Predicting Independent Functioning in an Elderly Population: The Evaluation of Working Memory Capacity as a Biomarker of Ageing.

Sara. L. Howard, B. Health Sc. (First Class Honours Psychology)

School of Psychology

The University of Adelaide,

South Australia.

This thesis is submitted to the Graduate Centre in fulfilment for the award of Doctor of Philosophy in Medicine.

Submitted: 16 March 2007

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The ageing process is characterised by declines in physical and cognitive ability and by a general increase in dependence in carrying out daily tasks. Maintenance of functional *in*dependence is critical to quality of life in elderly populations (Black & Rush, 2002; Sulander et al., 2005). Therefore, identifying measures that can predict functional ability is of particular interest to societies with an ageing population.

Due to increases in inter-individual and intra-individual variability with age, chronological age has been demonstrated to be a poor predictor of an individual's functional ability (Bauco et al., 1996; Willis et al., 1992). Consequently, other, more successful indicators, referred to as biomarkers, have been established (e.g. grip strength and visual acuity). However, of these more accurate measures, few are cognitive. This is surprising given the reported strong and positive relationship between independent functioning and intact cognition (Atkinson et al., 2005; Bäckman & Hill, 1996). Therefore, the current project investigated whether a task of working memory capacity (Reading Span), could predict a range of independent functioning outcome measures.

Employing a longitudinal study design (three measurement occasions over approximately 18 months), 150 community-dwelling participants, 70 years of age and over (99 females, 51 males), were tested on a range of cognitive and physiological tasks. Cross-sectional results from logistic and linear regressions showed that chronological age was in fact a significant predictor of all three functional outcome measures. In contrast, Reading Span was a significant predictor only of one outcome measure (reasoning ability). Some of the physiological and sensorimotor biomarkers were found to predict two of the three functional outcome measures. Therefore, cross-sectional results showed that all of the biomarkers were limited in their ability to predict

outcomes measured concurrently and, in the current sample, chronological age was the best predictor of some outcome measures. However, over time, Reading Span became a significant predictor of most of the outcome measures and explained a comparable amount of variance to age. Reading Span also often accounted for more variance than physiological and sensorimotor variables.

The current sample was healthy, independent functioning and cognitively intact. Performance variability was low initially and was further reduced by the presence of selective attrition (i.e. individuals with poorer reasoning and crystallised ability and lesser working memory capacity dropped out of the study). Based on this, it is not surprising that biomarkers were able to explain less than 10% of the variance in *any* outcome measure. In summary, the current study shows that working memory capacity, as measured by Reading Span, is a valuable addition to the assessment of functional ability in an elderly population and highlights the importance of cognition in this context. However, further investigations are required before Reading Span can be described as a *biomarker* of ageing.

DECLARATION

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being made available in all forms of media, now or hereafter known.

13/06/07

Sara Howard Date

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisors, Professor Ted Nettelbeck and Dr. Carlene Wilson. Their support and guidance throughout my PhD candidature was unwavering. I cannot thank them enough for teaching me sound research skills and providing me with an experience I will never forget.

I would also like to thank several academic staff from various institutes; Dr. Lynn Ward for general advice and support; Dr. Janet Bryan, Dr. Akira Miyake and Professor Timothy Salthouse for advice on compiling the test battery and/or provision of test stimuli; and Dr. Sally Record and Dr. Nick Burns for guidance with data analyses. Successful research is truly a collaborative effort.

I must also acknowledge Tess Gregory, a fellow PhD candidate, who worked on a project in a similar area. I am thankful that I had someone to share the joys and frustrations with; your friendship and sense of humour were much appreciated throughout this process. I wish you every success in the future.

Similarly, gratitude must be paid to the participants themselves; I big thank you for pursuing with a project that was at times, quite demanding. I met some wonderful people who inspired me and showed me that ageing doesn't have to be a negative experience. I wish you all health and happiness.

To my current employers, the Primary Health Care Research and Information Service (PHC RIS), I want to thank you for your understanding and flexibility in order for me to complete my thesis. Your faith in my capability is touching.

On a more personal note, I want to say a huge thank you to my family, in particular, my parents. They have always supported me in every way possible. I would not be who, or where, I am today without their love and friendship.

Last but not least, I would like to thank my fiancé, Anthony, who put up with me when I was moody and stressed and encouraged me when I was disheartened. Most of all, I would like to thank him for reminding me that there is *so* much more to life than a PhD thesis. I love you and hope we can grow old together, whatever outcomes that may bring.

CHAPTER 1: INTRODUCTION (PART I)

AGEING AND THE ABILITY TO FUNCTION INDEPENDENTLY

1.1 Dissertation overview: Independent functioning, chronological age, working memory and biomarkers

The ability to function independently is an important issue for older cohorts of the population, defined as those 70 years of age and older, as well as the wider community. Maintenance of this independence involves retaining both physical and cognitive abilities, many of which tend to decline with increasing age. Given that there is great variation within and between individuals as to which abilities decline and the rate of their decline, predicting functional ability from chronological age is frequently likely to be misleading. With this in mind, several other types of measures have been examined for their capacity to predict independent functioning. Of these different types of measures, cognitive measures have been employed in a limited fashion. The cognitive concept of working memory, broadly defined as the ability to simultaneously store and process information, has been shown to explain age-related deterioration in many memory-based activities and other types of cognitive tasks. On the basis of this empirical evidence, and the observation that working memory is ubiquitous in day-to-day tasks, the construct is hypothesised to covary with independent functioning in the elderly. If in fact working memory demonstrates such predictive or prognostic utility (ultimately indicated by longitudinal results), it may be a candidate for use as a prognostic biologically based marker (biomarker) for premature or unfavourable ageing. The main aim of this thesis has been to test this proposition.

This dissertation consists of seven chapters. Chapter 1 provides a context for the cross-sectional and longitudinal research and discusses how this work can contribute to

the general area of cognition and ageing. Chapter 2 focuses on working memory as a predictor of individual differences in independent functioning among elderly people (a construct that has not been employed in this context before). Chapter 2 also provides justification for the manner of operationalising working memory in this thesis. Chapter 3 is a methodological chapter that describes the test battery employed in the current study and the general study design. Chapter 4 provides a more in-depth assessment of the attributes of working memory, as operationalised in the current study, such as the factor structure of the working memory tasks and their association with chronological age. Chapter 5 presents initial cross-sectional results that address whether working memory can correlate concurrently with independent functioning measures. A discussion of some preliminary observations and limitations is also provided here. Chapter 6 discusses different methods of longitudinal analysis; and then presents the longitudinal results analysing the association of biomarker variables with independent functioning measures in an elderly population. A concise discussion of these results concludes this chapter. Finally, Chapter 7 provides an overall discussion of the objectives, achievements and limitations of the current study and the implications of these outcomes in the research areas of biomarkers, cognition and ageing. Future directions in these areas of research are also suggested.

1.2 Independent functioning

Independent functioning in elderly populations is commonly assessed by self-report or observational assessment of ability to perform common "activities of daily living" (ADL). These activities are usually divided into two categories; basic ADL and instrumental ADL. The former term refers to daily tasks involved in personal self-care such as bathing, dressing and feeding. The latter refers to more complex tasks such as cooking, cleaning and managing finances. Instrumental ADL tasks require a finer level

of motor-co-ordination and have a stronger cognitive component (Agüero-Torres, Hillerås & Winblad, 2001). Independent functioning is described as functional capacity or functional status (e.g. Bauco et al., 1996; Bortz, 1990) and incorporates the assessment of "a full range of physical, social, cognitive and psychological functioning" (Kaplan, Strawbridge, Camacho & Cohen, 1993, p.141). In contrast, some studies assess functional *de*pendence or *dis*ability rather than independent functioning. These two terms are used in reference to individuals who are in need of assistance in at least one area of basic ADL (e.g. Agüero-Torres et al.; Marengoni et al., 2004).

Finding strategies to maintain independent functioning is of great importance for several reasons. Firstly, the ability of elderly people to carry out daily activities is critical to their quality of life (Agüero-Torres et al., 2001; Black & Rush, 2002; Bortz, 1990; Sulander, Martelin, Rahkonen, Nissinen & Uutela, 2005; Wang, van Belle, Kukull & Larsen, 2002). Secondly, loss of independent functioning increases with increasing age, on average (Agüero-Torres et al.; Black & Rush; Wang et al.) and, given that many countries have ageing populations, there is great potential for the needs of elderly persons to outstrip the available social, medical and economic resources. The causes of a relative increase in the ageing population are two-fold: sustained low fertility, which creates a population with fewer children; and increased life expectancy, largely due to improvements in medicine, technology, and general health-related knowledge (e.g. life styles that include regular exercise and a low fat diet are beneficial to health). Awareness of the dilemmas associated with an ageing population is not new (Frank, 1946; Shock, 1947). In Australia, this issue has grown in magnitude from 1985 to 2005 with the proportion of people over 65 years of age increasing from 10.3% to 13.1% (Australian Bureau of Statistics; ABS, 2005). In the period from June 2004 to June 2005, this translated to an increase of 63,100 people in this cohort, producing a

total of 2.7 million people aged over 65 years (ABS). This has particular relevance for South Australia because, as of June 2005, this state had the highest proportion of people 65 years or older of any Australian state (15.2% compared to Tasmania, the next highest, with 14.5%; ABS). Therefore, facilitation of independence among elderly people has both individual and community benefits and is now considered a priority for government funded research (e.g. research grants funded by the Australian Research Council; Australian Government, 2004).

1.3 Physical and cognitive declines associated with increasing age

On average, difficulties in independent functioning tend to occur with increasing old age because many physical and cognitive abilities tend to decline with increasing age. Being aware of the aspects that decline with age allows for specific prevention strategies and interventions to be applied to these areas, hopefully thereby minimising further loss of daily functional ability. Details of the two main areas of decline, physical and cognitive, are discussed below.

1.3.1 Physical decline

The influence that ageing has on the human body is widespread and, in general, marked functional decline occurs in all areas (systems) of the body. Table 1 (pp.6-7) provides a short summary of the changes that are considered to be a part of "normal ageing" (Anderson & Craik, 2000; Benjamin, Garman & Funston, 1997; Clark, 1999). Where relevant, their relationship to independent functioning is highlighted. Normal ageing is defined in terms of absence of physical (or mental) disease (Atchley, 1989).

1.3.2 Cognitive decline

Human cognition can be described at a behavioural level in terms of the Horn-Cattell Gf-Gc theory. According to Carroll (1993), this hierarchical model of intellectual functioning provides the soundest approach to theory regarding the structure

of cognitive abilities[1]. "Gf" stands for fluid intelligence and refers to abilities involved in novel reasoning and problem solving. These are abilities that are held to be acquired principally as a result of genetic factors and are minimally influenced by education, learning and culture. In contrast, "Gc" stands for crystallised intelligence and reflects abilities that are influenced by learning and culture, such as vocabulary and general knowledge. Initially, this theory comprised Gf and Gc factors only (Cattell, 1941) but, over time, more abilities or broad cognitive domains have been included (see McGrew & Flanagan, 1999 for a short summary). Consensus on Gf-Gc theory is that it now incorporates eight main cognitive factors or abilities (Compton, Bachman, Brand & Avet, 2000; McGrew & Flanagan). These are described in Table 2 (p.8).

Age-associated decline is commonly observed for many of these factors from about the age of 67 (Schaie & Willis, 1993). Regardless of individual variation (as mentioned briefly in Section 1.1 and as will be discussed further in Section 1.4), a reasonably stable, average age-related pattern occurs involving Gf and Gc. Gf tends to commence declining from an early age (i.e. from the 30s) whereas Gc remains quite robust until the late 70s, when more rapid decline begins (Salthouse, 2004; Schaie, 1994). In summary, prior to Gc decline, Gf, Gs (processing speed) and Gsm (short-term memory) are considered susceptible to the effects of ageing, with evidence that Gf is particularly vulnerable (Li, Lindenberger & Sikstrom, 2001; Lindenberger & Baltes, 1994; Verhaeghen & Salthouse, 1997).

1.4 Chronological age is not a good indicator of individual variation in ageing

Chronological age is a simply-calculated and precise measure of the passage of time that has been and continues to be employed as the main criterion for important, life-altering decisions that are made by governments and other institutions. These

[1] Carroll's (1993) version of this model places more emphasis than has Horn on a general factor, defined at a higher order of factor analysis than that at which Horn's broad general factors are located. This debate, however, is not relevant to the aims of this thesis.

Table 1
Summary of normal changes that occur in the human body with increasing age

System within the body	Observed change with age	Relevance to independent functioning or daily activities
Sensory	The acuity of the majority of the senses declines, although touch is less effected	Poor vision can prevent activities such as reading and driving; poor hearing can prevent social interactions, etc.
Skeletomuscular	Atrophy occurs, resulting in bone and joint problems as well as a decrease in height and weight	Can make mobility and tasks requiring dexterity difficult (e.g. turning on taps, opening jars)
Digestive	There is thinning of mucosal walls, reducing peristalsis, gastric secretions and absorption of nutrients	Can cause swallowing difficulties and reflux problems, and a reduction in the actual amount of nutrient intake
Endocrine	Metabolic rate slows, increasing vulnerability to obesity	Obesity is linked with many health conditions, such as type II diabetes, heart disease and some types of arthritis
Cardiovascular	The ratio of muscle to collagen cells changes. The presence of more collagen cells makes the heart less flexible and decreases overall cardiac output (the blood flowing through the heart and consequently the body)	Situations where increased demands are placed on the heart (e.g. hot weather and physical activity) become more difficult and strenuous. This is likely to restrict behaviour in such situations (i.e. minimise activity in hot weather, less inclined to exercise)

Table 1 (continued)
Summary of normal changes that occur in the human body with increasing age

System within the body	Observed change with age	Relevance to independent functioning or daily activities
Immune	There is a general decline in the body's ability to recognise and respond to foreign antigens	In particular, a decrease in T-cell responses can result in an increased risk of autoimmune disease and cancer
Respiratory	The shape of alveoli in the lung changes, reducing the efficiency of gas exchange (less oxygen/more carbon dioxide in the blood). There is also a reduction in the number of cilia and amount of mucous in the membrane lining	Increases susceptibility to infections and inflammation
Central Nervous	Neuronal death occurs (due to general wear and tear since neurons here cannot regenerate) and there is a subsequent decrease in brain size and weight. There is also a decrease in brain metabolism and blood flow to the brain. Changes also occur to neurotransmitters	These events can cause reduced oxygen and nutrients reaching brain cells. This may result in a build up of toxins within cells and further cell (neuronal) death. This in turn could disrupt thought processes

Table 2
Current factors in Gf-Gc Theory

Name of factor	Abilities factor represents
Fluid intelligence (Gf)	Concerned with the ability to form concepts and manipulate rules and logical relations
Crystallised intelligence (Gc)	Involves answering questions or problems that incorporate relatively familiar materials and processes
Processing Speed (Gs)	Clerical speed and accuracy; speed of automatic processing. For example, perceptual speed is the ability to rapidly search for and compare visual symbols
Short-term memory (Gsm)	Storing information in immediate awareness and then using it within a few seconds. For example, repeating dictated numbers in correct, reversed order. Incorporates the concept of working memory
Long-term retrieval (Glr)	Storing information in long-term memory and having it retrieved readily via association. For example: being able to recall one part of a previously learned but unrelated pair of items when the other part is presented
Quantitative knowledge (Gq)	General math achievement and knowledge; ability to manipulate mathematical and numeric symbols
Auditory processing (Ga)	Perception, analysis and synthesis of auditory stimuli. For example, the ability to discriminate sounds or speech when it is distorted
Visual processing (Gv)	Perception, analysis, synthesis and manipulation of visual stimuli. For example, it looks at imagery, spatial relations and length estimation

decisions have included the following, as discussed by Anstey, Lord and Smith (1996) and Schaie (1994):

(1) mandatory retirement age;

> In recent years this has been abolished in Australia on the grounds of equal opportunity (Andrews, 2003). However, for example, in the USA, the government requires American pilots to retire at 60 years of age regardless of their physical and mental condition (Landphair, 2006).

(2) age-discrimination cases regarding employment;

(3) economic and resource projections; and

(4) driving licence restrictions

> In South Australia, all those aged 70 years and over must have yearly check-ups with their general practitioners to qualify for licences (Transport South Australia, 2004).

However, many researchers have observed that in both cross-sectional and longitudinal studies, chronological age is a poor indicator of an individual's functional ability (e.g. Bauco et al., 1996 and Willis, Jay, Diehl & Marsiske, 1992, respectively). Therefore, alternative measures of functional ability, rather than chronological age, should be used as the main decision-making criteria by governments and other institutions. The reason why chronological age is an unreliable predictor of an individual's functional ability stems from the fact that, as the population ages, the population becomes more heterogeneous. This phenomenon, that there is wide inter-individual and intra-individual variability in ageing, is well-documented (Anstey et al., 1996; Bäckman & Hill, 1996; Baker & Sprott, 1988; Bashkireva & Khavinson, 2001; Bauco et al.; Borkan & Norris, 1980; Dirken, 1972a; Frank, 1946; Heron, 1987; Jarvik, 1975; Piantanelli, Rossolini, & Basso, 1992; Pollock, 2002; Schaie, 1994; Schulz &

Heckhausen, 1996; Shock, 1947; Welford, 1958). For example, each type of body system (as illustrated in Section 1.3.1) can potentially age at different rates in the same individual. Moreover, these systems age at different rates between individuals. McClearn (1997) perhaps best summarises this phenomenon by stating, "there is an enormous range of individuality in the consequences of aging" (p.88). As a result, several alternatives to chronological age have been proposed as more accurate predictors of independent functioning.

1.5 Established factors that influence functional dependence in the elderly

As stated earlier, determination of independent functioning or functional capacity involves the assessment of "a full range of physical, social, cognitive and psychological functioning" (Kaplan et al., 1993, p.141). Clarification of factors that could prevent the loss of, or improve, the ability of elderly people to function independently would be of great benefit. The knowledge that disability or dependence in daily functioning increases with age does not itself explain the processes that are responsible for this dependence (Gill, Williams, Richardson & Tinetti, 1996). Consequently, many studies have investigated what factors may be associated with functional dependence. Table 3 (p.11) describes the main types of factors that have been investigated to date, including in longitudinal studies (Agüero-Torres et al., 2001; Atkinson et al., 2005; Black & Rush, 2002; Ishizaki, Watanabe, Suzuki, Shibata & Haga, 2000; Marengoni et al., 2004; Njegovan, Man-Son-Hing, Mitchell & Molnar, 2001; Wang et al., 2002).

On the whole, the aforementioned studies have consistently indicated that older age, female sex, having lower educational level, doing little or no exercise, the presence of chronic health conditions (and their comorbidities), depressive symptoms and impaired cognition are significantly related to increased functional dependence. From

this group of measures, it is evident that more than just physical abilities determine whether an elderly person can perform activities of daily living (ADL) such as dressing, shopping or meal preparation. Of particular note is the strong and well-documented positive relationship between independent functioning and intact cognition (Atkinson et al., 2005; Bäckman & Hill, 1996; Greiner, Snowdon & Schmitt, 1996; Ishizaki et al., 2000; Marengoni et al., 2004; Moritz, Kasl & Berkman, 1995; Njegovan et al., 2001). Even in high-functioning (independent), non-institutionalised elderly people, "low normal cognitive function may be a useful clinical indicator of older adults at increased risk for loss of independent physical function" (Greiner et al., p.65).

Table 3
Types of factors associated with functional dependence in elderly populations

General factor	Example of specific variables
Demographic	Age, sex, race, education, occupation
Psychological state/Social interactions	Depression, intellectual activity, internal locus of control, social roles
Health/Medical status	Chronic health conditions e.g. type II diabetes, arthritis, stroke
Lifestyle/Health Habits	Smoking, exercise, alcohol consumption
Physical condition	Strength (upper & lower body), mobility e.g. grip strength, walking speed, balance
Cognitive condition	Global status measures (e.g. Mini-Mental Status Exam), language, memory, processing speed

In summary, both physical and cognitive abilities make significant, independent contributions to the development of functional dependence, and in certain cases (e.g. the most physically frail), this contribution is cumulative (Gill et al., 1996). This finding supports the inclusion of cognitive assessment in determination of an individual's functional capacity, rather than principally relying on self-report or performance-based ADL capability (Moritz et al.; Tabbarah, Crimmins & Seeman, 2002).

1.5.1 Previous studies involving cognitive predictors of independent functioning: Limitations

There are two main limitations with research in this area to date. The first relates to the nature of the variables investigated. Most studies investigating predictors of ageing have looked at more than one type of factor (i.e. have included measures from demographic, lifestyle and health factors), allowing for the interrelationship among them to be assessed. Moreover, each of these factors has typically been operationalised by multiple variables, creating a detailed examination of that factor. However, an exception to this approach applies in the area of cognition. The overwhelming majority of studies (including longitudinal) have used only global measures of cognition, such as the Folstein Mini-Mental State Exam [(Folstein, Folstein & McHugh, 1975) Atkinson et al., 2005; Boyd & Dawson, 2000; Gill et al., 1996; Hébert, Brayne & Spiegelhalter, 1999; Kelly-Hayes, Jette, Wolf, D'Agostino & Odell, 1992; Marengoni et al., 2004; Njegovan et al., 2001; Wilson et al., 2002]. This limited inclusion of cognitive measures obscures the identification of potential sources of individual differences (Wilson et al.). In addition to this, global measures of cognition are not particularly sensitive in high-functioning, well-educated samples. The limited utility of the Mini-Mental State Exam (MMSE) has been emphasised by Atkinson et al. who argued that the MMSE "is not a sensitive means of detecting subtle declines in cognitive function in community-dwelling, cognitively intact persons" (p.1201).

One study, conducted by Binder, Storandt and Birge (1999), has included a more diverse range of cognitive measures. Their measures focused on psychomotor speed and memory and included the following cognitive tests: associative learning and delayed recall, word fluency and the Trail Making Test (parts A and B). Their results showed that a factor representing processing speed (established via factor analysis) was

significantly related to daily functioning performance, where slower processing speed was associated with increased dependence. Given that extremely few studies exist that have included a comprehensive range of specific cognitive measures in this area of research, there is a clear need for this particular issue to be resolved.

The second issue is centred on the design of most studies. A lot of studies looking at independent functioning in elderly populations have conducted cross-sectional analyses only; they have compared performance of individuals varying in age on various tasks at the same point in time. This research is important and has permitted the identification of factors that might be associated with functional dependence in the elderly. However, there has been limited longitudinal research on this topic (Willis et al., 1992). Longitudinal research is important in order to specify the relationship between change in one variable (e.g. cognition) and change in another (e.g. physical ability; Tabbarah et al., 2002). Moreover, longitudinal studies allow for the establishment of whether cognitive impairment, for example, at one point in time can predict subsequent onset of new limitations in activities of daily living at a second point in time (Moritz et al., 1995). It is demonstration of this predictive validity that is most valuable because it allows for preventive interventions to be developed. The research area of biomarkers is concerned with the predictive ability of different types of measures. However, before discussing biomarker research, the cognitive construct of working memory will be explored as a means for explicating the link between cognition and functional ageing.

1.6 Working memory

Working memory is "a limited capacity system which is responsible for the processing and temporary storage of information while cognitive tasks are performed" (Van der Linden, Brédart & Beerten, 1994, p.145). The working memory system is

called into action when performing activities like mental arithmetic where intermediate solutions need to be remembered while further processes are carried out. When the storage demands of a task (or amount of information that has to be kept readily accessible for further use) are high, then successful performance in the manipulation, updating or further processing of information becomes more difficult. In terms of the mental arithmetic example, long division (e.g. 5392/31) has increased storage demands compared to short division (e.g. 27/3). That is, there are more numbers to remember in general as well as remembering the numbers that are forming the solution. As a result, processing steps intrinsic to mental arithmetic, like "carrying numbers" may be forgotten, resulting in an incorrect answer. Conversely, if attention is focused on the processing steps, this may lead to forgetting the numbers stored as the partial solution, also resulting in an incorrect answer. The threshold for when difficulty is encountered on this type of dual-activity task appears to lower with increasing age.

