicians were aware of some degree of cognitive impairment but considered this impairment to be clinically insufficient to warrant a diagnosis of CIND. A third potential reason is that the NCI cases in this study had some mild psychopathological symptoms that prompted their referral. For example, some of these cases may have been overly anxious or concerned about experiencing the age-related cognitive decline
that accompanies the normal aging process. These types of individuals have been referred to by others as the “worried well” (Petersen, 2000b).

The diagnostic outcome of NCI individuals is an area of research that merits closer investigation in the future. It might be the case that there are subgroups of NCI individuals that differ with respect to diagnostic outcome, with some of these subgroups improving, some remaining stable, and some converting to CIND or dementia. Being able to predict diagnostic outcomes in NCI individuals would allow for more efficient channelling of clinic resources: limited time and money could be allocated to those NCI individuals who are most at risk for progressing to CIND or dementia.

In this dissertation, rates of conversion to dementia were reported for periods of 2 and 5 years. It is important for future research to monitor the five CIND subgroups over longer periods of time (e.g., 10 to 15 years) to provide more detailed accounts of the developmental pathways of each subgroup. Through extended follow-up, it would be possible to determine whether the patterns of diagnostic outcomes observed here remain stable or whether they change over longer periods of time. Although the Verbal Dysfunction and Visuospatial Dysfunction subgroups had similar conversion rates over periods of 2 and 5 years, it might be the case that these rates change over longer periods of time such that one of these subgroups ultimately has a higher risk for dementia. Increased knowledge of the developmental pathways of these five subgroups would also be useful in assisting patients and their families in making various lifestyle decisions and for planning ahead for long-term care if necessary.

Future research should also focus on determining whether the five CIND subgroups progress to different types of dementia. Over the last decade, there has been increasing awareness of non-Alzheimer types of dementia such as vascular dementia, frontotemporal dementia, dementia with Lewy Bodies, and primary progressive aphasia (Knopman, 2001). The patterns of cognitive deficits observed in these types of dementia have been shown to differ from
one another and from those seen in Alzheimer disease (Ballard et al., 1999; Kertesz & Clydesdale, 1994; Knopman, 2001; Mendez et al., 1996). Although differentiating the various types of dementia is a very difficult and diagnostically challenging process, it will be interesting to explore in the future whether individuals from the five CIND subgroups show elevated risks for converting to different types of dementia. For example, it may turn out that individuals from the Memory Dysfunction and Memory/Verbal Dysfunction subgroups are more likely to progress to Alzheimer disease, whereas individuals from the other three subgroups might be at an increased risk for converting to other types of dementia, such as vascular dementia, frontotemporal dementia, dementia with Lewy Bodies, and primary progressive aphasia. Testing such hypotheses would require very large samples, however, in part because Alzheimer disease typically accounts for the majority of dementia cases (Fratiglioni, 1999). In the present research, the percentage of dementia cases that were diagnosed with possible or probable Alzheimer disease were 61.6% and 75.0% in the CSHA and ACCORD study, respectively.
References