1.6.1 Age-related decline in working memory

The inferior performance of older adults compared to younger adults on tasks assessing working memory has been well-documented (Babcock & Salthouse, 1990; Craik, 1977, 1994; Dobbs & Rule, 1989; Kirasic, Allen, Dobson & Binder, 1996; Light & Anderson, 1985; Meguro et al., 2000; Morris, Gick & Craik, 1988; Salthouse, 1991b, 1994; Salthouse & Babcock, 1991; Van der Linden, Beerten & Pesenti, 1998; Van der Linden et al., 1994). However, what is less well agreed upon is what specific aspect of the working memory system is responsible for this age-association (Babcock & Salthouse; Light & Anderson; Myerson, Emery, White & Hale, 2003). Given that the concept of working memory is commonly defined as involving concurrent storage and processing, then age-related decline may stem from the storage capacity, efficiency of processing or both (Babcock & Salthouse). Which components of the working memory

system are most sensitive to the effects of ageing also depends on the particular working memory model or theoretical grounds selected. Details on theoretical aspects of working memory will be addressed in Chapter 2. In addition to deteriorating with age, working memory is a powerful mediator of other cognitive processes.

1.6.2 Mediating role of working memory

Many studies have shown that by statistically controlling for performance on a test assessing working memory, the relationship between age and other cognitive tasks is substantially reduced or eliminated entirely. That is, working memory has been shown to have explanatory power in regard to age deterioration in a range of cognitive abilities. For example, computation span (a test of working memory) was found to explain age-related performance on a test of inductive-reasoning (Gf) and a spatial paper-folding task [(Gv) Salthouse, Mitchell, Skovronek & Babcock, 1989]. Other cognitive tasks where age-related variance has been demonstrated to be mediated by performance on a working memory task include: language/discourse processing (Light & Anderson, 1985; Stine & Wingfield, 1990), episodic memory (Park et al., 1996; Van der Linden et al., 1999) and declarative learning (Kirasic et al., 1996). As a consequence of such findings, working memory capacity has been proposed to be a general processing resource. (Li et al., 2001; Park, 2000; Verhaeghen & Salthouse, 1997).

Processing resources are considered to be "general factors that are necessary for the execution of multiple processing components", which are involved in multiple cognitive tasks (Hultsch, Hertzog, Dixon & Small, 1998, p.40). Furthermore, they "are available in limited quantities at any specific point in the execution of a cognitive task" (Hultsch et al., p.40). Processing resources are indices of how much mental energy or processing power an individual has to draw upon when performing a cognitive task (Park, 2000). In other words, age-related deterioration in cognitive abilities can be

mediated by reductions in these processing resources (e.g. working memory) rather than by declines in task-specific components (e.g. mathematical skill or Gq, as per the earlier mental arithmetic example). In addition to working memory, other proposed processing resources are processing speed, inhibition and sensory function (Park)[2]. There is evidence to suggest that measures of these resources are related to each other (Li et al., 2001; Park) and conceivably it is unlikely that one type of processing resource is responsible for age-related deterioration across all tasks. Finally, it has been suggested that the amount of cognitive processing resource an individual can draw upon in a given situation influences functional ability in daily activities, including tasks such as managing finances and medications (Park, 1997, 1999).

1.6.3 Ecological validity of working memory in independent functioning

Given that a majority of daily tasks involve some degree of remembering information (storage) whilst thinking about or doing something else at the same time (processing), it is not difficult to appreciate the ubiquitous nature of working memory. Logie (1999) defined working memory as:

> …a cognitive function that helps us keep track of what we are doing or where we are moment to moment, that holds information long enough to make a decision, to dial a telephone number, or to repeat a strange foreign word that we have just heard (p.174).

Moreover, it is not only elderly people who have been reported to experience difficulties on tests of working memory. People with schizophrenia (e.g. Okada, 2002), multiple sclerosis (e.g. Ruchkin et al., 1994), depression (Geva, 2002), and Alzheimer's Disease (e.g. Collette, Van der Linden, Bechet, Belleville & Salmon, 1998) have all been shown to have reduced working memory capacity and impaired daily functioning.

[2] This thesis has formed part of a larger project that has included consideration of processing speed as a marker for cognitive decline. This possibility is well supported by the work of Salthouse and colleagues (e.g. Salthouse, 1991b).

The latter two groups hold particular significance in the current study because the incidence of these illnesses increases with age. Therefore, the presence and potential confounding influence of depression and dementia (i.e. Alzheimer's Disease) had to be controlled for in the current study (see Chapter 3 for a more detailed explanation of how this was achieved).

1.7 Relationship of independent functioning to the concept of 'biomarkers'

Research focused on independent functioning or functional ability in elderly populations has looked to the validation of biomarkers. Generally, a biomarker can be described as "a physical or laboratory measure with diagnostic or prognostic usefulness" (Pollock, 2002, p.644). Baker and Sprott (1988) defined a biomarker of ageing as "a biological parameter of an organism that either alone or in some multivariate composite will, in the absence of disease, better predict functional capability at some late age than will chronological age" (p.223). In other words, biomarker research captures the predictive or longitudinal aspect of independent functioning in the elderly; it has predictive power. It will therefore be clear that, as used here, a higher level cognitive construct like working memory can have potential as a biomarker (refer to Chapter 7, Section 7.2 for a detailed account of defining criteria). Measures or factors that can provide a more accurate picture of an individual's functional ability than chronological age are considered to be measuring an individual's functional age (Dirken, 1972a; McFarland, 1953). Synonyms of 'functional age' include 'physiological age' (Frank, 1946; Murray, 1951) and 'biological age' (Benjamin, 1947; Birren, 1959; Borkan & Norris, 1980).

In order to determine functional age, a large number of physiological and psychological variables need to be considered. This is due to increasing heterogeneity with age (as already described in Section 1.4). Consistent with this, it is unlikely that a

single biomarker will be found to predict or describe why two people of the same chronological age differ in regard to their health, robustness and so on (McClearn, 1997). Similar to independent functioning research, biomarker research has also employed several different categories of variables. Table 4 describes commonly-used measures in biomarker research. These are essentially similar to those listed in Table 3 (p.11) only with a different, slightly more specific nomenclature. Biomarkers have been applied in various areas of research. For example, literature on biomarkers can be found in relation to promoting health behaviour change (McClure, 2002), psychiatric treatment (Pollock, 2002), neurodegenerative diseases, such as Alzheimer's Disease (Klunk, 2002; Papassotiropoulos & Hock, 2002), and more recently in research on tobacco-related health outcomes (Hatsukam, Benowitz, Rennard, Oncken & Hecht, 2006).

Table 4
Examples of biomarker variables by different biomarker categories

Category	Example of specific variables
Sensorimotor	Grip strength, simple reaction time, visual acuity, auditory acuity
Physiological/biomedical	Blood pressure, vital capacity, forced expiratory volume
Anthroprometric/Morphologic	Skin elasticity, weight
Psychosocial	Socio-economic status, stress, personality
Behavioural	Diet, activity, sleep duration, health status
Dentition	Number of teeth, plaque index
Cognitive	Vocabularly (Gc), digit symbol (Gs)

1.7.1 Limitations of biomarker research

There are two main problems with biomarker research. The first has to do with the inappropriate use of chronological age as an outcome (dependent) measure. That is, several studies have used functional age variables (e.g. grip strength, blood pressure) as

independent variables in order to predict chronological age (Dirken, 1972b; Finkel, Whitfield & McGue, 1995; Furukawa et al., 1975; Murray, 1951; Webster & Logie, 1976). Borkan and Norris (1980) stated that this has been the most prevalent approach to functional age estimation historically. Moreover, this approach has continued in more recent times (e.g. Finkel et al.) despite the fact that Heron (1987) was one of the first to realise the circularity of such assessment and stated that "a combined measure or functional age score is not that useful if it equates with or predicts chronological age" (p.36). It is important for hypothesised biomarkers to demonstrate greater predictive validity than chronological age for functional outcomes (Lord, Anstey, Williams & Ward, 1995).

The second issue is similar to one previously raised in Section 1.5.1. From Anstey et al.'s (1996) review of biomarker research, it is clear that the majority of researchers have focused on sensorimotor or physiological measures with cognitive measures seldom used in prediction of functional age. However, unlike the studies included in Section 1.5.1, tests representing more specific or narrow cognitive domains have been included in biomarker research. The problem lies in their low frequency of use.

1.8 Summary and contribution of this dissertation

Independent functioning, as measured by activities of daily living, represents "a complex mix of physical and physiological capabilities and cognitive, social and motivational factors. Multifactorial determination is to be expected" (Kaplan et al., 1993, p.152). Given the number and nature of factors that have been shown to influence functional dependence, deterioration and loss of autonomy should not be viewed as a predestined or irreversible consequence of ageing (Black & Rush, 2002; Hébert et al., 1999). According to Anderson, James, Miller, Worley and Longino (1998), "a

substantial proportion of those with disabilities will remain stable or actually improve over time" (p.S24). The more factors that are discovered to be associated with and predict independent functioning (or functional *de*pendence), the more likely researchers are to develop successful intervention and prevention programs to keep the growing elderly population functioning independently.

A noticeable area for development in ageing and biomarker research is the investigation of possible roles and contributions that specific cognitive abilities make to functional capacity or independence. For reasons discussed in Section 1.6, working memory is a strong and plausible candidate in this context. Therefore the primary contributions of this dissertation to the field of cognitive ageing are via the examination of specific cognitive abilities (working memory) examined with longitudinal study design, as a predictor (biomarker) of functional ageing.

CHAPTER 2: INTRODUCTION (PART II)

WORKING MEMORY IN THEORY AND PRACTICE

2.1 Overview: Importance of clarifying working memory theory and tasks

Chapter 1 discussed the well-documented age-related deterioration in working memory (WM) performance, the importance of WM to high level cognitive tasks (such as language processing and reasoning), and the high ecological validity WM has with respect to daily activities. Prior to assessing WM's utility as a prognostic index for subsequent reduced functional ability of elderly persons (i.e. via regression results, see Chapter 5), it is important to discuss in more detail what the construct of WM represents and how it is best, or most commonly, operationalised. Although the primary goal of the current study is to assess the practical utility of WM tasks (and the related theoretical perspective adopted) as a predictor of ageing outcomes, operationalising WM for this purpose requires consideration of the theoretical concerns. As a result, this Chapter provides a definition of WM, a brief account of the development of the concept of WM, and discusses the particular theoretical model adopted for this study (Baddeley & Hitch's, 1974 tripartite model) and how this model is operationalised.

2.2 What is working memory?

Hultsch et al. (1998) stated that:

> In general, working memory refers to processes and structures involved in temporarily holding information in mind while manipulating that information or combining it with further incoming information to achieve a variety of goals, such as comprehension, problem solving, or learning (p.44).

The overwhelming majority of working memory (WM) definitions include this notion of simultaneous storage and processing of information (as evidenced by a

definition provided in Chapter 1, p.13). However, very few authors have specifically stated what is meant by "processes" or "structures". Perhaps this is because these two concepts are so closely associated, as recognised by Atkinson and Shiffrin (1968). Thus, it was appropriate to include a description of the processes and structures of memory as part of the definition of WM.

2.2.1 Memory structures

Any permanent feature of memory is referred to as a memory structure (Atkinson & Shiffrin, 1968). That is, structure includes the physical system/basic memory stores (e.g. sensory register) as well as unvarying processes that are present across all cognitive situations (e.g. the act of copying/transferring information from one store to another). Table 5 describes the three fundamental memory stores.

Table 5
Memory structure: A description of three basic memory stores

Name of store	Function
Sensory register	Registers sensory input. Visual stimuli have received most research attention. The register scans a visual image and holds a more or less photographic trace of this image for several hundred milliseconds, but thereafter it begins to decay. Subsequent visual stimuli can modify or remove prior images
Short-term store	Receives selected inputs from the sensory register and from the long-term store. Information here decays (is lost) in about 30 seconds unless a control process (e.g. rehearsal) maintains this information
Long-term store	Is a relatively permanent repository for information that has been transferred (copied) from the short-term store.

There may be direct transfer of information from the sensory register to the long-term store. As each element in the sensory register is scanned, a match-up procedure takes place that searches for relevant existing information in the long-term store. For example, if a person sees a dog, the verbal name and other information (e.g. smell, danger) is recovered from the long-term store and fed into the short-term store to

be used in the particular cognitive task occurring at that point in time. The act of transferring information is part of the memory structure; however, the amount and form of the information transferred varies greatly. The latter aspects are under the control of the individual and as such fall under the category of memory processes (Atkinson & Shiffrin, 1968).

2.2.2 Memory processes

Memory processes typically refer to control processes (e.g. coding procedures, mnemonics, rehearsal operations and search strategies). The use and exact nature of these processes are under the control of the individual. They are modifiable at the individual's discretion depending on the instructions or demands of the task and the individual's personal experience (e.g. meaningfulness of the material in question, exposure to types of processes like mnemonics). Therefore, the processes used by one individual may vary from one task to another (Atkinson & Shiffrin, 1968), as well as varying between individuals on the same task. Of particular salience are coding procedures and rehearsal operations. Coding procedures involve alterations to the information in short-term store as a result of a search of the long-term store. That is, strong, pre-existing associations already in long-term store are used to assist in remembering the information in short-term store. Rehearsal operations are carried out with the purpose of lengthening the period of time that information stays in the short-term store. By doing so, this information is more likely to be remembered permanently (i.e. in the long-term store); just like in the scenario of repeating the same telephone number, or person's name, over and over again. Control processes are also responsible for which particular sensory inputs or information in the sensory register receive further processing and subsequent transfer to the short-term store. As discussed in the next section, the key feature of the working memory construct is its focus on control

processes or how the stored information is maintained and used, rather than where the information is stored, as per the concept of short-term memory.

2.3 The development of working memory

In the 1960s, a popular and influential model of short-term memory was the Modal Model developed by Atkinson and Shiffrin (1968). This model included separate memory stores (as discussed above), with the short-term store playing the crucial role of transmitting information to and from the long-term store. This model, like most short-term memory models at the time, placed heavy emphasis on the structural aspects of memory. However, like most theories and their models, problems with the Modal Model became apparent over time. According to Baddeley (1990b), discrepancies between short-term memory theory and empirical data stemmed from four main areas of research:

(1) neuropsychological data;

(2) short-term storage and long-term learning;

(3) recency effects and short-term storage; and

(4) the coding of different types of information.

It is perhaps the double dissociation of memory evidenced by brain-damaged individuals (i.e. neuropsychological data) that has most shaped the development of more recent memory theories, in particular resulting in the fractionation of memory (Baddeley, 1990b; Logie, 1995). For a more detailed account of these issues consult Baddeley.

Baddeley and Hitch's (1974) working memory (WM) model provided an explanation for these discrepancies (Logie, 1995; Baddeley, 1990a). In essence, their model emphasised processing or functional aspects of memory and the importance of memory to cognition (Miyake & Shah, 1999; Parkin, 1993). Furthermore, this notion of

WM was considered to replace the short-term memory concept; it combined memory storage and processing within one system (Kintsch, Healy, Hegarty, Pennington & Salthouse, 1999). Alternatively, the concept of short-term memory can be retained when considered as a processing continuum. In this case, WM can be viewed at the "active" end of short-term memory and primary memory at the "passive" end (Craik, 2000). That is, tasks that require manipulation or transformation of stored information are referred to as tasks of WM and tasks that require information to be maintained over a short period of time are referred to as tasks of primary memory. As will be seen in Section 2.6, nomenclature surrounding WM and respective tasks can be problematic.

2.4 Working memory models

In the 30 years since the Baddeley and Hitch (1974) model of working memory (WM) was developed, many more theories and models of WM have been proposed. Almost a decade ago, the book, *"Models of Working Memory. Mechanisms of Active Maintenance and Executive Control"* (1999), edited by Miyake and Shah, provided a comprehensive discussion of 10 different models of WM. To date, there is nowhere else where such a detailed assessment of so many WM models has been made. Although the editors acknowledged that their book was not inclusive of all WM models, it continues to provide insight into a wide range of models. Several of the models in this book were based on symbolic computational architectures but were found to have attracted only limited empirical investigation, particularly in the field of human ageing.

In the area of ageing, working memory has been most frequently conceptualised as a general, limited-capacity system or resource. The Baddeley and Hitch (1974) tripartite model has been employed to a lesser extent in this research area; but it is extremely useful in this context because "it allows the possibility of a functional account of exactly how and why cognitive changes occur with age, rather than the basic

assertion that age differences are 'resource' differences" (Phillips & Hamilton, 2001, pp.101-102). Moreover, many studies have employed the tripartite model in this way to examine age-related cognitive decline (e.g. Salthouse, 1994; Salthouse, Kausler & Saults, 1988; Fisk & Warr, 1996). For some two decades, there has been widespread support for the general application of Baddeley and Hitch's model (e.g. Cantor, Engle & Hamilton, 1991; Lehto, 1996; Logie, 1999; Loisy & Roulin, 2003; Salthouse; Van der Linden et al., 1994) and this acceptance was the main justification for applying the framework to the current study.

2.5 Baddeley and Hitch's (1974) tripartite model of working memory

The original model (Baddeley & Hitch, 1974) included three components (hence "tripartite" model); the phonological loop, the visuo-spatial sketchpad and the central executive. Each component will be discussed in more detail shortly. It should also be acknowledged that modifications have been made to these components since the original model was first proposed and these modifications will also be described where relevant. More recently, Baddeley (2000a) proposed that a fourth component be included in this model. He termed this component the episodic buffer. The episodic buffer appears to be a combination of central executive functions and mnemonics. Given that this component is relatively new and has received only limited empirical investigation (Baddeley), it was not assessed in the current study.

In the tripartite model, "working memory forms an interface between memory, attention and perception" (Baddeley, 1998, p.237) and according to Baddeley and Logie (1999), working memory can be defined as comprising:

> ...those functional components of cognition that allow humans to comprehend and mentally represent their immediate environment, to retain information about their immediate past experience, to support the acquisition of new knowledge, to

solve problems, and to formulate, relate, and act on current goals (pp.28-29).

Figure 1 (p.33) illustrates how the three working memory components interact and how working memory forms an interface between perception, attention, memory, and action (Baddeley, 1996).

2.5.1 The phonological loop

The phonological loop is one of the two 'slave' or storage systems that were originally proposed by Baddeley and Hitch (1974). This component is termed a 'slave' system because information that is stored and undergoes rehearsal here (and likewise in the visuo-spatial sketchpad) is under the control of the central executive. The phonological loop is considered to be specialized for the temporary storage of verbal material (Logie, 1995). This component has been the most investigated and best understood of the three and is closest to the original concept of a short-term store (Baddeley, 2000b). It can be further divided into the phonological store and the articulatory control process. The phonological store holds speech-based information for a brief period of time (approximately two seconds) and then decays. This information trace does not decay if it is maintained by the other component – the articulatory control process. This second component employs articulatory rehearsal, which refreshes the memory trace and thus maintains the to-be-remembered information (Baddeley, 1990b; Baddeley 2000b; Gathercole & Baddeley, 1993). Rehearsal of to-be-remembered information may be subvocal. Essentially, the amount of information that the phonological loop can store equates to how much information (how many words) can be remembered or rehearsed in about two seconds. If the words are short and take little time to rehearse, then more information can be remembered. If the words are longer, then fewer words may be remembered. This is known as the word length effect. The articulatory control process also enables the registration of visually presented, but

nameable, material in the phonological store. That is, this process converts the visual information (grapheme) into phonological code [(phoneme) Baddeley, 2000b]. Consequently, "a memory trace in the phonological store might stem from either a direct auditory input, or from the subvocal articulation of a visually presented item" (Baddeley, 2000a, p.418).

It should be noted that in earlier versions of this model (Baddeley & Hitch, 1974; Baddeley, 1986) the phonological loop was termed the 'articulatory loop'. The name change was made to emphasise that this component is not restricted to articulation processes (Baddeley, 2002). Studies demonstrating a phonological similarity effect, irrelevant speech effect, word length effect, and articulatory suppression provided evidence for the phonological store, or a speech-based storage structure. Many studies have provided a more detailed description of these effects (e.g. Baddeley, 1990a; Baddeley, 1992; Baddeley, 2000a; Baddeley & Hitch, 1994; Gathercole & Baddeley, 1993; Logie, 1995). The phonological loop appears to play an important role in activities such as the development of language in children, and second language learning (Baddeley, 1998; Baddeley & Hitch).

The classic digit span procedure (i.e. Wechsler's digit span forwards) is used to assess the phonological loop (Baddeley, 2002). As the amount of information entering the phonological loop increases (e.g. there are eight or nine numbers to repeat back instead of three or four), a point is reached where the first item (number) presented has decayed before the last number has been processed or rehearsed. This causes errors to be made when repeating back the number sequence. It is in this way that memory span has a limited capacity (Baddeley, 1996). Only a certain amount of information can be maintained at any one time so in order to process new/incoming information, some previous information is lost. Anecdotal evidence suggests that verbal material may also

be represented in a visuo-spatial or imagery format (Baddeley, 1996). For example, in trying to remember numbers, the concept of a ladder may be employed, where each rung of a ladder represents each number presented. If this is the case, then there is the advantage of drawing upon the storage and rehearsal capacities of the visuo-spatial sketchpad for the purposes of maintaining successful performance on a task such as digit span forwards. The central executive (see below) can also employ strategies (e.g. grouping numbers into pairs) to improve or maintain memory performance.

2.5.2 The visuo-spatial sketchpad

The second slave system, which has received comparatively less research attention, is the visuo-spatial sketchpad. It is conceived as an approximate visual analogue to the phonological loop, only more complex (Baddeley, 1996). This component was formerly known as the visuo-spatial scratchpad, and also changed names in order to acknowledge the involvement of other components, such as visual elements in addition to spatial elements (Baddeley, 1986; Baddeley, 2002). Professor Robert Logie has conducted the majority of research into this component and has suggested that it also comprises two subsystems: the visual cache or 'inner eye' (visual working memory) and the inner scribe (spatial working memory). The former refers to a passive store that holds information about static visual patterns, such as information pertaining to the appearance of objects (e.g. colour, shape). The latter holds dynamic information about movement and movement sequences and is typically concerned with location and direction (Baddeley, 1998; Baddeley & Hitch, 1994; Logie, 1995). The inner scribe is also associated with the control of physical actions (Logie). Furthermore, Logie considered the spatial working memory component to provide a means for rehearsal, manipulation and transformation of visual and spatial information. He viewed this rehearsal mechanism to be akin to that of the phonological loop; however debate

continues over whether this is indeed the case (for example, see Baddeley, 2002, p.89 for a discussion of this point). According to Baddeley (2002), the visuo-spatial sketchpad is believed to provide "an interface between visual and spatial information, accessed either through the senses or from long-term memory" (p.88).

The visuo-spatial sketchpad is thought to play a role in the following tasks: sentence processing (namely in the area of spatial syntactic forms, such as 'below' or 'inside'; Baddeley, 2002); planning and executing spatial tasks; keeping track of changes in the visual perceptual world over time; directing spatial movement; and maintaining orientation in space (Baddeley & Hitch, 1994). The capacity of the visuo-spatial sketchpad has been assessed by many different types of tasks, including: remembering the location of dots in a grid (dot memory); remembering the sequence of blocks that have been tapped (Corsi Blocks); deciding whether one image is a rotated version of the other (rotation task); and following a target through space (tracking tasks). Finally, it is argued that the visuo-spatial sketchpad is more dependent or closely affiliated with the central executive than the phonological loop. This may be due to the fact that humans are predominantly verbal or phonological by nature and, consequently, the process of using visual imagery is relatively less automatised or well-practiced (Baddeley, 1996).

2.5.3 The central executive

The central executive is modality-free and is often considered to be the most important part of the working memory model (e.g. Baddeley, 1996; Baddeley, 2002; Kyllonen & Christal, 1990; Lehto, 1996). Baddeley (1986) suggested that the central executive can store some information like the slave systems, in addition to providing attentional resources. However, even currently, this remains the most poorly defined component of the tripartite model. Originally it was loosely described "as a limited

capacity pool of general processing resources" (Baddeley, 2002, p.89). It has been widely acknowledged that the central executive was initially modelled on, if not equated with, Norman and Shallice's (1980) neuropsychological model that described the control of information processing (Baddeley, 1986; Fisk & Warr, 1996; Lehto; Morris & Jones, 1990). The Norman and Shallice model comprises two controlling systems. The first system is termed contention scheduling. This activates semi-automatic schemata necessary to accomplish routine tasks (Lehto) and according to Shallice (1982) it is "a unit that can control a specific overlearned action or skill" (p.200). The second system is referred to as the supervisory attentional system and is required or called upon in novel situations, or poorly learned sequences; planning future actions and decision-making; and trouble-shooting in difficult or dangerous situations, such as when automatic processes are inappropriate (Baddeley, 1986; Lehto; Morris & Jones). It is the latter component to which the central executive is believed to equate.

Fortunately, more is known now about the central executive than in 1974, with various researchers attempting to reveal its component parts. Although there are still many questions remaining as to the exact functions of the central executive and how they may interact (Baddeley, 1996; Loisy & Roulin, 2003), the roles described in Table 6 (p.32) have been suggested. These roles are principally attentional in nature and some have been grouped according to similar or overlapping functions (Baddeley, 1996; Baddeley, 2002; Baddeley & Hitch, 1994; Gathercole & Baddeley, 1993; Loisy & Roulin). It is not difficult to appreciate that the central executive is the most complex working memory component. Not surprisingly, it plays a role in a multitude of tasks such as mental arithmetic, recall of lengthy lists of digits, logical reasoning and semantic verification (Gathercole & Baddeley).

Table 6
Summary of the proposed functions of the central executive

General function type	Description of function
Attentional control of action	This provides the capacity to override habitual response patterns when initiating a new behaviour, as modeled from the supervisory attentional system
Selective and focused attention	This provides the capacity to attend selectively to one stimulus and to discard or inhibit non-pertinent stimuli. Thus, the central executive directs what information an individual processes and stores
Dividing and switching attention and the co-ordination of performance on separate tasks	This relates to dual task performance or the ability to complete two different tasks simultaneously. One task may rely heavily on the phonological loop and the other on the visuo-spatial sketchpad. This can also refer to the ability to switch retrieval strategies
Forming an interface between slave systems and long-term memory	This is a more recent development and is related to, if not the same as, the episodic buffer. It looks at the capacity to hold and manipulate information in long-term memory

The most commonly used task to assess central executive functioning has been random generation (Miyake, Friedman, Rettinger, Shah & Hegarty, 2001; Oberauer, Süss, Schulze, Wilhelm & Wittmann, 2000). In this task a participant is required to produce a random sequence of numbers or letters at a given rate (e.g. one item per second). To do this he or she must inhibit well-learned sequences such as consecutive strings (e.g. 1,2,3 or A,B,C), groups (e.g. 3,5,7 – representing odd numbers) and abbreviations (e.g. RSVP – an acronym for a French phrase, meaning 'respond please'). This task may entail ignoring or inhibiting information retrieved from long-term memory. The participant must also keep track of the numbers or letters already said in order to keep the output as random as possible. In general, successful performance of

this task requires the selection of new retrieval strategies, one of the many central executive functions outlined in Table 6 (p.32).

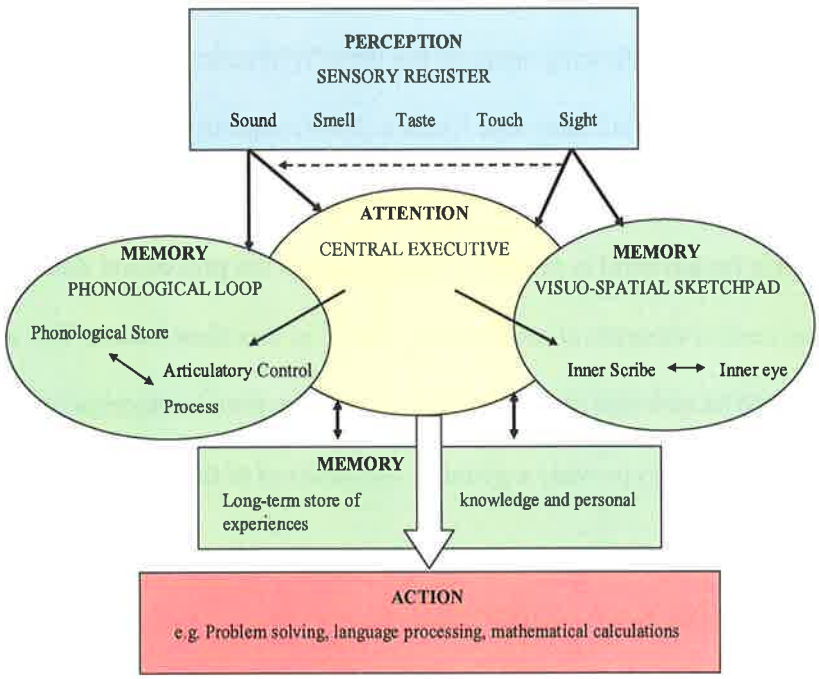


Figure 1 A representation of Baddeley and Hitch's (1974) working memory model

Information received by the senses undergoes multiple stages of processing. The central executive controls which information is processed and provides active searches for relevant information in the long-term store. The dashed arrow is a reminder that visual information can also be coded or transformed into verbal material (as described on p.28). The combination of perception, attention and memory components results in action.

2.6 Operationalisation of working memory in the current study

Working memory (WM) tasks are those that are considered to measure aspects of both storage (structure) and processing (function). A broad range of tests is available to measure the WM construct and the same tests have been used by researchers applying different theoretical perspectives. Thus, deciding how to operationalise WM is a critical decision and each task is likely to measure slightly different aspects of WM, or the same aspect to different extents (Dobbs & Rule, 1989; Lehto, 1996; Roberts &

Gibson, 2002; Waters & Caplan, 1996). Moreover, WM tests have been described under various nomenclatures, including WM capacity (e.g. Engle, 2001), WM span (e.g. Towse, Hitch & Hutton, 2000), and complex span (e.g. Cantor et al., 1991), which makes searching for tests a lengthy process. All of these terms have been employed for tasks that measure more than just storage (that is, more than the earlier concept of short-term memory). In the following sections the three WM tasks employed in the current study to operationalise Baddeley and Hitch's (1974) tripartite model (Reading Span, Dot Matrix and Digit Ordering) will be described, together with the rationale for their selection. This background is provided, in addition to the procedural details in Chapter 3, given the central theoretical importance of WM in this dissertation. The appropriate selection of a WM task that can predict important functional outcomes was critical, if the current study was to provide a genuine contribution to the area of cognition and ageing.

2.6.1 Reading Span: An overview

The Reading Span Test was developed by Daneman and Carpenter (1980) and many different versions or adaptations of it currently exist. In essence, this test consists of reading sentences and remembering the final word of each sentence for later recall. Fisk and Warr (1996) described it as "a good overall measure of the working memory system" (p.317). Reading span tests assess working memory (WM) span, "a measure of how much temporary information can be maintained while performing processing operations" (Baddeley & Hitch, 2000, pp.131-132). Reading span-type tasks are synonymous with WM assessment (Baddeley & Hitch; Meguro et al., 2000; Wareing, Fisk, Murphy & Montgomery, 2005; Whitney, Arnett & Driver, 2001). This is evidenced by their frequent use in the field of cognition (e.g. La Pointe & Engle, 1990), particularly for the study of cognitive ageing (e.g. Kirasic et al., 1996; Park et al., 1996;

Van der Linden et al., 1999) and individual differences (e.g. Oberauer et al., 2000). The current project incorporated all of the above aspects (i.e. cognition, ageing and individual differences) and Reading Span was therefore considered an important test to include. Before detailing which particular components of the WM construct the Reading Span Test measures, a brief review of its origins and more recent modifications is provided below.

2.6.1.1 The original Reading Span Test

Daneman and Carpenter (1980) investigated various tasks that could explain individual differences in language comprehension among undergraduate university students. The impetus for their work was the lack of correlation between traditional short-term memory tasks (otherwise referred to as the storage component of working memory), such as word span, and measures of reading comprehension. Their investigations resulted in the development of a new task that assessed working memory capacity (that is, both storage and processing components). This task was named the Reading Span Test. Daneman and Carpenter's study actually proposed several variations of the Reading Span Test; reading sentences aloud, reading sentences silently, and listening to sentences. Each of these variations will now be described, noting that the last two involved slightly different methods from the first.

2.6.1.2 Reading sentences aloud

This version of the Reading Span Test involved reading one sentence at a time from a card. According to Daneman and Carpenter (1980), simply reading the sentence involved sentence comprehension, or the processing characteristic of working memory (WM). Participants were instructed to remember the final word of each sentence; this involved maintenance and retrieval of words in memory, or the storage characteristic of WM. As soon as the participant had finished reading one sentence, it

was replaced with a new, unrelated sentence (i.e. a new card). This fairly quick presentation rate was employed to avoid participants overtly rehearsing the final words of the sentences. The final word of each sentence was to be recalled (verbally and in correct serial order) after the presentation of between two and six sentences. The number of sentences to be read increased by one after three attempts at each sentence level. Participants were alerted to the fact that the number of sentences would increase during the test. A blank card was the cue for recall of the final words (see Figure 2 for a pictorial representation of this procedure).

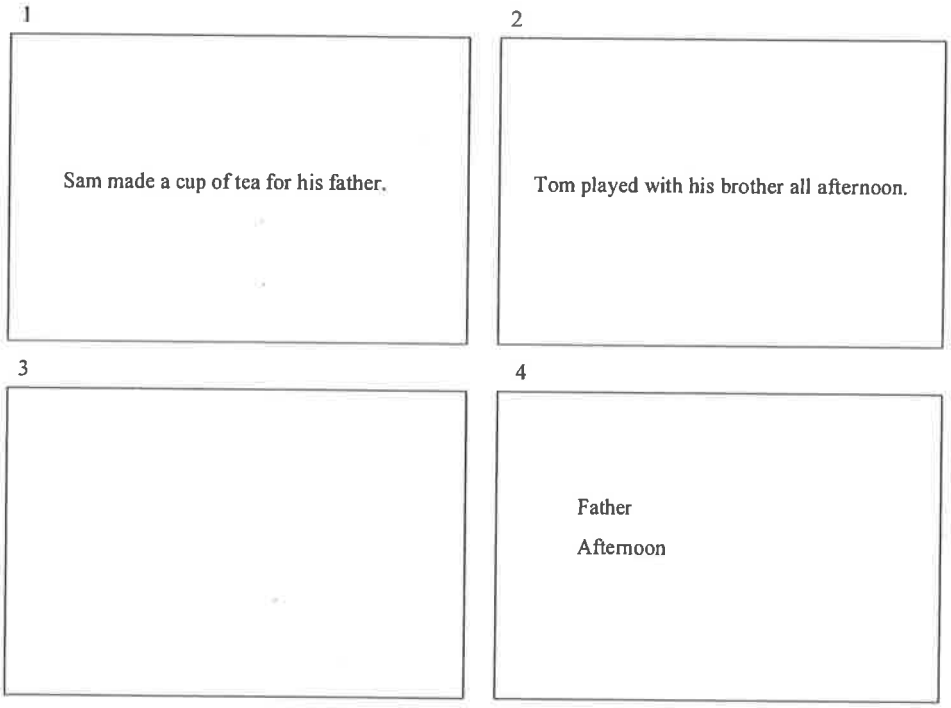


Figure 2 Daneman and Carpenter's (1980) Reading Span Test: Procedural steps in the reading sentences aloud version

Participants read a designated number of sentences aloud (steps 1 & 2). A blank card (step 3) is shown that prompts verbal, serial recall of the final word from each sentence seen prior to an earlier blank card (step 4).

The score achieved ('reading span') referred to the highest set of sentences read while correctly recalling the final word of each sentence. Reading Span is considered to reflect the capacity of WM that remains after meeting the demands for sentence-processing. The test was terminated when errors were made on all three trials of the

same sentence level.

2.6.1.3 Reading sentences silently

The administration process for this version of the Daneman and Carpenter's (1980) Reading Span Test was similar to that for reading sentences aloud except for one change. Participants also had to verify (via verbal response) whether the sentences they were reading were true or false prior to the recall of the final words. For example, the sentence may have been 'giraffes have long necks', in which case the answer is true. This additional step of sentence verification was included to ensure that participants were reading (processing) the entire sentence. Without this step, participants may have used the strategy of paying attention to the final word of the sentence only. By doing this, the task would amount to a simple storage task (i.e. word span). This particular strategy could not be employed in the reading sentences aloud scenario. Recall of the final word was verbal (as above) but the words recalled could be in any order. See Figure 3 (p.38) for an illustration of these steps.

2.6.1.4 Listening ('Listening Span')

Method for this Reading Span Test version was essentially the same as that for reading sentences silently. That is, sentence verification also occurred in this version, to ensure that sentences were being processed in their entirety. In this version participants heard sentences read aloud and instead of a blank card, a tone signalled the recall of final words.

2.6.2 Adaptations to the Reading Span Test

Daneman and Carpenter's (1980) Reading Span Test has been very influential (Lehto, 1996) and many tasks exist that have adapted the principles and format of the original test in some way. For example, operation span is a common transformation of this task. The format is similar to that of 'reading sentences silently' but operation span

uses numeric information (i.e. sums; 3 + 2 = 5) instead of word-based material (e.g. Kirasic et al., 1996; Turner & Engle, 1989). Verifying the sum is the processing aspect and remembering a target number in the sum is the storage or recall component. There are also several types of 'reading span' tests that combine sentences or words with other non-language stimuli. For example, Cantor and Engle (1993) combined the verification aspect of operation span with the recall of words; and Numminen, Service, Ahonen and Ruoppila (2001) utilised 'sentence-picture pairs' where the verification process refers to pictures shown at the same time as the sentences and recall refers to the last word of the sentence. Also, Tirre and Peña (1992) incorporated a general knowledge test into their version of the Reading Span Test via the sentence verification component. Unlike other studies, they also provided feedback on the sentence verification.

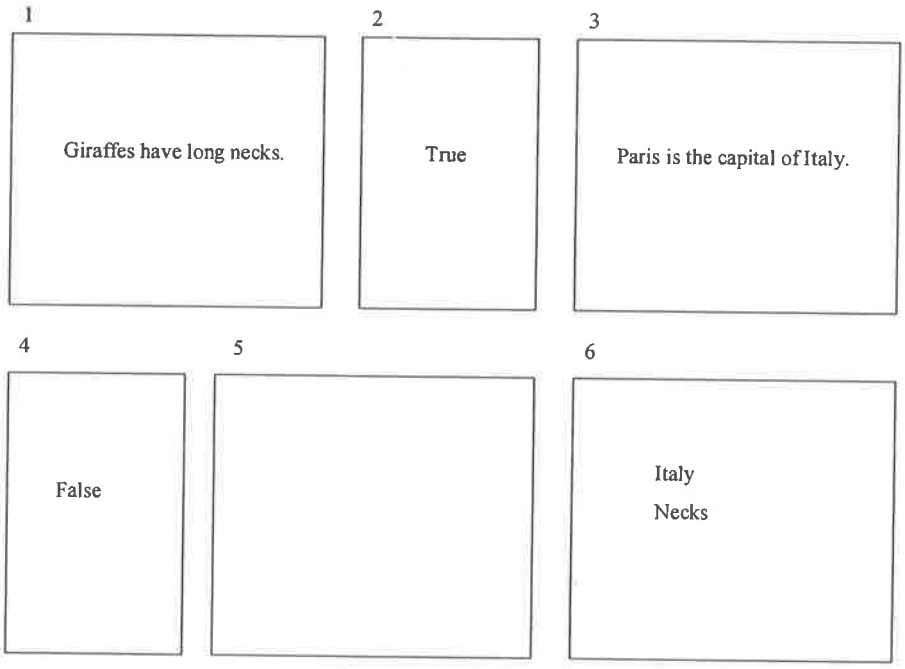


Figure 3 Daneman and Carpenter's (1980) Reading Span Test: Procedural steps in the reading sentences silently version

Participants read a designated number of sentences to themselves (steps 1 & 3). After each sentence, a verbal response is made as to whether the sentence is factual (steps 2 & 4) and then a blank card (step 5) is shown that prompts verbal recall of the final word from each sentence seen prior to an earlier blank card (step 6). Final word recall can be in any order, unlike the "aloud" version.

2.6.3 Analyses of how different versions of the Reading Span Test can vary

There are potentially many components of the Reading Span Test that can vary. Table 7 presents five main aspects that have varied from study-to-study. The degree of variation in these components was evaluated in an analysis of 15 studies. This analysis was primarily conducted in order to discover the dominant methodological trends for reading span-type tasks administered to participants over the age of 70 years (i.e. the target population for the current study). In addition to involving reading span-type tasks and a participant group over the age of 70 years, the other main inclusion criterion for studies incorporated in this analysis was sufficient methodological information/clarity of test procedures. The following Tables (8 through 13) and their accompanying descriptions summarise the results of this 15-study analysis. It should be noted that not all of the studies in the analysis provided information on every component investigated here.

Table 7
Administrative features that can vary across different versions of Daneman and Carpenter's (1980) Reading Span Test

Feature	Description
Sentences	This refers to aspects such as mode of presentation, number of words in and grammatical complexity of sentences, sentence-set length, and number of trials per sentence level
Use of questions or sentence verification	If this step is included, variation can be found in the mode of presentation and whether a time limit applies
Word recall	The word-to-be-recalled may be the final word or another target word in the sentence. The number of syllables, response mode and presence of a time limit can also vary
Stopping or discontinue rule	Refers to the number of errors made prior to test cessation
Scoring procedures	Three types of scoring techniques have been used: traditional span, number correct and absolute span

In regard to 'sentences', Table 8 (p.41) illustrates that previous studies incorporating participants over the age of 70 have most commonly used computer administration for Reading Span Test sentences. The time allocated for participants to read sentences was self-paced for all studies included in this analysis. That is, once the participant had finished reading one sentence, the next sentence appeared. This method takes into consideration individual differences in reading speed. The majority of studies employed sentences within a range of 6 to 16 words. Associated with sentence length is another variable, grammatical complexity. Longer sentences tend to be more difficult and allow for more complex grammatical elements such as plurals, prepositions and inflections (e.g. La Pointe & Engle, 1990; Lehto, 1996). However, many studies focusing on working memory ability or performance have not manipulated grammatical complexity; the aim is to measure working memory capacity rather than reading ability, intelligence or other cognitive processes. The range of sentence-set sizes (i.e. the minimum and maximum number of sentences seen prior to recall) used in various reading span-type tasks has not varied greatly (see Table 9, p.41). Similarly, there does not appear to be any particular preference for sentence-set size among studies involving people over the age of 70.

Most studies have employed a standard set format where the number of sentences viewed before recall increases by one after a given number of trials. An exception to this was a study by Lustig, May and Hasher (2001). In addition to the standard set format, they included a descending set format, where participants started with the largest sentence-set and, after a given number of trials, the set decreased by one sentence. By varying the number of sentences presented, the demands or load on memory also varies (Craik, Morris & Gick, 1990). The larger the sentence-set size, the greater the memory load and the demands on the working memory system.

Table 8
Sentences: Different modes of presentation for versions of the Reading Span Test

Paper and pencil (cards)	Arnett et al. (1999) Daneman & Carpenter (1980) Kail & Hall (2001) Light & Anderson (1985)* MacDonald, Almor, Henderson, Kempler & Anderson (2001)
Computer (screen)	Cantor et al. (1991) Fisk & Warr (1996)* Kirasic et al. (1996)* La Pointe & Engle (1990) McCabe & Hartman (2003)* Oberauer et al. (2000) Park et al. (1996)*
Verbal/ Oral	Park et al. (2002)* Salthouse & Babcock (1991)*

*studies that included participants over the age of 70 years

Finally, the majority of studies have provided three trials at each sentence-set length. Exceptions to this have included Kail and Hall (2001; two trials per sentence-set length), Roberts and Gibson (2002; five trials per sentence-set length), and Cantor et al. (1991; varied number of trials from two to seven so that it was not predictable). However, none of these exceptions involved participants over the age of 70 years.

Table 9
Sentences: Variability of set length (number of sentences seen prior to recall) across different versions of the Reading Span Test

Minimum 1; Maximum 5	Fisk & Warr (1996)* Kail & Hall (2001)
Minimum 2; Maximum 5	Kirasic et al. (1996)* La Pointe & Engle (1990) Light & Anderson (1985)* MacDonald et al. (2001) Roberts & Gibson (2002)
Minimum 2; Maximum 6	Arnett et al. (1999) Daneman & Carpenter (1980) McCabe & Hartman (2003)*

*studies that included participants over the age of 70 years

Another aspect of the Reading Span Test that can vary across studies is the processing part of the task. Sentence processing is assessed by way of questions or sentence verification and is required when sentences are not read aloud to ensure processing of the sentence(s) is taking place. In terms of the mode of presentation for

sentence verification, Table 10 demonstrates that paper-and-pencil and computer formats have been employed equally across studies, including those involving individuals over the age of 70 years. Generally, a time limit for responding to sentence verification (or questions relating to reading the sentence) has not applied. Exceptions to this have been Daneman and Carpenter (1980; who used a 1.5 second limit) and Oberauer et al. (2000; who used a 4 second limit). Again, neither of the exceptions applied to studies using similar age groups to the current study.

Table 10
Sentence verification: Different modes of presentation on versions of the Reading Span Test

Paper and pencil (card/booklet)	Daneman & Carpenter (1980) Fisk & Warr (1996)* Kail & Hall (2001) Salthouse & Babcock (1991)*
Computer (screen)	McCabe & Hartman (2003)* Oberauer et al. (2000) Park et al. (2002)* Roberts & Gibson (2002)

*studies that included participants over the age of 70 years

The overwhelming majority of studies sampled here employed recall of the final word of the sentence rather than some other target word (see Table 11, p.43). This also applies to studies involving people over the age of 70 years. In regard to the number of syllables that the word-to-be-recalled contains, most studies have employed one-syllable words (e.g. Arnett et al., 1999; La Pointe & Engle, 1990). Some studies have involved recall words with a range of syllables. Light and Anderson (1985) and McCabe and Hartman (2003) used words with one to three syllables; Salthouse and Babcock (1991) used words with either two syllables or one; and Oberauer et al. (2000) used words with three or fewer syllables. In other words, elderly people have been exposed to final words of more than one syllable.

Table 12 (p.43) shows that both verbal and written response mode for word recall have been used fairly evenly across all studies sampled, as well as across studies

involving participants over the age of 70 years. Regardless of response mode for word recall, the majority of studies have stipulated that recall must be given in correct serial order. The only exception to this was the study by La Pointe and Engle (1990) and some versions by Daneman and Carpenter (1980). Many of the studies did not mention whether participants were under a time limit when recalling words. A 4 second limit per final word was nominated by Salthouse and Babcock (1991).

Table 11
Frequency of final word versus target word recall on different versions of the Reading Span Test

Recall Final Word of Sentence	Daneman & Carpenter (1980)
	Fisk & Warr (1996)*
	Kail & Hall (2001)
	Kirasic et al. (1996)*
	Light & Anderson (1985)*
	MacDonald et al. (2001)
	McCabe & Hartman (2003)*
	Oberauer et al. (2000)
	Park et al. (1996)*
	Park et al. (2002)*
	Roberts & Gibson (2002)
	Salthouse & Babcock (1991)*
Recall Target Word	Arnett et al. (1999)
	Cantor et al. (1991)
	La Pointe & Engle (1990)

*studies that included participants over the age of 70 years

Table 12
Frequency of different response modes for word recall on versions of the Reading Span Test

Verbal	Arnett et al. (1999)
	Daneman & Carpenter (1980)
	Kail & Hall (2001)
	Kirasic et al. (1996)*
	Light & Anderson (1985)*
	MacDonald et al. (2001)
	McCabe & Hartman (2003)*
Written	Fisk & Warr (1996)*
	La Pointe & Engle (1990)
	Oberauer et al. (2000)
	Park et al. (1996)*
	Park et al. (2002)*
	Salthouse & Babcock (1991)*

*studies that included participants over the age of 70 years

In terms of a stopping or discontinue rule, a popular rule (including for elderly age groups) has been for participants to continue until two or more errors in word recall are made, typically out of three trials (e.g. Light & Anderson, 1985; McCabe & Hartman, 2003; Park et al., 2002). Other stopping rules have been similar; the task ends when errors are made in two consecutive trials (e.g. Arnett et al., 1999) or when recall errors are made in more than half of the trials (e.g. Roberts & Gibson, 2002). Daneman and Carpenter's (1980) stopping rule required word recall errors to be made on all trials at the same sentence-set length and La Pointe and Engle (1990) did not use a stopping rule; completing the maximum sentence-set length was the end of the task.

The final main aspect where differences have occurred among reading span-type tasks is in scoring procedures. There have been three popular methods for scoring these tasks. 'Traditional Span' refers to the level or sentence-set length at which a participant was correct on two out of three trials (e.g. Daneman & Carpenter, 1980). 'Number Correct' refers to the total number of words correctly recalled (usually in serial order; e.g. Kirasic et al., 1996). 'Absolute Span', otherwise known as the 'weighted method', refers to the total of correctly recalled words (in order), for those trials that were perfectly recalled (e.g. Chiappe, Hasher & Siegel, 2000; May, Hasher & Kane, 1999; Rosen & Engle, 1998). Slight variations of these measures have also been used. For example, Kail and Hall (2001) combined 'traditional span' and 'absolute span' to produce a measure of the longest sentence-set length where all trials were correct. Table 13 (p.45) illustrates that 'traditional span' has been the most preferred method (including for participants over the age of 70 years), followed by 'absolute span' and 'number correct'.

Some authors have also suggested awarding 'half-marks' (e.g. Daneman & Carpenter, 1980; La Pointe & Engle, 1990; McCabe & Hartman, 2003). For example, if

a participant correctly answered the final words from two out of three trials involving a set of four sentences, then the traditional span score is four. However, if a participant correctly answered the final word from only one of the three trials involving a set of four sentences, then the traditional span score is three and a half. Alternatively, Roberts and Gibson (2002) awarded an additional 0.2 marks for each trial participants recalled correctly at the next sentence-set size.

Table 13
Frequency of different scoring methods used on versions of the Reading Span Test

Traditional Span	Daneman & Carpenter (1980) Fisk & Warr (1996)* Kirasic et al. (1996)* La Pointe & Engle (1990) Light & Anderson (1985)* McCabe & Hartman (2003)* Roberts & Gibson (2002) Salthouse & Babcock (1991)*
Number Correct (items)	Kirasic et al. (1996)* LaPointe & Engle (1990) Oberauer et al. (2000)
Absolute Span	Cantor et al. (1991) La Pointe & Engle (1990) McCabe & Hartman (2003)* Park et al. (1996)* Park et al. (2002)*

*studies that included participants over the age of 70 years

It is worth noting that several studies have used more than one scoring method (e.g. La Pointe & Engle, 1990; McCabe & Hartman, 2003; Turner & Engle, 1989). These studies established that analyses employing different working memory scoring methods produce similar results; the study by McCabe and Hartman (2003) involved elderly participants (i.e. individuals over 70 years of age). More recently, Friedman and Miyake (2005) clarified this view by stating, "It may be the case that different scoring methods will result in qualitatively similar patterns of results, but it is likely that these results will be clearer when more continuous measures are used" (p.588). Continuous scoring methods, such as number correct, were endorsed previously by Friedman and

her colleagues (e.g. Miyake, Emerson & Friedman, 1999) because they considered these measures to maximise variance and capture subtle differences among individuals.

A final point regarding the variability in Reading Span Test procedures refers to both the stopping rule and scoring. Both of these aspects are complicated by reading span-type tasks that include the sentence verification process. For these studies, sometimes the verification response has determined whether the trial was considered correct. For example, in many studies (e.g. Fisk & Warr, 1996; Park et al., 1996; Park et al., 2002; Salthouse & Babcock, 1991), if all the final words were correctly recalled but some part of the sentence verification was incorrect, then marks were only given for those trials where both components were correct. By contrast, in other studies (e.g. Daneman & Carpenter, 1980), scoring has not been affected by sentence verification accuracy. Finally, like most tasks, practice has typically been provided for versions of the Reading Span Test. La Pointe and Engle (1990) appeared to be the only exception to this.

2.6.4 What do reading span-type tasks measure?

The various versions of the Reading Span Test have commonly been used as measures of working memory capacity (Arnett et al., 1999; Just & Carpenter, 1992; La Pointe & Engle, 1990; Salthouse, 1994), and more specifically as measures of verbal working memory (e.g. Haavisto & Lehto, 2004; Kirasic et al., 1996; Meguro et al., 2000; Wareing et al., 2005; Waters & Caplan, 2003). However, the Reading Span Test is a complex task (Lustig et al., 2001; Meguro et al.; McCabe & Hartman, 2003) that draws upon many aspects of the working memory system and human cognition. According to Baddeley (1990b), the task is likely to involve strategy selection, the phonological loop, knowledge of vocabulary and coordination of these aspects. Considering that strategy selection draws upon the central executive (Fisk & Warr,

1996), the Reading Span Test has been acknowledged to measure also the operation of the central executive, rather than simply the phonological loop (Gathercole & Baddeley, 1993, Just & Carpenter; Lehto, 1996). Some authors have argued that reading span-type tasks are a measure of central executive functioning or attentional resources more than anything else (see Whitney et al., 2001 for a discussion on this).

2.6.5 Performance on reading span tests by age group

As discussed in Chapter 1, "it is generally assumed that there are at least moderate declines in working memory capacity with age" (Waters & Caplan, 2003, p.553). That is, younger adults show superior performance on tasks of working memory compared to older adults[3]. The following studies provide examples of differences found in performance on reading span-type tasks across different age groups. The examples are presented in accordance with the scoring methods used (i.e. grouped into studies that have employed traditional span, number correct or absolute span; refer to p.44 for a description of these terms).

Light and Anderson (1985) reported a traditional sentence span mean score of 3.08 for older individuals (age range = 56 - 80 years; mean = 68.7 years) and a mean of 3.60 for younger individuals (age range = 21 - 34 years; mean = 25.8 years). This difference was significant, with a moderate - large effect size. The results of Meguro et al. (2000) also reflected a similar distinction. Their study employed a Japanese version of reading span and compared performance across three age groups (see Table 14, p.48, for details). One-way analysis of variance showed a main effect for age. Post hoc comparisons indicated that younger adults (age range = 20 - 39 years) performed significantly better than both middle-aged adults ($p < .01$) and elderly ($p < .01$); and

[3] Not all studies have shown evidence of this trend, and the extent of differences in performance across these age groups also varies. This variability among results is perhaps due to the wide variety of working memory tests employed across studies and their respective reliabilities (Waters & Caplan, 2003). These issues will be discussed in more detail in Chapter 4.

middle-aged adults (age range = 40 - 59 years) also performed significantly better than elderly adults (age range = 60 - 82 years; $p < .05$).

Table 14
Results from Meguro et al. (2000): Age differences on a reading span test

Age group	Age (years)		Traditional Span score	
	Mean	SD	Mean	SD
Young adults	28.8	6.07	3.50	1.7
Middle-aged adults	49.4	5.68	2.53	1.50
Elderly adults	68.3	6.47	2.02	1.00

In terms of the likely range of performance shown by elderly and young adults on reading span-type tasks, Salthouse and Babcock (1991) demonstrated that the majority of older participants (those in their 70s) achieved traditional span scores between 0 and 2 whereas younger adults in their 20s or 30s achieved span scores between 2 and 4.

Van der Linden et al. (1999) demonstrated that, in a French version of the reading span test, younger adults (age range = 30 - 39 years; mean = 33.3 years) showed superior performance (mean = 31.63 words, SD = 12.28), or recalled more correct final words (number correct), compared to older adults (age range = 70 - 80 years, mean = 72.9 years), with a mean of 21.27 words (SD = 9.32). Age was significantly and negatively associated with reading span performance ($r = -.34$, $p < .01$, $N = 151$).

Lustig et al. (1999) employed the absolute span scoring method for their reading span task and found that younger adults (age range = 18 - 24 years; mean = 19.1 years) performed better (mean = 26.20 words, SD = 8.86) than older adults (age range = 61 - 75 years; mean = 68.9 years), whose mean performance was 20.10 words (SD = 8.07). This difference was significant and of moderate effect size. Using the same scoring method, McCabe and Hartman (2003) also reported superior performance by younger adults (mean = 20.1 years, SD = 2.4) on a composite measure of reading span-type tasks (mean = 13.13 words, SD = 8.31). Older adults (mean = 72.3 years, SD = 5.90)

performed significantly lower with a mean of 8.88 words (SD = 5.82).

2.7 Dot Matrix: An overview

Dot Matrix is the second of three working memory tasks employed in the current study. This task is based on Law, Morrin and Pellegrino's (1995) matrix task involving both visuo-spatial storage (remembering the location of a dot in a matrix grid) and processing (verifying an addition equation comprising lines connecting dots in a matrix). This task is representative of Baddeley and Hitch's (1974) visuo-spatial sketchpad. Although the Dot Matrix task is not the most commonly used visuo-spatial working memory task (see Fischer, 2001 for a discussion on the use of the Corsi Blocks), "matrices containing items in their cells have been frequently used as stimuli in studies of memory of spatial information" (Ichikawa, 1981, p.69). The importance of including a task from the visuo-spatial domain was not only to improve understanding of this working memory component, but also to ascertain whether there is in fact differential decline with age on tasks from different content domains. That is, age-related effects in some domains may be greater than those in others (e.g. Jenkins, Myerson, Joerding & Hale, 2000; Myerson et al., 2003; Salthouse, 1995). For example, age-related deficits are thought to be greater on tasks measuring spatial information compared to tasks involving verbal information (Jenkins et al., 2000). The following section details the Dot Matrix and similar tasks.

2.7.1 Development of the Dot Matrix Task

Ichikawa (1979[4]; cited by Ichikawa, 1981) was among the first to use the presence of dots in a matrix to test short-term visual memory. He showed that the proportion of dots correctly recalled decreased as the number of dots to remember increased. Ichikawa suggested that such a memory test would be suitable for use in

[4] Primary source is unavailable because it is a Master's thesis written in Japanese (University of Tokyo)

studies of individual differences. Ichikawa (1983) later proposed that "if the visual span expresses the capacity of visuo-spatial representation rather than that of verbal immediate memory, then the correlation between visual span and digit span should be considerably weak (sic)" (pp.173-174). This is exactly what he found, providing evidence for two independent memory mechanisms. Moreover, Ichikawa (1983) concluded that "visual span is one of the most fundamental measures for assessing visuo-spatial ability" (p.179).

Over the period from the early 1980s to the mid-1990s, the "dot matrix task" has become known as a "simultaneous storage and processing" task, in line with Baddeley and Hitch's (1974) principles of working memory. Law et al. (1995) employed a task using the recall of shaded cells in a matrix (akin to Ichikawa's use of dots), in addition to using a verification or processing step. They instructed participants to:

> ...memorise three to five individually presented 3 x 3 matrices that had one of nine squares shaded....Between individual matrix presentations, subjects verified an addition equation in which two line matrices were to be added together to form a third line matrix (e.g. Matrix 1 + Matrix 2 = Matrix 3). Each line matrix consisted of a 3 x 3 dot matrix with a line connecting the dots in a variety of configurations (pp.213-215).

This task was administered on computer and participants had 8 seconds to verify each equation. After the sets of matrices and verification equations were completed, a 3 x 3 matrix appeared with each cell containing a number (i.e. digits 1 to 9). Recall of the shaded cells presented at the beginning of the task was done by pressing the number on the keyboard that corresponded to the location of the shaded cell(s). The position of the numbers (1 – 9) in the cells varied across trials to avoid people using a number recall strategy. This procedure is illustrated in the article by Law et al. and a pictorial

representation of the version employed in the current study is provided in Chapter 3.

Several years later, Miyake et al. (2001) employed a very similar task to Law et al.'s (1995). This more recent task differed in the following ways: the to-be-recalled matrices contained dots instead of shaded cells; the equation processing procedure was the same except that participants were only given 4.5 seconds to verify the equation (rather than 8 seconds); a 5 x 5 matrix was used instead of a 3 x 3 matrix; and the range of set sizes was marginally bigger (i.e. two to five instead of three to five).

2.7.2 Performance on visuo-spatial tasks by age group

Compared to "reading span" tasks, the possible names that visuo-spatial working memory tasks can take is extremely variable (e.g. block or cube tasks, tracking tasks, integration tasks, rotation tasks and imagery tasks). This is further complicated by the evidence that these tasks may measure visual working memory and spatial working memory separately or to different degrees. Overall, these issues make identifying studies that have investigated age differences and visuo-spatial working memory more difficult. Moreover, readily identifiable studies incorporating visuo-spatial working memory tasks (i.e. studies that contain the key term 'visuo-spatial working memory' somewhere in the article), do not often focus on age differences, or report mean performance across different age groups (e.g. Fischer, 2001; Kirasic et al., 1996; Law et al., 1995; Miyake et al., 2001; Oberaurer et al, 2000).

One exception to the above is a study by Vecchi and Cornoldi (1999). Their study incorporated three age groups: young adults (age range = 19 - 29 years; mean = 22 years), young elderly (age range = 60 - 70 years; mean = 66 years) and old elderly (age range = 72 - 84 years; mean = 76 years). Participants completed three visuo-spatial working memory (or 'active manipulation') tasks; two of these involved matrix stimuli (i.e. similar to Dot Matrix) and performance on these tasks only is reported here. For

both tasks, there were significant differences ($p < .01$) in mean traditional span scores across all three age groups, established via one-way analysis of variance with post hoc comparisons. Vecchi and Cornoldi's results are shown in Table 15 and support the trend for superior working memory performance by younger age groups, as demonstrated earlier for reading span-type tasks.

Table 15
Results from Vecchi and Cornoldi (1999): Age differences on visuo-spatial working memory tasks (traditional span scores)

Age group	Task 1		Task 2	
	Mean	SD	Mean	SD
Young adults	4.65	1.05	9.81	2.49
Young elderly	3.23	.86	6.72	3.45
Old elderly	2.01	1.03	4.54	2.10

2.8 Digit Ordering: An overview

The third working memory task used in the current study, Digit Ordering, was first developed by Cooper, Sagar, Jordan, Harvey, and Sullivan (1991). It requires the ordering of a random series of digits so that they are in ascending, numerical order. The Digit Ordering Task (DOT) has been reported to measure the manipulation or processing aspect of working memory (Camicioli, Lea, Nutt, Sexton & Oken, 2001; Müller, Werheid, Hammerstein, Jungmann & Becker, 2005) or more generally, executive function (Cooper et al.; Hoppe, Müller, Werheid, Thöne, von Cramon, 2000). It has principally been employed in studies involving clinical populations, in particular, those with Parkinson's Disease (Camicioli et al.; Cooper et al.; Hoppe et al.). However, it should be noted that, within general population studies, the DOT has not been commonly used. Given that the current study had the potential of including a clinical group (Alzheimer's Disease patients), this task was considered highly appropriate (see p.55 for further discussion of the reasoning behind this)[5]. Furthermore, the inclusion of

[5] Discussions were held with colleagues at the Royal Adelaide Hospital regarding a study involving this clinical group. However, due to limited access to participants and the timeframe of the current study, this did not eventuate.

this task would allow for rarely seen comparisons of performance in a 'normal' or non-clinical elderly population, as well as comparisons with other tests of working memory. Inclusion of the DOT in the current study also serves to provide a range of task stimuli; in this case numeric stimuli, compared with words (Reading Span) and visuo-spatial stimuli (Dot Matrix). The form of this task has changed somewhat over the last several decades, starting out as a pen and paper task and, more recently, taking on a similar format to the widely used Wechsler Digit Span tasks (Lezak, 1995; Wechsler, 1945, 1987). A brief review of the development and use of the DOT follows in the next section.

2.8.1 Development of the Digit Ordering Task

Corkin (1968) was the first to develop a 'serial ordering of digits' task, albeit that the format of this task more closely approximated a speed of processing task rather than a task of working memory. Her study compared the rate of motor learning in a man with severe amnestic syndrome to that of normal males (matched for age and dominant hand). In this study, the digit task was "thought to be sensitive to personal tempo" (Corkin, p.256). It was a timed task (15 minutes) where participants were given groups of seven digits on paper. They were asked to re-write the groups of digits so that the digits were in order from smallest to highest value. Throughout the task, a light flashed at 60 second intervals. This was a signal for participants to mark a line on the page indicating how much they had completed since the last flash. This procedure permitted calculation of the number of lines (groups of digits) completed in each 5 minute period.

Cooper et al.'s (1991) study, involving newly diagnosed and never-treated Parkinson's Disease patients, was the first to modify Corkin's (1968) task. The Digit Ordering Task (DOT) was now a verbal task (DOT-V) where participants "were read a random selection of seven digits and were required to reorder the items by memory and

repeat them back in ascending fashion" (Cooper et al., p. 2100). The task was no longer timed and "1 point was awarded for each digit placed in its correct position until a response broke the ascending sequence" (Cooper et al., p. 2100). The entire test consisted of 15 sets of digit sequences, with a presentation rate of seven digits in 5 seconds, a speed akin to natural speech (Hoppe et al., 2000). Thus the format of DOT-V is very different to Corkin's (1968) task. However, there were several omissions in Cooper et al.'s reporting of DOT-V methods and procedures, such as the range and frequency of digits used. Hoppe et al.'s study investigated some of these aspects by conducting important analysis of the psychometric properties of DOT-V.

Hoppe et al. (2000) replicated Cooper et al.'s (1991) results and investigated a new scoring procedure and different rates of digit presentation. However, more relevant here is their finding that DOT-V was structurally unbalanced and too difficult for their participants. For example, where digits were given twice in a seven-digit string (e.g item = 4735856; solution = 3455678), the repeated digits were more likely to be of a higher numerical value; and placed at the end of the solution series (i.e more likely to have two '7s' in an item than two '2s'). Moreover, despite items being of constant length (i.e. seven digits), item difficulty was quite variable. Items with consecutive digits and no repeated digits (e.g item = 3512674; solution = 1234567) were easiest for participants to perform, but such items were infrequent. Hence, "80% of all items required negative feedback" to be given to individuals (Hoppe et al., p.43), and low levels of motivation were a concern on this task.

In order to investigate the effect of presentation rate, Hoppe et al. developed a new experimental version of DOT-V. They referred to this as DOT-EXP. The procedure for DOT-EXP was similar to the Wechsler Digit Span tasks. The final version of this task involved re-ordering three digits initially, increasing by one digit after every two

trials at the same length, to a maximum of eight digits. Only one of the two trials contained a repeated digit. DOT-EXP was considered to be structurally balanced and the standard rate of presentation was also that of natural speech (seven digits in 5 seconds).

MacDonald et al. (2001) employed a slightly different version of the DOT-V. This time, instead of a DOT analogous to Wechsler's Digit Span (i.e. DOT-EXP), MacDonald et al. proposed a DOT analogous to Reading Span (termed here as DOT-R). MacDonald et al. investigated language processing and comprehension in Alzheimer's Disease patients and normal adults. The main impetus for employing DOT-V was the poor performance of Alzheimer's Disease patients on more commonly used tests of working memory, such as the Reading Span Test (Daneman & Carpenter, 1980). According to MacDonald et al., "patients do poorly on this task [reading span] simply because they have tremendous difficulty in understanding and remembering instructions to do two things at once....patients forget one or another part of the task" (pp.18-19). A more reliable task without such floor effects would be desirable. They suggested that DOT-V would be a good solution to this problem because DOT-V maintained the fundamental features of Reading Span (that is, storage and processing components) but lacked the complex instructions that appeared to give Alzheimer's Disease patients so much trouble. Therefore, in order to compare performance on both Reading Span and the DOT-V in Alzheimer's Disease patients, MacDonald et al. used the following DOT-R method:

> Digit ordering trials were developed using the digits 1 - 9. Difficulty ranged from two to six digits to order, with four trials at each level. The digits were sampled equally often across all levels of difficulty....Digit span testing proceeded through increasing levels of difficulty, similar to the reading span

task. Digits in each trial were presented auditorily at the rate of about one digit per second. Participants were instructed to report the digits in ascending numerical order. At each testing level, testing was stopped if the participant failed to sequence the digits in two or more trials correctly (pp.21-22).

The most recent modification of DOT-V was proposed by Werheid, et al. (2002). The Adaptive Digit Ordering Test (DOT-A) "was developed to provide an economic measure for assessment of working memory with maximal comparability to the Wechsler Digit Spans" (Werheid et al., p.550). It involves ordering, in ascending numerical order, three to eight digits with two trials at each sequence length (thus the entire test consists of 12 sets of digit sequences). One of the two trials contains a repeated digit. The task is stopped if errors are made in both trials at the same sequence level (as for Wechsler Digit Span instructions). Unlike Hoppe et al. (2000; seven digits in 5 seconds), presentation rate is one digit per second. The most recent study involving DOT was conducted by Müller et al. (2005). In that study DOT-A was administered to individuals with schizophrenia.

2.8.2 Performance on digit ordering tasks by age group

Although digit ordering tasks have successfully been shown to differentiate between various clinical and non-clinical populations (e.g. Hoppe et al., 2000; MacDonald et al., 2001; Werheid et al., 2002), no studies have been found that directly investigated age differences in normal, healthy populations. There are only limited data on digit ordering performance in young and older adults. However, there is some indication that performance on digit ordering tasks across age groups follows the age-related patterns demonstrated by other working memory tasks (i.e. Reading Span and Dot Matrix type tasks). This evidence primarily comes from the results of MacDonald et al. (2001).

The main purpose of MacDonald et al.'s (2001) study was to demonstrate that Digit Ordering is capable of distinguishing between dementia classifications and dementia progression. As aforementioned, Digit Ordering was successfully used to achieve this goal. In the process of their investigations (this particular study involved four experiments), data were also generated that supported the conclusion that younger adults are superior to older adults for Digit Ordering performance (see Table 16).

Table 16
Results from MacDonald et al. (2001): Differences on a digit ordering task by age and cognitive status groups

Age group	Age (years)		Number Correct	
	Mean	SD	Mean	SD
Young adults#	-	-	18.10^	1.59
Healthy elderly	77.4	7.71	8.31	1.30
Questionable dementia	78.6	8.21	7.44	1.55
Mild dementia	78.9	6.95	4.22	2.86
Moderate dementia	79.6	5.26	3.00	2.87

#Specific age details (e.g. mean age, age range) were not available; the only information provided was that participants were university students

^ Maximum number correct was 20 whereas the maximum number correct for the remaining scores was 10. Therefore, a comparable mean score for young adults could be approximated to 9.05 correct digit sequences. The maximum score of 10 referred to the number of correct *trials*; this equates to ordering six digits correctly (i.e. digit sequence length ranged from two to six digits and there were two trials at each sequence length)

2.9 Summary of working memory test selection

Generally, the major goal of task selection in the current study was to operationalise each component of Baddeley & Hitch's (1974) tripartite model; that is, to include tests that primarily measure phonological loop capacity or verbal working memory (e.g. Reading Span); visuo-spatial sketchpad capacity or visuo-spatial working memory (e.g. Dot Matrix); and central executive or more attentional/processing functions (e.g. Digit Ordering). Achieving this would permit the determination of whether different aspects of the working memory construct are more or less vulnerable to the ageing process - and thus which components may have more utility in biomarker research (see Chapter 1 for a discussion of biomarkers). Moreover, the inclusion of tasks that use different types of materials or content domains [i.e. words/verbal

(Reading Span), numbers/quantitative (Digit Ordering) and spatial (Dot Matrix) stimuli], should also provide the opportunity to analyse the differential decline hypothesis (as considered earlier on p.49).

However, it is accepted that task selection based mainly on content domains or model components may confound the influence of storage and processing aspects of working memory. That is, it will be difficult to ascertain whether age effects (and other relationships involving working memory) are domain or function related. Other researchers have acknowledged that storage and processing aspects of memory are difficult to separate (e.g. Atkinson & Shiffrin, 1968). Consequently, the selection of tasks based on domain seemed to provide the most appropriate course of action. However, it was planned, where possible, to attempt to separate storage and processing components of working memory by statistical manipulation (this is explained further in Chapter 4).

To assess the utility of the three working memory tests (and hence the construct of working memory) within the current context where the aim was to test the adequacy of the tests as biomarkers, various other types of measures (i.e. apart from cognitive) were included in the current study. That is, measures of psychological well-being, health, lifestyle and physical condition were also included to ascertain the optimal variables that might explain independent functioning of elderly persons. Chapter 3 will describe all of the measures included in the current study, as well as other relevant aspects of method.

CHAPTER 3: STUDY DESIGN AND METHOD

INITIAL, INTERMEDIATE AND FINAL MEASUREMENT BATTERIES

3.1 Chapter overview

This Chapter provides details about the study design, participants and the questionnaires and tests employed in the current study. The study design was longitudinal; and the rationale behind this selection will be discussed, in addition to the qualities that make the specific design employed advantageous. The attributes of participants and how they were recruited are then discussed. The last main topic presented in this Chapter is the test battery. In addition to describing the tests and questionnaires administered throughout the study, the importance and overall relevance of each measure in the test battery is considered, and the psychometric properties of the tests are described.

3.2 Study design

As defined in Chapter 1, a biomarker is "a physical or laboratory measure with diagnostic or prognostic usefulness" (Pollock, 2002, p.644). Therefore, in order for a task assessing the construct of working memory to be considered a biomarker of independent functioning in an elderly population, it must first be shown to demonstrate prognostic utility[6]. To determine this, the study design was longitudinal, meaning that the same sample of individuals was assessed at more than one measurement occasion (Anstey & Hofer, 2004); on three occasions over a period of 19 months. Details of time intervals between measurement occasions are shown in Figure 4 (p.60). This longitudinal design has distinct methodological advantages over both cross-sectional studies and other longitudinal methods, as discussed in the following section.

[6] There are additional criteria surrounding whether a measure can be classified as a biomarker and these will be discussed in Chapter 7.

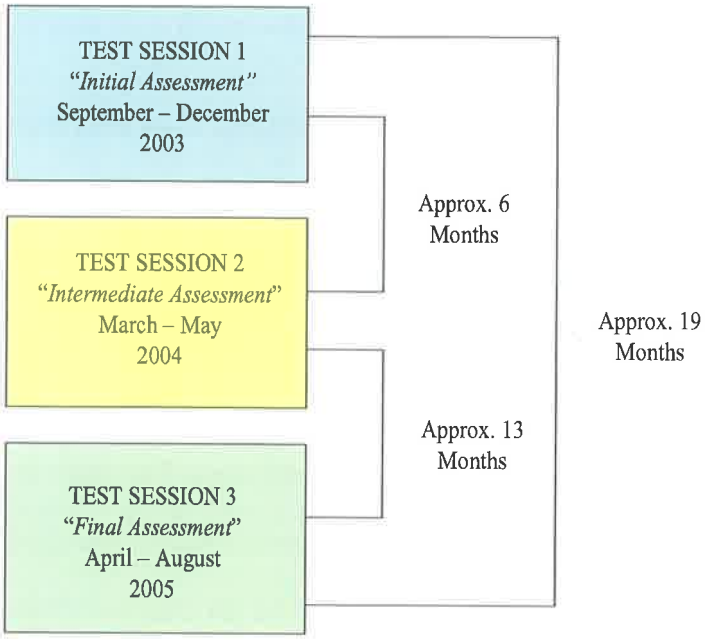


Figure 4 Timeline for longitudinal data collection in the current study

3.2.1 General qualities of longitudinal studies

"The prime advantage of a longitudinal study is its effectiveness for studying change" (Diggle, Liang & Zeger, 1994, p.17). More specifically, longitudinal studies are useful when investigating the following aspects of change, amount, nature (i.e. linear vs. quadratic), determinants (i.e. what causes change to happen) and consequences (Collins, 1996). Longitudinal studies are superior to cross-sectional studies, which involve one measurement occasion only, because the former can establish the predictive validity of measures.

It is important to bear in mind that not all longitudinal studies are created equal. Longitudinal studies involving two measurement occasions, commonly referred to as 'follow-up' studies, have several limitations compared to 'multi-wave' studies (more than two measurement occasions). A main limitation to follow-up studies is that they show only linear developmental trajectories. In summary, "two time points provide an inadequate basis for studying change" (Bryk & Raudenbush, 1987, p.147). This view is

common among researchers (e.g. Alder, Adam & Arenberg, 1990). Therefore, additional measurement occasions potentially permit a more accurate picture of change (Zimprich & Martin, 2002). For example, three or more measurement occasions can show whether developmental trajectories are nonlinear and they provide information about fluctuations between the first and last measurement occasions (Anstey & Hofer, 2004; Willett, 1988). Thus, including three assessment occasions in this dissertation has permitted investigation of the rate at which elderly individuals change on various measures.

In addition to the specific advantages of the longitudinal design employed here, the current longitudinal study was planned to contribute more broadly to the field of cognitive ageing because:

(1) Generally, there have been fewer longitudinal studies than cross-sectional studies (e.g. Willis et al., 1992). This is probably due to the greater resource requirements of the former, by way of cost and time;

(2) Longitudinal data on working memory are even more limited (e.g. Hultsch, Hertzog, Small, McDonald-Miszczak & Dixon, 1992), particularly in older populations; and

(3) Conducting longitudinal studies involving assessment periods over short time intervals (i.e. < 2 years) have been overlooked in favour of longer time intervals (e.g. 10+years). However, longitudinal studies with shorter time intervals can demonstrate reliable changes (Hertzog, Dixon & Hultsch, 1992; Hultsch et al., 1992) and be more sensitive to influence from subtle changes.

3.3 Participants

People 70 years of age and older and living in the community (within the Adelaide metropolitan area) were recruited following exposure of the study's aims

through various media. These included local newspapers (The Messenger, The Adelaidean), radio programs (ABC Radio, Adelaide University Student Radio) and a television news report (Channel 7 Nightly News). Initially, 150 participants were recruited, with ages ranging from 70 to 91 (mean = 77.59 years, SD = 4.39). The overall sex ratio, approximately 2:1 in favour of females (99 females and 51 males), was higher than but still reasonably representative of the State demographics for elderly persons. That is, in South Australia in 2003 there were almost 30% more females than males over the age of 65 and this percentage was higher (more than 50% more females) among elderly people over the age of 85. This outcome reflected the greater life expectancy of females (Australian Bureau of Statistics, 2004). Table 17 details this same information for the current sample, where it can be seen that the current sample did have a higher proportion of females than the state as a whole.

Table 17
The numbers of each sex in the current sample across age groups

	Females	Males	Percentage more females
70 – 85 years	91	49	53.9
Over 85 years	8	2	75.0

Figure 5 (p.63) shows the distribution of participants by location. The Adelaide metropolitan suburbs were divided into directional areas (according to Qpzm Australia, 2004). As can be seen, participants were sampled widely (i.e. unbiased recruitment) from across the Adelaide metropolitan area. The lower proportions of participants from the Adelaide Hills and the city business district were likely due to these areas having lower residential occupancy.

Demographic information on socio-economic status, marital status and nationality/ethnicity was obtained via questionnaire. Socio-economic status was measured indirectly by education where 'years of formal education' was the principal measure; highest level of formal education reached was also assessed for descriptive

purposes. Over a third of participants (38%) indicated that high school was their highest educational level reached. However, academic attainment ranged from primary school (12.7%) to a higher tertiary degree (5.3%). Average years of education received was 11.68 years (SD = 4.13). The majority of participants were either married or widowed (see Table 18); 67% were born in Australia and, of those not born in Australia, most (23%) came from the United Kingdom.

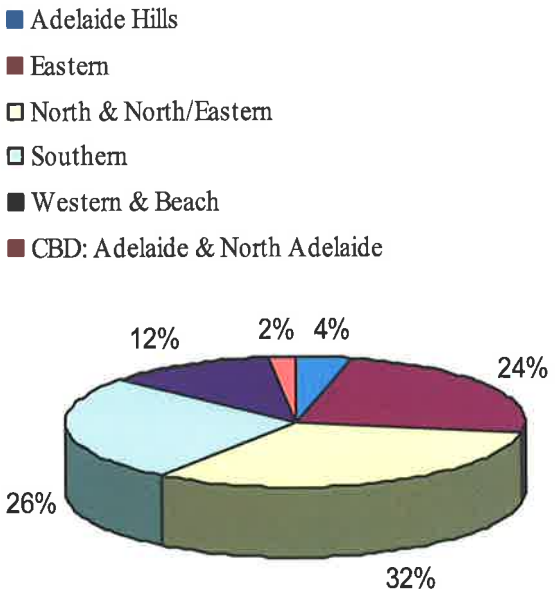


Figure 5 An overview of where participants resided

Table 18
Number of people in each marital status category (N = 150)

Marital Category	Frequency	Percent
Married or Defacto	69	46
Widowed	64	43
Separated or Divorced	10	7
Never Married	7	5

3.3.1 Selection and exclusion criteria

Age (70 years of age and older), fluency in English and absence of dementia were all selection criteria in the current study. Age was a criterion based on the fact that decline in many physical and cognitive tasks (and consequently dependence in daily functioning) is not always observable or significant in younger cohorts (e.g. Fillenbaum,

1985; Spector, Katz, Murphy & Fulton, 1987). Fluency in English and absence of dementia were additional selection criteria because both of these can confound performance on questionnaires and cognitive tasks (e.g. Bäckman & Hill, 1996).

3.3.2 Demographic information across test sessions

Sample attrition, or selective drop out, is an inherent problem with longitudinal studies (Anstey, Hofer & Luszcz, 2003; Frerichs & Tuokko, 2006; Rabbit, Diggle, Smith, Holland & McInnes, 2001; Salthouse, 1991a; Schaie, 1975). This term refers to the trend whereby less motivated or less healthy individuals (physically or mentally) discontinue participation (Anstey & Hofer, 2004). Over the 18 month period of the current study, an attrition rate of 15% was anticipated, based on earlier studies. For example, Hultsch et al. (1992) had an attrition rate of 32.2% over a 3 year interval (initial N = 484 and final N = 328); and Anderson et al. (1998) had an attrition rate of 40% over a 4 year period (initial N = 5079 and final N = 3046). In fact, the rate was 15.3% (test session 1, N = 150; test session 3, N = 127). Figure 6 (p.65) summarises demographic information for all three test sessions. From these data, it appears that more females than males dropped out over the course of the study. A detailed analysis of factors associated with attrition is provided in Chapter 6.

3.4 Apparatus and measures

In addition to assessment regarding selection criteria (i.e. presence of dementia), the test battery consisted of three main categories of variables: outcome, covariate and biomarker[7]. Outcome and covariate measures were assessed at the initial and final test occasions (although not all measures were assessed on both occasions); and biomarker variables were measured on all three test occasions. Age was included as a 'biomarker'

[7] These variable categories are helpful in subsequent descriptions of analyses in Chapter 5 and 6. Hierarchical regression analysis is the main analytical method used where outcome measures feature as dependent variables and biomarker measures act as independent variables. Covariate measures are entered as independent variables prior to entering biomarker variables.

measure because it is used as a benchmark against which to compare the predictive performance of other biomarker variables (see Chapter 5 for the rationale behind this). The specific measures in the test battery are described below.

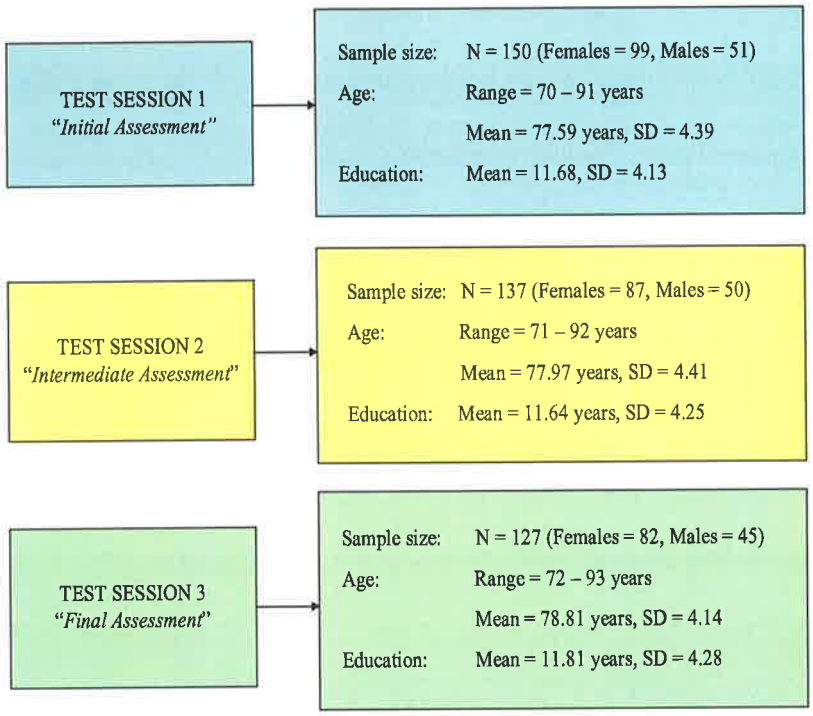


Figure 6 Sample size, age and education demographics for all three test sessions of the current study

3.4.1 Selection criterion assessment: Screening for the presence of dementia

Age is a prime risk factor for dementia, irrespective of its aetiology [e.g. vascular dementia, senile dementia of the Alzheimer's type (Blansjaar, Thomassen & Van Schaik, 2000)]. Low education and depression have also been associated with incidence of dementia (Gallassi, Morreale & Pagni, 2001; Kiosses, Klimstra, Murphy & Alexopoulos, 2001). Senile dementia of the Alzheimer's type is the most frequently diagnosed form (Harman, 1998). The Alzheimer's Disease Assessment Scale (ADAS) is a neuropsychometric test that evaluates 21 symptoms of dementia (Rosen, Mohs & Davis, 1984). The ADAS is divided into two parts, cognitive and non-cognitive. The

non-cognitive section involves conducting a structured clinical interview assessing aspects such as depression, delusions, increased motor activity, tremors and change in appetite. This part of the ADAS is only of use for describing symptoms (Ihl, Grass-Kapanke, Jänner & Weyer, 1999). The cognitive section (ADAS-Cog) involves 11 items focusing on memory, language and praxis (generally defined as putting instructions/thoughts into action, such as drawing objects). Seven of these items are assessed directly from a participant's performance and five of them are assessed indirectly (i.e. rated by the test administrator).

ADAS-Cog can be used to reflect the course of the dementia (Ihl et al., 1999) and has higher sensitivity in detecting mild degrees of cognitive dysfunction or borderline cognitive impairment, compared to more widely used global cognitive measures like the Mini-Mental State Exam (Tafaro et al., 2001; Weyer, Erzigkeit, Kanowski, Ihl & Hadler, 1997). Moreover, the ADAS-Cog does an excellent job at differentiating healthy controls from patients with senile dementia of the Alzheimer's type (Mohs & Cohen, 1988; Mohs, Rosen & Davis, 1983). Consequently, ADAS-Cog was included in the current study (see Appendix A for a copy of this scale). It should also be acknowledged, however, that even though the ADAS was originally designed to evaluate the severity of cognitive and behavioural dysfunctions that are typical of Alzheimer's Disease, there is no claim that this measure is only of use for identifying those persons with senile dementia of the Alzheimer's type (Kim, Nibbelink & Overall, 1994; Rosen et al., 1984). In summary, the ADAS-Cog is more comprehensive and sensitive than other rating scales (Yesavage, Poulsen, Sheikh & Tanke, 1988) and is "a valuable screening test" (Peña-Casanova, 1997, p.105).

The ADAS-Cog is scored by the number of errors made, with a higher score denoting poorer performance. The maximum score achievable is 70. In terms of

classification of cognitive performance, a score between 0 and 13 errors is conventionally accepted as indicating normal cognitive functioning; a score between 14 and 17 errors suggests borderline impairment; and a score greater than 17 errors indicates definite impairment (Tafaro et al., 2001). The ADAS-Cog has demonstrated high reliability and validity. For example, internal consistency, as indicated by Cronbach's alpha, ranges from .75 (Kim et al., 1994) to .81 (Weyer et al., 1997); and test-retest reliability correlations have average .90 whether the interval spans several weeks or several months (Kim et al.; Mohs et al., 1983; Rosen et al., 1984; Weyer et al.). In the current study, internal consistency was adequate (Cronbach's $\alpha = .67$); and test-retest reliability over the 18 month period was moderate ($r = .75$) but appropriate, given that this measure was used to assess the potential development of cognitive impairment. Concurrent criterion validity has been demonstrated by moderate, significant, correlations between the ADAS-Cog and other global cognitive measures. For example, Weyer et al. and Ihl et al. (1999) found correlations between ADAS-Cog and the Mini-Mental State Exam of $r = .66$ ($p < .05$) and $r = .77$ ($p < .01$), respectively.

3.4.2 Outcome measures

The current study included five functional ability outcome measures considered to reflect key attributes of independent functioning. These measures included: activities of daily living, reasoning ability, life satisfaction, crystallised ability and memory-based activities of daily living. The operationalisation of these measures will now be described. Where relevant, the test-retest (or parallel form) reliabilities of outcome measures are reported in Chapter 6.

3.4.2.1 Activities of daily living

Independent functioning in elderly populations is primarily assessed in terms of basic and instrumental activities of daily living (ADL). Basic ADL refer to

physical task demands of personal care (e.g. dressing, mobility) whereas instrumental ADL refer to more complex tasks involving a more obvious cognitive component (e.g. preparing meals, managing personal finances). Spector et al. (1987) suggested that disability or dependence in instrumental ADL would be a more sensitive predictor of functional decline and impending death in elderly persons than disability in basic ADL alone. Thus basic and instrumental ADL are conceived in terms of a hierarchical relationship; basic, well practised tasks are more resistant to age-related decline and impairment in such activities tends to be preceded by impairment in instrumental ADL (e.g. Judge, Schechtman & Cress, 1996; Spector et al.). Moreover, Ishizaki et al. (2000) demonstrated that the capabilities to carry out each type of ADL are predicted by different types of measures. For example, for persons older than 75 years of age, having weak grip strength or a history of hospitalisation predicted decline in both types of ADL, whereas poor intellectual functioning was a unique predictor of instrumental ADL, and not walking regularly was a unique predictor of basic ADL. These results are consistent with the view that instrumental ADL are strongly associated with cognitive functioning (e.g. Aguero-Torres et al., 2001; Lawton & Brody, 1969). Therefore, when assessing ADL, it is important to make a distinction between each type of activity.

Basic and instrumental ADL were assessed in the current study via a composite, self-report questionnaire. The composite questionnaire was based on Lawton's Instrumental Activities of Daily Living Scale (Lawton & Brody, 1969) and the Bristol Activities of Daily Living Scale (Bucks, Ashworth, Wilcock & Siegfried, 1996). The Lawton scale contains six basic and eight instrumental ADL items. Participants' ability in these tasks is rated by a third-party (e.g. carer, social worker), where one point is allocated if the participant can complete basic ADL items unassisted and instrumental ADL items with some assistance. The maximum possible score is 14, with higher scores

indicating more independence. Lawton and Brody observed that although inter-rater reliabilities ranged from .87 to .91, extensive reliability testing had not been carried out on the Lawton scale. In terms of concurrent criterion validity, Lawton and Brody reported moderate to high correlations between physical function classifications provided by a physician and both basic (r = .62) and instrumental (r = .40) ADL items. The items in the Lawton scale "were found to have practical utility in widely diverse settings, with a range of population groups of aged, and for a variety of goals" (Lawton & Brody, p.185).

The Bristol scale was designed to provide a brief, accurate and sensitive assessment of daily functioning ability in individuals suffering from memory problems and dementia. Typically, carers are asked to complete the 20 items of this scale, based on the individual's general ability over the previous 2-week period. Each item is allocated 0 (not applicable), 1 (some assistance), 2 (lots of assistance), or 3 (totally dependent) points. The maximum score achievable is 60, indicating that the individual is totally dependent on others to complete daily activities. The Bristol scale has demonstrated good test-retest reliability (r = .95) and concurrent criterion validity, correlating significantly with an observational scale (r = .65; Bucks et al., 1996). Moreover, the Bristol scale "is sensitive to a wide range of ADL performance, from individuals who require assistance or are totally dependent on most tasks, to individuals who are completely independent on all tasks" (Bucks et al., p.117).

The composite ADL questionnaire used here was developed to maximise the range of activities assessed and to provide a more detailed scoring system in order to distinguish the functional ability of the current sample (i.e. non-demented elderly persons living in their own homes). Items from the Lawton and Bristol scales with different names but assessing similar activities were represented by one item in the

composite questionnaire. For example, "hygiene" and "bathing" represent similar capabilities and were labelled "bathing" in the composite questionnaire (see Appendix B for a copy of this). "Orientation" items from the Bristol scale were excluded from the composite questionnaire because of duplication with other test measures (i.e. items from the dementia screening test, ADAS-Cog). In total, the composite ADL questionnaire included 9 basic and 12 instrumental ADL items. Similar to the Bristol scale, the composite questionnaire allocated a range of points to each level of assistance required for each activity. The range of assistance options varied for each item and consequently so did the maximum score allocated; Appendix B (Part A) also includes a break-down of the point-allocation system.

The composite ADL questionnaire was administered at both initial and final measurement occasions. However, due to feedback on the initial measurement occasion an abridged version was employed on the final test session[8]. As a consequence of the study design employed here (i.e. requiring calculation of change scores – see Chapter 6 for details of this), initial and final measurements needed to be comparable. Therefore, in all subsequent descriptions and analyses of ADL, performance on the abridged questionnaire (shown in Appendix B, Part B) is reported. For this version, there were two basic ADL items (with a total maximum score of 7) and seven instrumental ADL items (with a total maximum score of 27). Therefore the maximum total score achievable is 34, with higher scores denoting more independent functioning.

3.4.2.2 Reasoning ability

Impaired cognition has been reported to be significantly related to increased functional dependence (as discussed in Chapter 1). Fluid ability is a central

[8] For example, initial results indicated no variance in performance on the majority of basic ADL items, such as 'eating', 'dressing' and 'bathing'. This is believed to be a product of the sample. Furthermore, some instrumental ADL items (e.g. 'gardening' and 'hobbies') tended to receive 'not applicable' responses, which did not add to the clarity of participants' functional ability. A more detailed discussion on these issues is provided in Chapter 5.

cognitive factor that is closely associated with ability in other cognitive domains, such as working memory and processing speed (e.g. Conway, Cowan, Bunting, Therriault & Minkoff, 2002; Fry & Hale, 2000; Kyllonen & Christal, 1990; Süss, Oberauer, Wittmann, Wilhelm & Schulze, 2002). Being able to correctly identify specific relationships between constructs and to problem-solve appropriately are important in daily life. In fact, poor reasoning skills could generate life-threatening situations. Three tests representing fluid ability were selected for the current study: Raven's Standard Progressive Matrices (De Lemos, 1995), the Culture Fair test (Cattell & Cattell, 1959) and Concept Formation (Woodcock & Johnson, 1989). These tests are considered to strongly reflect the fluid ability factor of intelligence (McGrew, 1997). All three tests were computer-administered[9] on both the initial and final measurement occasions. Raven's Standard Progress Matrices (Raven's) was selected as the representative task of reasoning ability for the following reasons:

(1) it is commonly used as a measure of abstract reasoning and is considered to reflect Spearman's g (Jensen, 1982; Lezak, Howieson, Loring, Hannay & Fischer, 2004; Salthouse, 1993);

(2) Salthouse (1993) has reported that substantial age-related differences in Raven's Progressive Matrices performance exist;

(3) there were some missing data on the Culture Fair test (Chapter 4 discusses the sources of missing data); and

(4) the distribution of performance on Concept Formation was bi-modal (see Appendix C), making it undesirable for regression analyses (see Chapter 5 for further explanation of this issue).

Raven's is typically used as a pen and paper, non-verbal reasoning task. It

[9] All computerised tasks were programmed with the software Presentation (version 7.1) and administered on a Dell PC laptop.

consists of five sets (A, B, C, D, and E), each containing 12 visuo-spatial abstract puzzles or matrices (for a total of 60 items). Each puzzle has part of it missing. Participants are required to select which pattern, out of six or nine options, would best fit or complete the puzzle. The puzzles require participants "to conceptualise spatial, design, and numerical relationships ranging from the very obvious and concrete to the very complex and abstract" (Lezak et al., 2004, p.579). More specifically, puzzles in Set A involve incomplete figures, which require pattern matching skills; and puzzles in Set B onwards require reasoning by analogy. The Sets get more difficult as they progress, with many involving mathematical concepts (Lezak et al.). Participants are allowed a total of 20 minutes to complete as many puzzles as possible (2 minutes for Set A, 3 minutes for Set B, 4 minutes for Set C, 5 minutes for Set D and 6 minutes for Set E). Higher scores indicate better performance, i.e. greater reasoning ability. Test-retest reliability data for Raven's are generally high, with a median correlation from extensive studies of $r = .80$. This reliability is slightly lower with retest intervals longer than one year (Spreen & Strauss, 1998). Concurrent criterion validity between Raven's and other conventional tests of intelligence such as the Wechsler scales has been shown to be moderate (e.g. $r = .70$; O'Leary, Rusch & Guastello, 1991).

On account of the exhaustive nature of the test battery assembled for the current study, a short form of Raven's was used. For this, the even and odd items were divided, to comprise parallel forms of this test. The parallel form from even items was used on the initial assessment occasion and the parallel form from the odd items on the final (see Chapter 6, p.169 for parallel form reliability). Each form consisted of six abstract reasoning puzzles from each of the five sets. Therefore, the maximum score achievable on one form was 30. Accordingly, the participants had a total of 10 minutes to complete as many puzzles as possible (practice was provided and was not timed). The instructions

featured on the computer were consistent with the relevant instructions from the technical manual (De Lemos, 1995). Participants responded to puzzles by pressing the appropriately-numbered key on the keyboard.

3.4.2.3 Life satisfaction

Life satisfaction is another outcome measure considered to reflect key attributes of independent functioning. Life satisfaction is closely associated, if not interchangeable, with other terms such as positive affect, emotional well-being and quality of life. For example, positive affect "reflects an individual's emotional relationship with his or her environment, communicating qualities such as happiness, personal satisfaction, optimism and morale" (Ostir, Markides, Black & Goodwin, 2000, p.473). Moreover, "the concept of life satisfaction is a construct universally accepted as a valid measurement of quality of life in the elderly" (Salamon, 1988, p.45). The subscales of the Life Satisfaction Scale [formerly known as the Life Satisfaction in The Elderly Scale by Salamon & Conte (1984)] mirror the above description of quality of life. The Life Satisfaction Scale (LSS) measures eight categories pertaining to quality of life. These include: taking pleasure in daily activities, regarding life as meaningful, goodness-of-fit between desired and achieved goals, positive mood tone, positive self-concept, perceived level of health and physical well-being, financial security, and social contacts. As discussed in Chapter 1, the maintenance of independent functioning is associated with quality of life. Similarly, the presence of positive affect appears to "protect individuals against physical declines in old age" (Ostir et al., p.473).

The LSS was designed as a general measure of well-being relevant to older adults (Salamon, 1985). Its development stemmed from a lack of existing scales at that time that demonstrated good reliability and validity (Salamon, 1988). It contains 40 statements with each of the eight categories containing 5 statements. The LSS can be

administered in an interview format or as a self-report questionnaire. The latter applied in the current study where participants were asked to respond to statements on a 5-point Likert scale. For example, options ranged from "completely dissatisfied" (1 point), through 'dissatisfied', 'partially satisfied', 'satisfied', to 'very satisfied' (5 points). The wording of the response options varied with each statement but the point-values remained the same. The maximum score achievable is 200, with higher scores denoting greater life satisfaction. Psychometric properties of the scale were assessed by Salamon (1988). He found internal consistency, measured by Cronbach's alpha, ranged from .50 to .90 for the subscales and total scale; test-retest reliability over a 6 month interval produced a correlation of r = .67; and concurrent criterion validity of r = .85 was found with ratings made by health care and social service workers. Salamon (1988) interpreted the moderate test-retest correlation as reflecting the sensitivity of the scale to change over time in older adults. Test-retest reliability was not available for the current study because this questionnaire was administered on the initial assessment occasion only (see Chapter 5 for an explanation); and internal consistency of this questionnaire was high in the current sample (Cronbach's alpha ranged from .73 to .83 for the subscales and total scale).

3.4.2.4 Crystallised ability

Crystallised tasks are considered to reflect processes that are familiar, part of human culture and learned or influenced by education. Such tasks include those that are language, vocabulary and general knowledge-based. The finding that performance on tests of crystallised ability remains relatively stable with increasing age is robust (e.g. Ryan, Sattler & Lopez, 2000; Willis, 1996); and in older populations, improvement over time on these tasks is also not out of the ordinary. Generally, deterioration on these types of tasks does not occur until the seventh decade or later, but

when this decline begins, it is often rapid (Salthouse, 2004; Schaie, 1994). Such rapid decline has been closely associated with mortality and is believed to be an indicator of 'terminal drop' (Palmore & Cleveland, 1976). Based on this, it was anticipated that participants experiencing decline over the 18 month period of the current study may be closer to death and subsequently be more dependent in their daily functioning. That is, cognitive impairment has been associated with dependence in activities of daily living. Therefore, change in performance on crystallised ability over 18 months has been employed as an outcome measure in the final measurement occasion. In order for the change measure to be calculated, crystallised ability was assessed on the initial and final measurement occasions.

On both measurement occasions, three tests of crystallised ability were administered: Similarities (Wechsler, 1997), Information (Wechsler, 1997) and Spot-the-Word (Baddeley, Emslie & Nimmo-Smith, 1992). The first two tests are considered to be strong indicators of crystallised ability (McGrew, 1997) and Spot-the-Word is a reliable and valid measure of premorbid intelligence (Baddeley, Emslie & Nimmo-Smith, 1993). For all tasks, higher scores signify better performance. A composite score (sum of crystallised ability scores divided by three) was used as the outcome measure in the current study[10].

Similarities is a verbal task involving the identification of elements that two items have in common. It consists of 19 word pairs ranging in difficulty from obvious similarities (e.g. orange and banana; they are both fruit) to more abstract connections (e.g. hibernation and migration; they are both seasonal activities). A range of points (0 to 2) is allocated depending on the quality of the answer provided. The instructions to

[10] A composite score was employed to minimise task-specific variance. This particular composite score was used in favour of a factor score or average z-score measure because the latter measures would make the calculation and use of change scores (see Chapter 6 for the change score calculation method) less interpretable.

participants, procedure and scoring method employed in the current study were consistent with those outlined in the instruction manual for the Wechsler Adult Intelligence Scale (WAIS) - III (Wechsler, 1997). The maximum possible score is 33. Similarities has good reliability, with a split-half reliability of r = .86 and a test-retest reliability of r = .83 over a period of 5 weeks (Kaufman & Lichtenberger, 1999).

Information is another verbal task from the WAIS-III (Wechsler, 1997). The task involves answering a series of short questions testing knowledge about objects, places, people and common events (Kaufman & Lichtenberger, 1999). Information has shown very high reliability, both split-half (r = .91) and test-retest (r = .94) over a five week interval (Kaufman & Lichtenberger). The general knowledge task used in the current study is based on Information. The current version is a computerised, multiple-choice test that consists of 40 short questions, each accompanied by four alternate answers (see Appendix D for test material). Participants responded to the questions by pressing a designated key on the keyboard. The test was not timed but a minimum of two questions out of every six needed to be answered correctly otherwise the test ended automatically. Questions from Information were modified in order to make them relevant for an Australian, elderly population and a multiple-choice format was added to prevent the task becoming too difficult. The questions became more difficult as the task progressed. One point was allocated per correct answer.

Spot-the-Word involves identifying which of two words is a real word. There are 60 pairs of words and each pair consists of a real, legitimate word from the English language, and an invented word. Selection of the real word is based on lexical decision. In other words, the ability to identify the real word may stem from several features of the word, such as "its meaning, its orthographic appearance, its sound, or indeed a general feeling of familiarity based on all of these" (Baddeley et al., 1993, p.56). The

real words in each pair range in difficulty, beginning with words that are commonly encountered (e.g. daffodil) moving to less commonly known words (e.g. shako) and finishing with fairly obscure words (e.g. strubbage). Practice items are provided. One point is allocated for each real word selected (i.e. maximum score is 60). Spot-the-Word has two parallel forms; Form A was used on the initial measurement occasion and Form B on the final occasion (see Baddeley et al., 1992 for actual test and specific instructions). The two forms have moderate to high (r = .78) alternate-form reliability (Baddeley et al., 1993). In terms of concurrent criterion validity, Form A was found to correlate moderately (r = .60) with the Mill Hill Vocabulary Test (Baddeley et al., 1993). The Mill Hill Vocabulary Test has been shown to be highly correlated with verbal ability (Raven, 1958). Similarly, Spot-the-Word (Form A) correlated highly (r = .83) with the National Adult Reading Test (Nelson, 1982), an established measure of premorbid verbal ability (Baddeley et al., 1993).

3.4.2.5 Memory-based activities of daily living

This outcome measure was administered on the final measurement occasion only. The aim of including this questionnaire was to provide a measure of activities of daily living that was more able to differentiate the functional ability of a well-educated, healthy and independent functioning sample (see Chapter 5 for a discussion on this). The memory-based activities of daily living (ADL) questionnaire is based on the Subjective Scale to Investigate Cognition in Schizophrenia (SSTICS; Stip, Caron, Renaud, Pampoulova & Lecomte, 2003). The SSTICS comprises 21 items that assess the following areas of cognition: working memory, explicit long-term memory, attention, language and praxia (Stip et al.). The scale relies on self-report and investigates whether individuals perceive themselves as having difficulties remembering various types of information. Responses are made on a 5-point Likert

scale that ranges from 0 (never experience difficulty remembering things) to 4 (have difficulty remembering things very often). The SSTICS includes items such as, 'Do you have difficulty doing household chores or repairs? For example, do you ever forget how to cook things or what ingredients go into a recipe?'. Psychometric properties of the SSTICS are sound. Stip et al. reported an internal consistency (Cronbach's alpha) of α = .86 for the total scale (subscales ranging from α = .57 to α = .72) and a test-retest reliability of r = .82 over an average interval of 11 days. Moreover, on most areas of cognition assessed by SSTICS, subjective memory problems correlated with performance on objective neuropsychological tests (see Prouteau et al., 2004 for details).

The only difference between the memory-based ADL questionnaire employed in the current study and the SSTICS is the scoring method. In the current study, a 5-point Likert scale was used where 1 represented experiencing memory problems 'very often' and 5 represented 'never' experiencing memory problems. The Likert scale was reversed to maintain consistency with other measures in the test battery where higher scores reflect better performance or more independent functioning. The maximum score achievable is 105, indicating the perception of no memory problems. Test-retest reliability could not be assessed in the current study because this measure was assessed only once; and the internal consistency in the current study was moderate to high with Cronbach's alpha ranging from .73 to .75 for the subscales and total scale.

3.4.3 Covariates

This category includes variables that have been reported to influence either outcome and/or biomarker variables, thereby confounding relationships between these measures. Four types of covariates were assessed: demographic variables, depression, health and lifestyle. Each type of covariate is described below.

3.4.3.1 Demographic variables

In terms of the demographic measures already mentioned (see earlier this Chapter, pp.62-63), education (years of formal education and highest level attained) and sex are among those with the strongest influence on a wide range of measures. In particular, education correlates positively with performance on all types of cognitive tasks, especially with tasks of crystallised ability (e.g. Albert et al., 1995; Anstey & Christensen, 2000; Baltes, 1987). Similarly, cognitive decline has been reported to be slower in individuals with higher educational attainment (e.g. Wiederholt et al., 1993). Low levels of education have also been associated with higher functional dependence in elderly populations (Harris, Kovar, Suzman, Kleinman & Feldman, 1989; Parker, Thorslund, Lundberg & Kareholt, 1996). Years of education is the covariate applied in the current study. With respect to the influence of sex on cognition, it has been reported that women tend to perform better on verbal tasks compared to men, whereas men tend to perform better on visuo-spatial tasks compared to women (e.g. Schaie, 1994; Wiederholt et al.). In regard to independent functioning, females are more likely to be dependent in daily functioning compared to males (Guralnik, Leveille, Hirsch, Ferrucci & Fried, 1997; Parker et al.).

3.4.3.2 Depression

Among individuals 65 years of age and older, depression is the most common psychiatric disorder (Silvestri et al., 2001). The identification of depression is important because, apart from being a potential risk factor for dementia, depression can influence cognition and the ability to perform activities of daily living. For example, depression can cause cognitive impairment, affecting abilities such as attention, executive function, psychomotor activity, visuo-spatial skills and memory, including working memory (Channon & Baker, 1994; Gallassi et al., 2001; Lockwood,

Alexopoulos & van Gorp, 2002; Rockey, 1997). "Even in samples of healthy older adults, those who score higher on measures of depression have lower cognitive scores, particularly on memory tasks" (Dawson, Winocur & Moscovitch, 1999, p.96). Furthermore, individuals suffering from depression are more likely to indicate that they are experiencing difficulties with both basic and instrumental activities of daily living (Grigsby, Kaye, Baxter, Shetterly & Hamman, 1998). Similarly, individuals with high levels of depressive symptoms at initial testing are more likely to develop dependence in activities of daily living at subsequent testing (e.g. after a 2.5 year interval; Bruce, Seeman, Merrill & Blazer, 1994).

Depression is also associated with sex and health (other covariates in the current study). For example, depression tends to be more prevalent in females and individuals with at least two chronic physical conditions have been shown to have an increased rate of severe major depression compared with those without chronic physical conditions (Gareri et al., 2001). Results from Kiosses et al. (2001) also support the relationship between depression, activities of daily living and chronic medical conditions.

According to Brink et al. (1982), there are three main categories of symptoms fundamental to depression; affective (e.g. crying, sadness, apathy), cognitive (e.g. thoughts of hopelessness, helplessness, suicide, worthlessness, guilt), and somatic (e.g. fatigue, appetite, sleep, libido). However, the constituent symptoms of depression change in later life. Thus, somatic symptoms may be present in non-depressed elderly and this particular category of symptoms has therefore reduced diagnostic power in elderly samples (Brink et al.). Similarly, depressed elderly persons may report memory problems and cognitive impairment, both of which can also be seen in non-depressed persons. Consequently, there can be confusion between dementia and depression in elderly populations (Yesavage et al., 1983).

The Geriatric Depression Scale (GDS; Brink et al., 1982) is one of the most commonly used scales for the identification of depression in elderly persons (Perlis et al., 2002). The GDS was used in the current study. Its development was necessitated by the lack at that time of existing scales that emphasised symptoms characteristic of depression during old age (Brink et al., 1982), earlier scales emphasising only somatic and cognitive symptoms. As a result, the 30 items of the GDS focus more on affect. The items involve questions referring to topics such as lowered affect, inactivity, irritability, withdrawal, distressing thoughts, and negative judgements about past, present and future (Brink et al., 1982). In this self-report scale, participants are asked to respond 'yes' or 'no' to the 30 questions. This dichotomous form of response was considered an improvement on the graduated responses (e.g. a Likert scale such as never, sometimes, usually, always) employed in previous scales. Graduated responses have been criticized for confusing elderly persons because they require decisions to be made about subtle differences (Brink et al., 1982). In the GDS, one point is allocated per item when the response indicates a depressive thought (i.e. there are 20 'yes' responses and 10 'no' responses, which generate a maximum score of 30). Generally, a score of 0 - 10 can be considered the range for normal, non-depressed elderly persons; scores of 11 - 20 indicate mild depression; and scores of 21 - 30 indicate moderate to major depression (Brink et al., 1985). The GDS has demonstrated good reliability and validity (Yesavage et al., 1983). Internal consistency (Cronbach's $\alpha = .94$) and test-retest reliability (over a one week interval, $r = .85$, $p < .01$) were both high and concurrent criterion validity has been demonstrated, with the GDS having strong, significant correlations with other depression scales; for example, $r = .84$ with Zung's Self-Rating Depression Scale and $r = .83$ with the Hamilton Rating Scale for Depression (Yesavage et al.).

3.4.3.3 Health

In addition to an increased risk of suffering from depression, health problems can have an independent negative influence on cognition and the ability to carry out daily activities. For example, cross-sectional studies have shown that physical health has a strong negative impact on fluid and crystallised abilities, but more so on fluid (Christensen, Korten, Jorm, Henderson & Scott, 1996). However, some studies have shown that the influence of physical health is negligible after controlling for the demographic variables education and sex (Anstey, Stankov & Lord, 1993; Luszcz, Bryan & Kent, 1997; Salthouse et al., 1988). Health problems also have a detrimental effect on capacity to carry out activities of daily living independently. For example, Mulrow, Gerety, Cornell, Lawrence and Kanten, (1994) found that poorer perceived health was associated with more dependence in activities of daily living ($r = .39$, $p < .01$); and the absence of chronic disease symptoms at one point in time has been found to be associated with independent functioning 19 years later (Guralnik & Kaplan, 1989). In summary, health is an important covariate when investigating independent functioning outcome measures in elderly people.

Health is a broad topic and some issues are more relevant than others to elderly samples. Consequently, the topics included in the self-report health questionnaire used in the current study (see Appendix E for exact questions) are based on information from a variety of sources. These sources include: Benjamin et al. (1997), the Personal Health Inventory from the Byrd Health Attitude Scale (Byrd, 1941), and health measures used in previous studies examining independent functioning in elderly persons (e.g. Hébert et al., 1999; Mor et al., 1989; Wolinsky, Stump & Clark, 1995). The specific psychometric properties of the sourced health items are unknown. However, in general, studies have found congruence between self-reported health status and physicians' reports, at least in

community-dwelling, older populations (e.g. Fillenbaum, 1979; Kaplan & Kotter, 1985; La Rue, Bank, Jarvik & Hetland, 1979). The composite health questionnaire in the current study is divided into three main areas: general health, disease symptoms and disease history (health habits are also included in the questionnaire but will be discussed separately in the following section on lifestyle). The general health section includes items on vision, doctor and hospital visits and overall perception of health. However, the sections on disease symptoms and disease history are of particular relevance in the current study[11].

For assessment of disease symptoms, participants were asked to indicate which of 51 symptoms of illness or disease they had experienced in the past 6 months. Unlike the Personal Health Inventory (Byrd, 1941), participants were also asked to rate the severity of each symptom experienced on a 4-point scale (ranging from 1, 'never experienced' to 4, the symptom bothered them 'extremely'). The higher the score, the more severe the symptoms experienced. The symptoms surveyed represent five main systems of the human body: (1) head, ears, nose and throat; (2) circulatory system; (3) gastro-intestinal and genito-urinary systems; (4) musculoskeletal system; and (5) central nervous symptoms; a breakdown of which items pertain to which system is provided in Appendix G (Part A). A total score was calculated where a minimum score of 51 indicates that all symptoms were 'never experienced' and a maximum score of 204 indicates that all symptoms bothered the respondent 'extremely'. The inclusion of assessment of both symptoms and disease history had three main purposes. Firstly, it served as a safety net, potentially identifying individuals who suffered from a disease or ailment but who did not recognise their situation based on the name of the disease alone. Secondly, it allowed for the identification of specific parts of the body that may

[11] On the whole, participants were very healthy and the general health section showed the least amount of response variance. General health measures also had a limited influence on other variables in the test battery (see Appendix F).

strongly influence the ability to function independently. Thirdly, the opportunity for participants to indicate their perceived severity of illness may have assisted in establishing a threshold for when particular symptoms impact on independent functioning outcome measures.

Disease history required that a participant indicate which of 17 chronic diseases they had suffered from in their lifetime (Appendix F includes a list of the diseases). In addition to this, and for each condition, participants were asked if the disease was ongoing (i.e. suffering from it currently) and at what age they began suffering from the condition. The diseases were selected based on the fact that:

(1) their prevalence increases with increasing age (e.g. Anderson et al., 1998; Goldberg, Dengel & Hagberg, 1995);

(2) they have been shown to influence cognitive and/or physical performance (e.g. Atkinson et al., 2005);

(3) they have frequently been used in other studies on ageing and daily functioning (e.g. Carlson et al., 1999; Mulrow et al., 1994; Stuck et al., 1999);

(4) and moreover, the presence of many of these diseases has been found to be directly related to an increase in dependence in activities of daily living (e.g. Boult, Kane, Louis, Boult & McCaffrey, 1994; Chen et al., 1995; Ensrud et al., 1994; Furner, Rudberg & Cassel, 1995; Guccione et al., 1994).

The 17 chronic diseases represented the following disease categories (based on those used by Stuck et al., 1999): (1) cardiovascular; (2) gastro-intestinal; (3) haematological; (4) metabolic; (5) musculoskeletal; (6) pulmonary; (7) neurologic; and (8) miscellaneous (see Appendix G, Part B for a breakdown of which diseases represent which category). The total number of diseases *ever* suffered from was used as the score for this section (maximum score = 17).

3.4.3.4 Lifestyle measures

"Health habits or behaviours strongly contribute to functional ability" (Sulander et al., 2005, p.197). Health habits such as smoking, physical activity, alcohol consumption and diet also interact with overall health status so it can be unclear whether these variables contribute directly to functional ability or indirectly via health. The interdependent or circular relationship of these aforementioned variables is well-documented (Albert et al., 1995; Clarkson-Smith & Hartley, 1990; Ensrud et al., 1994; Goldberg et al., 1995; LaCroix, Guralnik, Berkman, Wallace & Satterfield, 1993). For example, the prevalence of chronic diseases increases with increasing age and the development of a chronic disease can reduce the ability to exercise. What is more, current or previous smoking can often be the cause of disease. Health status can also place restrictions on diet or even impair nutrient uptake during digestion. Therefore, it is important to include both health and lifestyle measures in a study such as this. Four aspects of lifestyle were included in the current study: smoking, exercise, alcohol consumption and diet. Their relevance to this research project and how they were measured will now be discussed.

There is no question as to the far reaching deleterious effects of cigarette smoking. For example, cigarette smoking increases the prevalence of peptic ulcers and lung and bladder cancer; it is also a major risk factor for atherosclerosis, hypertension and other chronic diseases (Benjamin et al., 1997; Wang et al., 2002). The detrimental effects of cigarette smoking also extend to activities of daily living (ADL), physical ability and cognition. Current smokers have been shown to be more dependent in ADL than both former smokers and people who have never smoked (Ensrud et al., 1994; Kaplan et al., 1993; Nelson, Nevitt, Scott, Stone & Cummings, 1994; Stuck et al., 1999; Sulander et al., 2005; Wang et al., 2002). A similar relationship has been established in

regard to physical ability. That is, Tabbarah et al. (2002) found that, compared to individuals who had never smoked, current smokers were more likely to show decline in grip strength over a 7 year period. Atkinson et al. (2005) also found that decline in physical ability (i.e. walking speed) was associated with former cigarette smoking. Moreover, they found that current cigarette smoking was associated with both physical and cognitive decline over a 3 year period. However, according to Schinka et al. (2002), few studies have investigated the relationship between cigarette smoking and cognition in elderly populations and of those studies that have, there is little consensus on the nature of the association between smoking and cognition. Therefore, given the salient association that cigarette smoking has with ADL in particular, smoking category (i.e. current smoker, former smoker and never smoked) was included as a covariate in the current study.

Participation in exercise (e.g. walking, swimming) and physical activity (e.g. cleaning, mopping the floor) in elderly populations has been linked to relative advantage in three main areas: the ability to function independently; cognitive performance; and likelihood of experiencing certain chronic diseases. Lack of exercise has been associated with functional dependence in many studies (Ensrud et al., 1994; Hébert et al., 1999; Kaplan et al., 1993; LaCroix et al., 1993; Mor et al., 1989; Parker et al., 1996; Simonsick et al., 1993; Stuck et al., 1999; Wolinsky et al., 1995). For example, Wang et al. (2002) demonstrated that individuals who exercised three times or more a week experienced less dependence in basic and instrumental ADL; and Sulander et al. (2005) found that those who exercised regularly (e.g. walked for 30 minutes or more, four times a week or more) were on average, almost three times more likely to be functionally independent.

The benefits of exercise also extend to cognitive performance. Previous studies

have demonstrated that higher levels of physical activity are associated with higher scores on various measures of cognition. For example, Clarkson-Smith and Hartley (1989) demonstrated a significant positive relationship between exercising and performance on tests of fluid ability and working memory. In terms of fluid ability, individuals (aged in their 60s - 70s) in a high exercise group achieved a higher score on Raven's Advanced Progressive Matrices (mean = 15.55, SD = 2.56) than those in a low exercise group (mean = 11.82, SD = 4.45). This trend continued with performance on a working memory task ('Letter Sets'), where a high exercise group achieved a mean score of 46.42 (SD = 14.98) compared to a mean score of 34.16 for a low exercise group (SD = 16.61). Similar results have been reported on other studies measuring these cognitive domains (Christensen & Mackinnon, 1993; Elsayed, Ismail & Young, 1980; Powell & Pohndorf, 1971). However, there is no consensus as to which specific domains of cognition are most likely to show an association with physical activity (Chodzko-Zajko, Schuler, Solomon, Heinl & Ellis, 1992). For example, studies have both found (Clarkson-Smith & Hartley, 1989) and not found (Elsayed et al., 1980) positive associations between physical exercise and crystallised ability. From a longitudinal viewpoint, elderly people who participate in regular, strenuous activity have been shown to experience less cognitive decline than those individuals who are less active (Dawson, et al., 1999).

The effects of physical activity appear to be widespread. In addition to improvements in cognitive performance and the capability to carry out daily activities, exercise is also reported to decrease the prevalence of chronic diseases. For example, regular exercise can reduce the incidence of cardiovascular disease, stroke, diabetes and some cancers. This relationship may be mediated by the fact that regular exercise can lower body mass index and blood pressure, common risk factors for diabetes, stroke and

heart disease (Bassey, 2000; Goldberg et al., 1995). In the current study, participants were asked to report how many times per week they exercised (frequency) and how long they exercised each time, on average (duration). From this information, total time spent exercising per week was calculated as a measure of exercise. Participants also reported the type of exercise they engaged in but this could not be meaningfully analysed beyond descriptive statistics (see Chapter 5 for a discussion of this point).

Like cigarette smoking, alcohol consumption has been associated with higher incidence of certain diseases; peptic ulcers, cirrhosis of the liver (Benjamin et al., 1997), breast cancer, stroke, and hypertension (Hochschild, 1990). Conversely, light to moderate alcohol consumption has been associated with lower incidence of cardiovascular disease (Klatsky, Armstrong & Friedman, 1990; Rimm et al., 1991). Thus, alcohol consumption can have both positive and negative health effects. The consumption of alcohol can also impact on the ability to carry out daily activities and cognitive performance; the critical factor in determining whether the effects of alcohol are beneficial or detrimental appears to be the quantity of alcohol consumed. For example, several studies have demonstrated that moderate drinkers, who on the whole can be described as having no more than two alcoholic beverages a day, are more independent in daily activities than both abstainers and heavy drinkers (e.g. Guralnik & Kaplan, 1989; Nelson et al., 1994; Wang et al., 2002). Sulander et al. (2005) also found that both abstainers and heavy drinkers reported more limitations in basic ADL than light to moderate drinkers.

In terms of the effect of alcohol on cognition, individuals who consumed one or two drinks a day were at a lower risk of memory problems than those who consumed alcohol in greater amounts, as well as those who abstained (Holford, 2003). In support of this finding, Perkins et al. (1999) found that lifetime abstention from alcohol resulted

in a greater occurrence of memory impairment. This type of relationship has been documented as both a J-, and U-shaped curve. That is, where light to moderate alcohol consumption is associated with better cognition (or a reduction in cardiovascular disease or independence in daily activities), heavy drinking is associated with poorer cognition (or an increase in cardiovascular disease or dependence in daily activities), and the effects of alcohol abstention lie somewhere in between these two extremes (Perkins et al., 1999). Hendrie, Gao, Hall, Hui and Unverzagt (1996) also found this type of association between alcohol and cognition; on a delayed recall task and a global measure of cognition, heaviest drinkers had the lowest scores, light to moderate drinkers obtained higher scores and abstainers scored intermediately. A study by Stampfer, Kang, Chen, Cherry and Grodstein (2005), using similar types of cognitive measures, demonstrated a similar trend. Based on these findings, drinking status was employed as a covariate in the current study. Participants were categorised as either abstainers (0 drinks per week), light drinkers (.5 to 7 drinks per week) or moderate drinkers (8 to 25 drinks per week). The sample included no heavy drinkers[12].

Nutrition was the final lifestyle measure included in the current study. Nutrition, or rather malnutrition, is a major issue in ageing populations. Malnutrition, which consists of both too much and too little of a nutrient, has serious implications, including (depending on the nutrient involved), increased cognitive impairment and increased incidence of physical impairments and diseases. Whilst malnutrition is considered more prevalent for those in retirement or nursing homes and those who are house-bound, five to 10 percent of community-dwelling individuals are believed to be at risk (Griep, Mets, Collys, Ponjaert-Kristoffersen & Massart, 2000).

---

[12] Five participants (4 males) reported an alcohol intake more than the equivalent of 2 standard drinks a day. However, considering that the highest intake (25 drinks per week) equates to 3.5 standard drinks a day, this is not considered excessive alcohol consumption by Australian standards (Australian Government Department of Health and Ageing, 2006).

According to Stuck et al. (1999), most studies that have looked at the relationship between nutrition and functional ability in elderly people have measured nutrition via body mass index. Individuals with either a high (i.e. overweight) or low (i.e. underweight) body mass index have been shown to be more dependent in daily activities over time than individuals within a normal range for body mass index (Launer, Harris, Rumpel & Madans, 1994). Studies that have measured nutrition via food diaries and/or blood analysis have primarily investigated the relationship between nutrition and cognition in elderly people, as will be discussed below. The current study assessed nutritional status, using a self-report, three-day diet diary (see Appendix H). The diet diary was based on those developed by the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Division of Human Nutrition (Adelaide, Australia). The diary included general diet questions such as 'what type of bread do you most often eat', as well as worksheets for participants to enter their total food consumption each day, over three days. To ensure a balanced view of the participants' diets, the days of the week on which food intake needed to be recorded were pre-selected and specified for each diet diary. One of the three days was always a day during the weekend (i.e. Saturday or Sunday) and the other two days were week days, selected randomly. An example of how to complete the diary was also included. The final component of the diet diary involved participants listing any medications and dietary supplements normally taken because both of these items can influence the overall level of nutrients consumed. For example, a magnesium supplement can include nutrients other than magnesium.

The information provided in these diaries was used to form estimates of average daily intakes of various types of nutrients. For each nutrient, these estimates were used to classify participants into two groups: those who received nutrient intakes *below* the

Australian recommended daily allowance (National Health & Medical Research Council, 1991), and those who received nutrient intakes *above* the recommended daily allowance. In addition to controlling for nutritional status, the relationship between level of nutrient intake and measures of cognition was assessed. This was primarily due to the observation that sub-clinical deficiencies in certain nutrients have been associated with decreased cognitive performance in elderly people (e.g. Goodwin, Goodwin & Garry, 1983). That is, lower nutrient intake may influence the functioning of elderly people, despite this level satisfying the recommended daily allowance (La Rue et al., 1997). It has been proposed that, the greater the nutrient intake above the recommended daily allowance, the greater the benefits to health, cognition, etc (Benton, Fordy & Haller, 1995).[13]

To date, there has only been limited research on the relationship between nutrition and cognition in elderly populations (La Rue et al., 1997). Similar to the research investigating independent functioning, nutrition research that has incorporated cognitive measures has tended to use global indices such as the Mini-Mental State Exam, or employ only one type of task (Fioravanti et al., 1997; La Rue et al.). The current study included specific cognitive domains and therefore it was possible to investigate the relationship between cognition and nutrition in elderly participants in more detail. Not all nutrients have the same influence on cognition (see Appendix I for a list of different categories of nutrients). Based on previous studies (see Table 19, pp. 93-95), it can be seen that the micronutrients, in particular B-vitamins and anti-oxidants, show the strongest relationship with cognitive measures (predominantly tasks involving reasoning and memory) in elderly populations. Moreover, many of these relationships have been reliably demonstrated, based on results from both blood samples and

[13] Presumably this view does not apply to nutrients where large doses cause toxicity, such as Vitamin A.

participants' reports of their diet. This suggests that self-report diet diaries are a valid and reliable method of providing dietary information.

3.4.4 Biomarkers

Chapter 1 detailed the predominant use of sensorimotor and physiological measures in biomarker research. This is due to their reliable association with outcome measures relevant to independent functioning (e.g. physical ability and cognition) and the fact that they are relatively quick and easy to measure. Of these biomarkers, systolic and diastolic blood pressure (physiological measures), and grip strength and visual acuity (sensorimotor measures) have particularly been widely used (Anstey et al., 1996). These traditional measures, in addition to body mass index (an anthropometric measure), have been included in the current study in order to evaluate the utility of a less commonly used type of measure, a cognitive biomarker. The proposal of working memory as a suitable candidate for a cognitive biomarker was outlined in Chapter 1. Each 'biomarker' measure will now be discussed in more detail.

3.4.4.1 Blood pressure

It is widely accepted that blood pressure increases with increasing age (Birren & Speith, 1962; Gardner & Poelham, 1997; Heron & Chown, 1967). This is particularly true for systolic blood pressure (Goldberg et al., 1995; Murray, 1951). Blood pressure is measured by millimetres of mercury (mm Hg), with the systolic reading listed first followed by the diastolic value. Systolic pressure is the pressure developed during contraction of the ventricles of the heart; diastolic pressure is the pressure that exists when the ventricles of the heart are at rest (Benjamin et al., 1997). Hypertension is the term employed for individuals who suffer from persistently high blood pressure. Individuals with blood pressure readings above 140/90 mm Hg are considered to be suffering from hypertension (Benjamin et al., 1997; Waldstein, 2003).

Table 19
Nutrient levels and their impact on cognitive functioning in elderly populations

Name of nutrient	Cognitive domain	Research findings
Thiamin	Abstraction Test (reasoning)	Higher self-report levels of dietary intake were positively associated with test performance (La Rue et al., 1997)
Riboflavin	Nonverbal abstract thinking (reasoning)	Low blood levels were associated with worse performance (Goodwin et al., 1983)
	Abstraction Test (reasoning)	Higher self-report levels of dietary intake were positively associated with test performance (La Rue et al., 1997)
Niacin	Abstraction Test (reasoning)	Self-report dietary intake was positively associated with performance on this test (La Rue et al., 1997)
	Rey-Osterrieth Recall Test (memory)	Self-report dietary intake was positively associated with performance on this test (La Rue et al., 1997)
Vitamin B-6	Backward Digit Span (working memory)	Higher plasma (blood) concentrations were associated with better performance (Riggs, Spiro, Tucker & Rush, 1996)
	Incidental memory	Higher plasma (blood) concentrations were associated with better performance (Riggs et al., 1996)
Folate	Nonverbal abstract thinking (reasoning)	Low blood levels were associated with worse performance (Goodwin et al., 1983)
	Abstraction Test (reasoning)	Plasma (blood) level, biochemical activity and dietary intake measures were all correlated significantly. Lower levels were related to poorer performance (La Rue et al., 1997)

Goodwin et al. (1983): sample age range was 60 - 94 years; La Rue et al. (1997): sample age range was 66 - 90 years;
Riggs et al. (1996): sample age range was 54 - 81 years

Table 19 (continued)
Nutrient levels and their impact on cognitive functioning in elderly populations

Name of nutrient	Cognitive domain	Research findings
Folate (continued)	Immediate and delayed recall (memory)	Low blood levels were associated with poorer performance (Hassing, Wahlin, Winblad & Bäckman, 1999)
	Spatial copying (visuo-spatial ability)	Low plasma (blood) concentrations were associated with poorer performance (Riggs et al., 1996)
	Global measure of cognition (i.e. MMSE)	Higher self-report dietary intake was associated with better cognitive function (Ortega et al., 1997)
Vitamin B-12	Nonverbal abstract thinking (reasoning)	Low blood levels were associated with worse performance (Goodwin et al., 1983)
	Immediate and delayed recall (memory)	Low blood levels were associated with worse performance (Goodwin et al., 1983)
	Spatial Copying (visuo-spatial ability)	Low plasma (blood) concentrations were associated with poorer performance (Riggs et al., 1996)
Vitamin C	Nonverbal abstract thinking (reasoning)	Low blood levels were associated with worse performance (Goodwin et al., 1983)
	Immediate and delayed recall (memory)	Low blood levels were associated with worse performance (Goodwin et al., 1983)

MMSE = Mini Mental State Exam
Goodwin et al. (1983): sample age range was 60 - 94 years; Hassing et al. (1999): sample age range was 90 - 96 years;
Ortega et al. (1997): sample age range was 65 - 90 years; Riggs et al. (1996): sample age range was 54 - 81 years

Table 19 (continued)
Nutrient levels and their impact on cognitive functioning in elderly populations

Name of nutrient	Cognitive domain	Research findings
Vitamin C (continued)	Rey-Osterrieth Copy Test (visuo-spatial ability)	Higher plasma (blood) concentrations were associated with better test performance; dietary intake of vitamin C approached significance for this same test (La Rue et al., 1997)
	Global measure of cognition (i.e. MMSE)	Higher self-report dietary intake was associated with less cognitive impairment (Paleologos, Cumming & Lazarus, 1998)
		Higher self-report dietary intake associated with better cognitive function (Ortega et al., 1997)
Vitamin E	Delayed recall (memory)	Lower serum (blood) levels were associated with poorer memory performance (Perkins et al., 1999)
	Global measure of cognition (i.e. z-score of four tests)	Self-report vitamin E intake, from foods or supplements, was associated with less cognitive decline with increasing age (Morris, Evans, Bienias, Tangney & Wilson, 2002).
Beta-carotene	Global measure of cognition (i.e. MMSE)	Higher self-report levels of dietary intake were associated with better cognitive performance (Jama et al., 1996)

MMSE = Mini Mental State Exam
Jama et al. (1996): sample age range was 55 - 95 years; La Rue et al. (1997): sample age range was 66 - 90 years;
Morris et al. (2002): sample age range was 65 - 102 years; Ortega et al. (1997): sample age range was 65 - 90 years;
Paleologos et al. (1998): sample age range was 69 - 91 years; Perkins et al. (1999): sample age range was 60 - 80+ years

Hypertension of unknown aetiology ('essential hypertension') represents 90% of cases (Benjamin et al.; Waldstein). 'Secondary hypertension' is the term used when the source of the hypertension is known. For example, hypertension is often associated with the presence of type II diabetes. In general, hypertension is also a risk factor for stroke and cardiovascular disease (Goldberg et al., 1995; Waldstein). This association between blood pressure and health adds support to the inclusion of health as a covariate in the current study.

In addition to health complications, hypertension has been shown to have a negative impact on cognitive performance (Goldberg et al., 1995; Waldstein, 2003). A review by Waldstein, Manuck, Ryan and Muldoon (1991) indicated that hypertension has a particularly strong influence on tests of reasoning and memory. By contrast, Anstey and Christensen (2000) suggested that the association between elevated blood pressure and impaired cognitive performance is most consistently found for measures of speed of processing and global measures of cognition. Conversely, measures reflecting crystallised and verbal ability have appeared less reliable for differentiating individuals with hypertension from those without (Waldstein et al.).

In studies investigating independent functioning, higher readings of systolic and diastolic blood pressure have been associated with increased dependence in activities of daily living (Ensrud et al., 1994). Similarly, the absence of high blood pressure was associated with maintenance of mobility (e.g. going up and down stairs unaided, walking half a mile unassisted) over a 6 year period in a study by Guralnik, et al. (1993). Results from several other studies have supported this negative relationship between hypertension and mobility and daily activities (e.g. Guralnik & Kaplan, 1989; Harris et al., 1989; Mor, et al., 1989).

In the current study, blood pressure (both systolic and diastolic) was measured

using an automatic blood pressure monitor (sphygmomanometer; Omron T5). Brachial blood pressure was measured three times, with administration consistent with the instructions in the manual (Omron, 2001)[14]. Average systolic and diastolic blood pressure was calculated from these readings and this continuous score was employed as the blood pressure variable (blood pressure categories are detailed in Chapter 5 for descriptive purposes). Each of the three readings was taken about five to 10 minutes after the preceding reading. The initial reading was taken approximately 25 minutes into the test session to minimise the effects of stress or rushing around prior to commencement of the test session. Any tasks administered before the first blood pressure reading was taken were relatively easy, or perceived to be less complicated. Many things can cause blood pressure to fluctuate including exercise, mental stress and talking (Omron, 2001).

Therefore, the procedure outlined above, including taking multiple readings, was necessary in order to obtain the most reliable blood pressure reading possible. In the current study, the test-retest reliabilities of blood pressure over both 6 months (systolic: $r = .69$; diastolic: $r = .71$) and 18 months (systolic: $r = .69$ and diastolic: $r = .67$) were moderate (also see Chapter 6). These reliability values are greater than those reported by Anstey, Smith and Lord (1997) over a 3 month interval and, in contrast to their results, suggest that both types of blood pressure have acceptable reliability.

3.4.4.2 Grip strength

Muscle mass and strength decline with increasing age (e.g. Goldberg et al., 1995). In particular, hand-grip strength is reduced in older adults (Fone & Lundgren-Linquist, 2003; Murray, 1951). Grip strength is dependent on both forearm strength and finger joint function (Judge et al., 1996). Loss of muscle strength has

[14] Participants were asked to relax and sit with feet flat on the floor; the blood pressure cuff was positioned so that it was approximately level with the heart

adverse implications for activities of daily living (ADL), including balance, walking and many tasks of dexterity. Evidence of this can be found in several studies where reduced grip strength has been associated with increased mobility problems (Femia, Zarit & Johansson, 1997), increased dependence in instrumental ADL (Judge et al., 1996) and greater decline in both instrumental and basic ADL over a 3 year period (Ishizaki et al., 2000). Moreover, several studies have reported a positive relationship between grip strength and cognition (Anstey, 1999; Anstey et al. 1993; Clark, 1960; Heron & Chown, 1967; Salthouse, Hambrick & McGuthry, 1998).

Grip strength was measured in units of kilograms (kg) of force by a 'My-Gripper' squeeze dynamometer (made in Japan, refer to Figure 7, p.99[15]). Participants were asked to stand with their arms straight (but flaccid) and by their side. The 'My-Gripper' was then placed in one hand and participants were asked to squeeze the device as hard as possible, without raising their arm or causing pain. For improved reliability (as per the measurement of blood pressure), three readings were taken for each hand, alternating hands each time to avoid muscle soreness and fatigue. Average grip strength was calculated for each hand and a record of the dominant hand was also made. The average score from the *strongest* hand was employed as the grip strength variable as this measure was considered to best reflect daily functional capacity (i.e. the dominant hand was not always the strongest hand and thus the strongest hand conceivably would be called upon in day-to-day tasks). Anstey et al. (1997) reported an "acceptable" test-retest reliability of $r = .84$ for left grip strength and $r = .81$ for right grip strength, over a 3 month period. In the current study, test-retest reliability for grip strength was very high over 6 ($r = .96$) and 18 ($r = .94$) months; also see Chapter 6.

[15] The manufacturer of this device is unknown; other studies reporting use of this device (e.g. Jensen, Andersen & Jensen, 1997) have also not included this information.

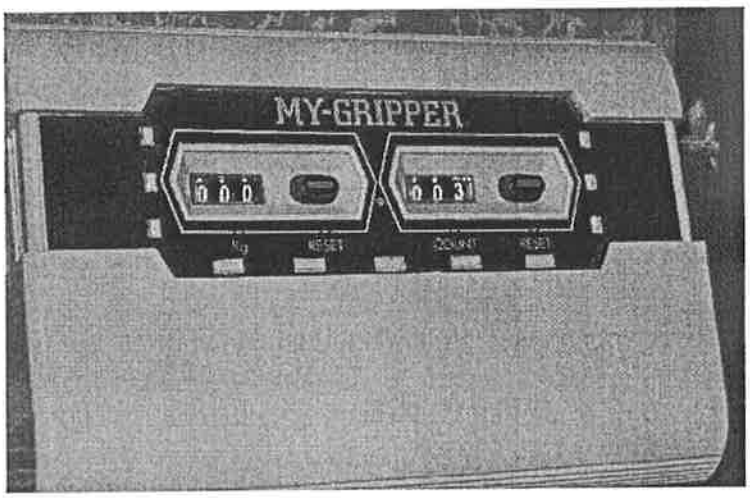


Figure 7 The 'My-Gripper' grip strength device used in the current study
Device measurements: length = 10.3 cm, width = 6.6cm, depth = 2.2 cm.
(The dark area compresses when squeezed).

3.4.4.3 Body mass index

Body mass index (BMI) was calculated from height and weight measurements; BMI = weight (in kilograms) divided by height (in metres) squared. A Soehnle electronic measurement scale was used to measure weight; height was measured by tape measure with the participant standing against a wall. Participants were encouraged to stand as upright as possible when measuring height. Both height and weight were measured without participants wearing shoes and heavy clothing, if applicable (e.g. removing a heavy jacket if in colder weather). As mentioned earlier in the section on nutrition, BMI is frequently used as an overall measure of nutrition, in particular, as an indicator of malnutrition (Bédard, Molloy, Bell & Lever, 2000; Galanos, Pieper, Cornoni-Huntley, Bales & Fillenbaum, 1994). Needless to say, it is likely that BMI is an indicator of several factors, including education/socio-economic status and health. In the current study, BMI was used as a continuous variable and as a separate measure to nutritional status. Descriptive statistics on BMI category membership are provided in Chapter 5. Test-retest reliability of BMI in the current study was high; r = .97 over 6 months and r = .95 over 18 months. Studies reporting test-retest reliability for BMI could not be found.

Generally, increased functional dependence has been associated with both high and low BMI (Stuck et al., 1999). Moreover, BMI has also been shown to have a u-shaped relationship with various outcome measures, including physical ability, falls and incontinence (Galanos et al., 1994; Harris et al., 1989; Tinetti, Inouye, Gill & Doucette, 1995). Moreover, much research has established the impact of obesity on health and functioning; both Ensrud et al. (1994) and Sulander et al. (2005) found that individuals with a higher BMI were more likely to experience disability or reduced functional independence. Furthermore, obesity was associated with a 20 per cent increased risk of losing mobility in males and a 40 per cent increased risk in females, compared to individuals with a normal to overweight BMI (LaCroix et al., 1993). A similar association has been found between obesity and cognition. For example, in a longitudinal study, Whitmer, Gunderson, Barrett-Connor, Quesenberry and Yaffe (2005) found that "the prevalence of a subsequent diagnosis of dementia was significantly higher for those who were obese or overweight at mid-life" (p.2). This result persisted when controlling for demographic and health variables and was stronger in women than men. Obese women (BMI $\geq$ 30) were twice as likely to develop dementia as women of normal weight (BMI = 18.6 – 24.9).

It is interesting to note that results from the studies by LaCroix et al. (1993) and Whitmer et al. (2005) do not support a u-shaped relationship between BMI and various outcome measures; they found no significant negative association between low BMI and mobility or dementia, respectively. This was therefore in contrast to a study by Bédard et al. (2000) who found that low BMI (BMI < 21) was associated with limitations in activities of daily living and poor cognition (as measured by a global measure of cognition). Even though this result was found in patients with moderate dementia of the Alzheimer's type, the prevalence rate of low BMI was the same as in

the general elderly population (Bédard et al., 2000), suggesting that similar results may be found in populations of cognitively intact individuals.

3.4.4.4 Visual acuity

"Acuity refers to the ability of the visual system to resolve fine spatial detail" (Schieber, 1992, p. 255). There appears to be a gradual decline in far distance visual acuity with increasing age (McFarland, 1968); uncorrected far distance visual acuity declining between the ages of 30 and 80 years, and corrected far distance visual acuity declining from 55 to 60 years of age (Gittings & Fozard, 1986; Pitts, 1982). It has been suggested that a substantial portion of age-related decline in visual acuity is the result of neural deterioration beyond the organ of the eye itself (Lindenberger & Baltes, 1994). This idea will be discussed further in regard to the relationship between visual acuity and cognition. Meanwhile, sensory abilities are pivotal to the capacity to perform daily activities independently. An obvious illustration of this relationship is driving. Impaired vision (or hearing for that matter) can make driving more difficult, and possibly dangerous. Generally, Marsiske, Klumb and Baltes (1997) and Stuck et al. (1999) found that vision was a significant predictor of ADL. More specifically, Branch, Horowitz and Carr (1989) found that individuals reporting visual impairments were more likely to indicate dependence in activities such as shopping and paying bills. Moreover, visual decline has been associated with decreased likelihood of living independently in old age (Salive, Guralnik, Glynn, Christen & Wallace, 1994).

A positive relationship between visual acuity and cognition also exists (Baltes & Lindenberger, 1997; Marsiske et al., 1997; Salthouse, Hancock, Meinz & Hambrick, 1996). This association between a sensory measure and cognition is frequently described as illustrative of the 'common-cause hypothesis' (Anstey, 1999; Baltes & Lindenberger; Lindenberger & Baltes, 1994). That is, the positive relationship between

sensory impairment (e.g. poor visual acuity) and cognitive impairment (e.g. slower speed of processing) has been found to strengthen with increasing age. In older populations (e.g. over 70 years), sensory impairment can explain a large proportion of variance in various tests of cognitive ability. It is considered that "individual differences in visual [and auditory] functioning are not only peripheral phenomena but also reflect age-based changes in the central nervous system" (Baltes & Lindenberger, 1997, p.12). Age-based changes in the central nervous system are also believed to influence cognitive ability, thus linking deterioration of both cognition and sensory ability to age-related decline in the central nervous system (Baltes & Lindenberger, 1997).

In the current study, far distance visual acuity was measured by a standard Snellen chart at a distance of 4 metres from the participant. The Snellen chart consisted of eight rows of letters of varying sizes. A distance of 4 metres is within the range used by other studies (e.g. 6 metres: Anstey & Smith, 1999; Park et al., 2002; and 2.5 metres: Baltes & Lindenberger, 1997; Marsiske et al., 1997). It was considered that a distance of 6 metres may not be achievable in many participants' homes (rooms may be quite small depending on the type of residence), whereas 2.5 metres as viewing distance may result in ceiling effects. Participants were instructed to read the chart with both eyes (binocularly), starting with the top row (largest letters) and continuing to the bottom row (smallest letters) or until they were unable to repeat correctly more than half the letters in a row. The same procedure was then completed with the right eye alone and then the left (monocularly).

Prior to starting, the participant was instructed to wear any glasses that may improve their vision at this distance (corrected vision). This instruction was given for two reasons; firstly, corrected vision is more representative of the acuity that individuals use in daily activities (Marsiske et al., 1997), a primary outcome measure of

independent functioning in the current study; and secondly, corrected vision is likely to minimise the influence of peripheral variance (i.e. individual differences in the refractory properties of the lens), and to maximise the variance due to more central processes [i.e. sensory loss stemming from age-related decline in the central nervous system (Anstey & Smith, 1999; Baltes & Lindenberger, 1997; Lindenberger & Baltes, 1994)]. The participant's initial score reflected the size of the letter correctly repeated at a distance of 4 metres; lower scores denoted better visual acuity. To be consistent with other measures used in the test battery, the visual acuity score was transformed so that larger scores reflected better visual acuity. This transformation was [100 subtract the participant's initial score]. Anstey et al. (1997) reported a test-retest reliability of r = .82 for corrected visual acuity over a 3 month interval. As will be discussed in Chapter 6, visual acuity in the current study demonstrated poor test-retest reliability, particularly over 18 months (r = .47 compared to r = .65 over 6 months).

3.4.4.5 Working memory

It is hypothesised that working memory should predict various independent functioning outcome measures, including activities of daily living, better than chronological age. As discussed in Chapter 2, working memory tasks were primarily selected to represent the content domains of the three components of Baddeley and Hitch's (1974) tripartite model. Reading Span was selected as a measure of verbal working memory (employing the phonological loop); Dot Matrix was selected as a measure of visuo-spatial working memory (employing the visuo-spatial sketchpad); and Digit Ordering was selected as a measure of central executive functioning. The specific details of these tasks will now be described.

Reading Span involved three main steps. Firstly, the participant was instructed to read a sentence on a computer screen. Reading the sentence out loud was encouraged.

This was done at the participant's own reading speed, with the sentence disappearing only when the participant pressed a key (the space bar) on the keyboard. Secondly, once the participant had read the sentence s/he was asked to answer a question about that sentence. This question related to the content of the sentence and a multiple-choice answer (three alternatives) was provided on the computer screen. Again, the participant responded by pressing a designated key on the keyboard. There was no time limit for answering the question and if the participant was unsure of the answer, s/he was encouraged to guess. Thirdly, the participant was asked to write down the final word of the sentence (in the correct order if more than one sentence) in a response booklet. The participant was not given an actual time limit for final word recall; however a 4 second limit per word (as employed by Salthouse & Babcock, 1991) was used as a guide. These three steps are illustrated in Figure 8.

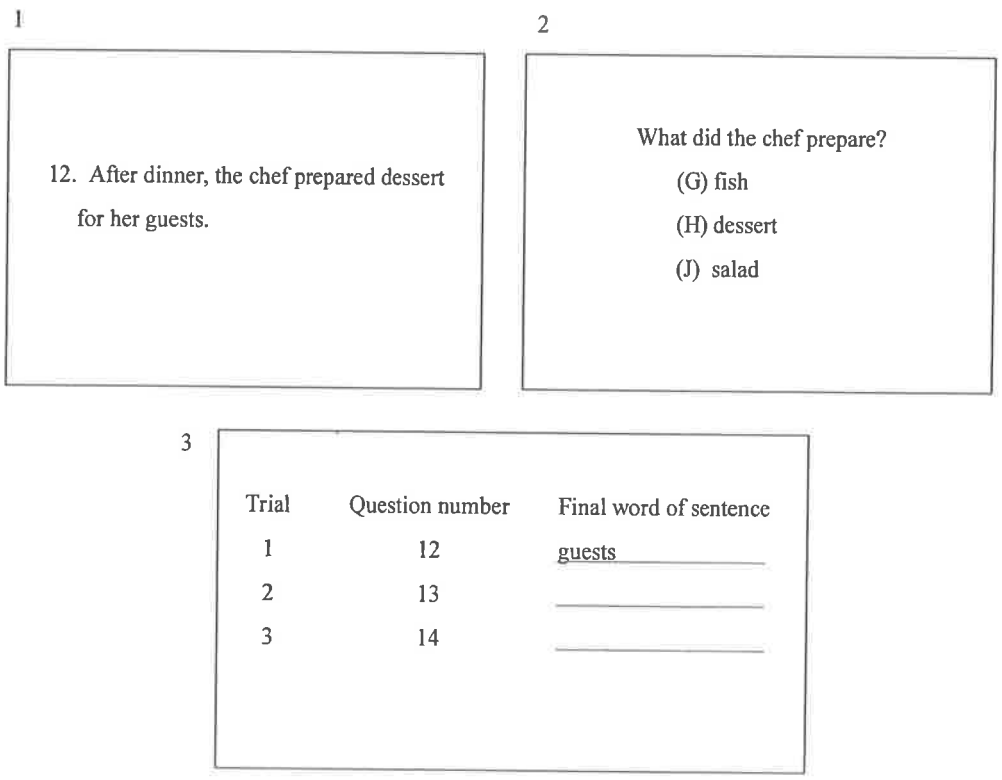


Figure 8 The three main steps in Reading Span
(1) reading the sentence; (2) answering a question about the sentence (answer is H); and (3) writing down the final word of the sentence.
Boxes 1 and 2 appear on a computer screen and box 3 is in a paper & pen answer booklet.

The answer-response keys of G, H, and J were selected because of their central and perceived easy-to-find location on the QWERTY keyboard. The letter representing the correct answer was randomly distributed. The computer recorded the answer to the multiple-choice question and the test administrator monitored the correctness of the final word recall. For a trial to be considered correct, both the answer to the question and the final word needed to be correct. The procedure outlined in Figure 8 continued at the same sentence-set length (i.e. one sentence) for three trials. After this, the number of sentences presented in a set increased to two. That is, a sentence and related multiple-choice question was followed by another sentence and multiple-choice question. The final word from each sentence could only be recorded after both sentences were read and their respective multiple-choice questions answered. These steps are illustrated in Figure 9a (below) and 9b (p.106). The second sentence appeared as soon as the participant had answered the question relating to the first sentence.

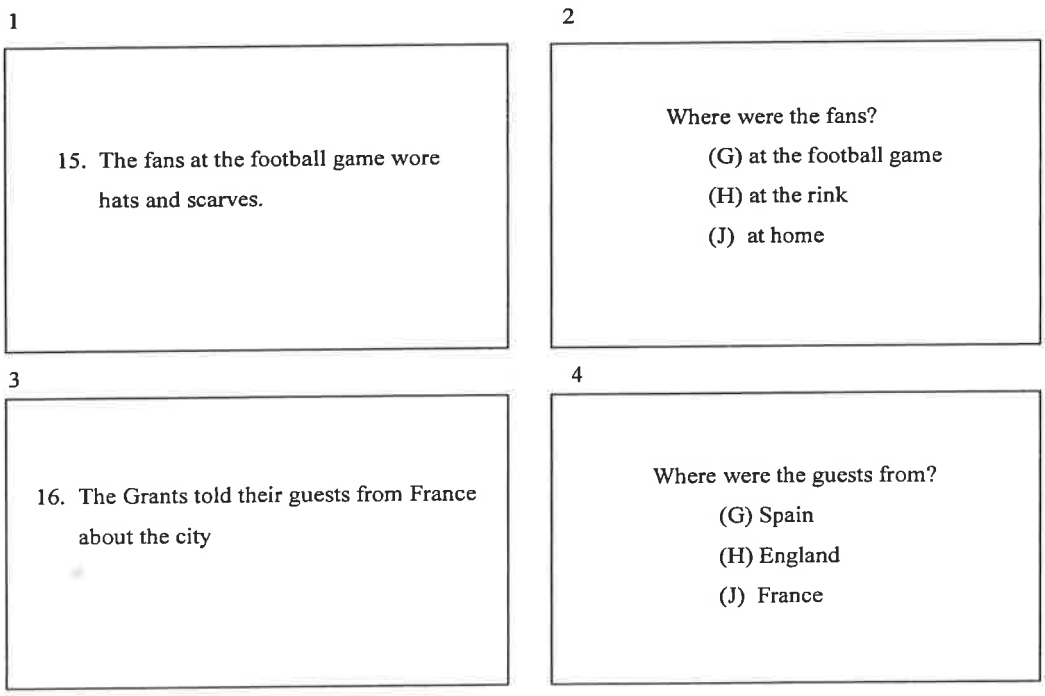


Figure 9a Sentence presentation: An example of a two sentence-set in Reading Span

Boxes 1 to 4 are shown on the computer screen and correspond to the steps of (1) reading the sentence (boxes 1 and 3) and (2) answering the multiple-choice questions (boxes 2 and 4). Answers to the questions are G and J, respectively.

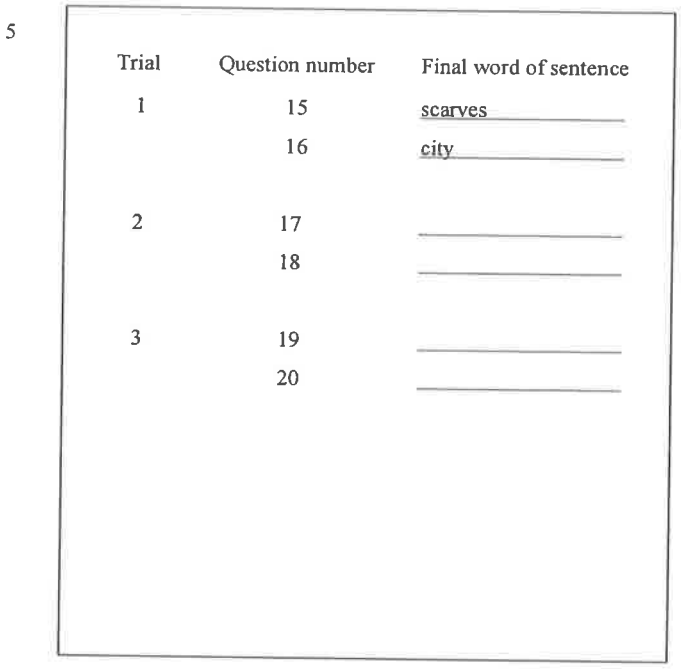


Figure 9b Final word recall: An example of a two sentence-set in Reading Span
Box 5 represents the final word recall component of the task, or step (3) as described on p.104. This is completed in a booklet. Each final word may only be written after the second multiple-choice question is answered.

This pattern continued with the sentence-set length increasing by one after every three trials (minimum set length was one sentence; maximum set length was four sentences). The task was discontinued when the participant failed to recall the final word from two out of the three trials correctly (answers to the multiple-choice questions were assessed post hoc). Participants were given practice sets of both one and two sentence-set lengths and were informed that the number of sentences and hence final words to remember would increment, as just described. Participants were encouraged to answer both the question and recall components correctly. Number correct was used as the scoring method with higher scores denoting better performance[16]. The maximum score is 30.

The sentences used in Reading Span were obtained with permission from

[16] As discussed in Chapter 2 (pp.45-46), the use of a continuous score, such as number correct, is favourable for working memory tasks. The use of number correct is also consistent with the scoring of other tasks in the test battery.

Professor Timothy Salthouse[17] and specifications for these sentences have been described by Salthouse and Babcock (1991). Specifications included: sentences between 6 and 10 words in length; final words contained no more than two syllables; final words were familiar and could be found in a children's dictionary; no word appeared more than once as the final word of a sentence; questions about the sentences were relatively simple (e.g. who, when, where); and neither questions nor multiple-choice answers contained sentence final words.

The administration and methodological aspects of the Reading Span task used in the current study are consistent with those used by many other researchers (for a detailed summary, see Chapter 2). Test-retest reliabilities on reading span-type tasks have ranged from r = .66 over a period of several months (Waters & Caplan, 1996) to r = .73 over the period of a week (Tirre & Peña, 1992). Other reliability measures for these tasks, including internal consistency (Cronbach's α = .84; Oberauer et al., 2000) and split-half reliability (r = .88; Park et al., 2002) have been moderate to strong. In terms of concurrent criterion validity, Waters and Caplan (2003) demonstrated that a reading span-type task had moderate correlations with other tasks of verbal working memory (e.g. r = .61 with a digit span backwards task). Test-retest reliability of Reading Span in the current study was poor to moderate but increased over a longer time interval (r = .58 over 18 months, compared to r = .56 over 6 months).

The format of the Dot Matrix task was very similar to that of Reading Span. Instead of sentences, the task used visuo-spatial material. Also, the order of steps differed slightly from Reading Span. For Dot Matrix, participants were asked to complete the following three steps; (1) answer a true or false picture puzzle (illustrated in Figure 10, p.108); (2) remember the position of a single black dot in a grid; and (3)

[17] Reading Span stimuli were obtained from Prof. Salthouse's website; the password to obtain these was exchanged via email circa 12/05/03.

draw the dot in the position it appeared in a grid provided. The picture puzzles comprised line segments in certain positions that needed to be either added together, or one subtracted from the other (achieved via mental imagery). The solution from this process then had to be compared with the puzzle solution provided. If participants considered their image to match the one provided, then they pressed the letter T (true) on the keyboard; if their image did not match the solution provided, then they pressed the letter F (false). The picture puzzle stayed on the screen until the participant entered their response.

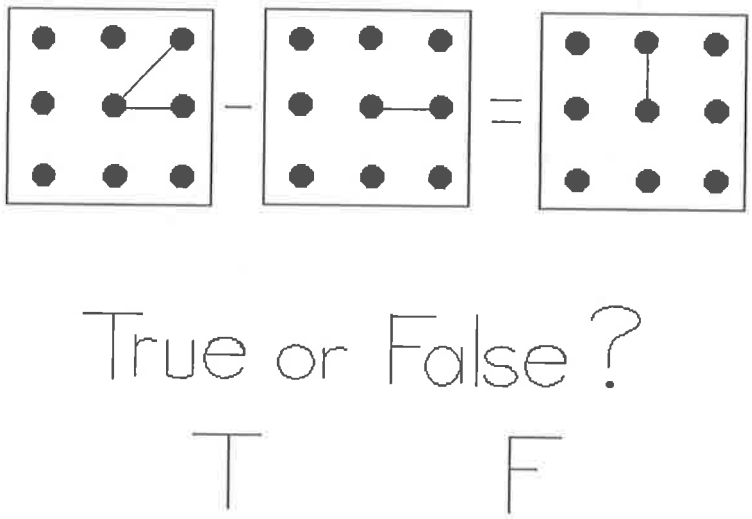


Figure 10 The information processing component of Dot Matrix
This is both viewed and answered on computer. Sums could either be additions (+) or substractions (-). This is an example of the latter and the answer is false.

Following the true or false response to the picture puzzle, a screen appeared with a 5 x 5 square grid on it (see Figure 11, p. 109). One of the squares within the grid contained a single black dot. Participants were instructed to remember the position of this dot. The grid (and dot) remained on the screen for approximately 4 seconds. After it disappeared, participants were instructed to replicate the dot's position in a booklet of blank 5 x 5 square grids. The number of trials, the range of puzzle and dot-set lengths, time limits, scoring procedures and practice were as described in Reading Span.

The material used in the Dot Matrix task was obtained with permission from Dr. Akira Miyake[18]. The method used in the current study varied slightly from his method (refer to Miyake et al., 2001). Picture puzzle solutions were presented randomly with approximately half of the solutions being true and half being false. In terms of internal consistency, a similar Dot Matrix task had a Cronbach's $\alpha = .79$ (Miyake et al., 2001). In terms of concurrent criterion validity, visuo-spatial working memory tasks (e.g. line span and letter rotation) employed by Park et al. (2002) correlated moderately with other measures of working memory, such as a reading span task (i.e. $r = .54$ and $r = .62$, respectively). Psychometric properties for the current Dot Matrix task are not available (see Chapter 4 for an explanation).

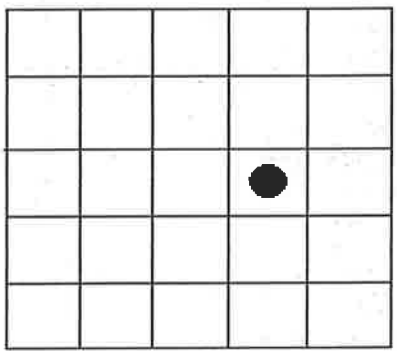


Figure 11 The information storage component of Dot Matrix
This is viewed on computer and later recorded in a booklet. For puzzle/dot-set lengths greater than one, each picture puzzle was followed by a single grid with only one dot in it. That is, participants answered a picture puzzle and saw one dot in a grid; they then answered a second picture puzzle and saw another single dot in a grid. Participants had to wait until the second dot disappeared before they could draw both dot positions in the booklet.

Digit Ordering involved the administrator reading out loud a sequence of digits, varying in length from two digits to eight. The participant was instructed to repeat the digits back, first re-arranging them so that the digits were in ascending order. For example, for the digit sequence '8, 4, 7, 3' the correct answer is '3, 4, 7, 8'. In keeping with the previously mentioned working memory tasks, there were three trials at each set

[18] Upon request, stimuli were obtained via email circa 26/05/03 from Dr. Miyake

length. This resulted in a total of 21 digit sequences (provided in Appendix J). The digits for these sequences were obtained from two sources. Some came from Werheid et al. (2002, p.561); and the remaining sequences were randomly generated by a computer. Digit sequences in the current study did not include any zeros or double-digits, thereby excluding some of Werheid et al.'s sequences. This decision was made based on the results of MacDonald et al. (2001), who found that the inclusion of extra digit names and harder trials, such as those including the digits 0, 10, 11 and 12, were "not necessary for testing elderly adults or Alzheimer Disease patients, as these participants rarely approach ceiling on this task" (p.23). Digits were presented at the rate of one digit per second. Furthermore, digits were read out "with equal stressing and prosody", with a lowered pitch on the last digit in each sequence (Werheid et al., p.561). The task was discontinued if an error was made on two out of three trials at the same digit sequence length. One point was allocated for every complete digit sequence ordered correctly (maximum score is 21).

Digit ordering tasks have largely been employed as tests of working memory in clinical populations (see Chapter 2 for details). Therefore this task's psychometric properties in 'normal' populations are of particular interest. Digit ordering performance in the control group of Hoppe et al.'s (2000) study produced a split-half reliability of $r = .64$ and a Cronbach's $\alpha = .61$. Moreover, Werheid et al. (2002) described their digit ordering task as having sufficient reliability (Spearman-Brown split-half reliability of $r = .75$). In terms of concurrent criterion validity, digit ordering tasks have been found to correlate moderately with more widely utilised tasks of working memory such as a digit span backwards task, $r = .47$ (MacDonald et al., 2001) and a two-back task, $r = -.33$ (Hoppe et al.). Psychometric properties for the current Digit Ordering task are not available (also see Chapter 4).

3.5 Procedure

The test battery was administered individually for all three measurement occasions. Participants were contacted by phone and a 4 hour meeting was arranged, usually at the participant's home. A total of 29 participants elected to come into the University to be tested[19]. A week or so prior to the test session taking place, a package of questionnaires (all questionnaires were printed in a large font size for easy reading) was sent out and the participants were asked to complete them, ready for collection at the test session (this occurred on initial and final measurement occasions). A contact number was provided in case participants had any queries regarding the questionnaires and in case session times needed to be cancelled or rescheduled.

The order of tests in the battery was compiled such that tests perceived to be more difficult were evenly distributed across the session with 'easier' tasks being administered before and after 'difficult' ones. This was orchestrated in order to maintain motivation and general enjoyment. Similarly, tests from the same domain (e.g working memory, crystallised ability) and type of administration (i.e computer, pen-and-paper, verbal) were spaced out across the test session. The presentation order of tests was as follows and remained the same for each participant: ADAS-Cog, blood pressure, Spot-the-Word, Digit Symbol*[20], Reading Span, Raven's Standard Progressive Matrices[21], weight, height, Digit Ordering, Similarities, Pattern Comparison*, Concept Formation, Visual Matching*, Dot Matrix, grip strength, visual acuity, Culture Fair, and Information. The total test battery duration was estimated at 3 hours and 40 minutes, per person for initial and final assessments and 2 hours for the intermediate assessment

[19] The characteristics of these two groups will be explored in Chapter 4.
[20] Data for the current study were co-collected with another PhD student, whose thesis focused on speed of information processing; tasks marked with an asterix were of interest to her. However, some of these data are utilised (and described) in Chapter 4. Arguably, processing speed could provide a strong explanation for results reported here (see Introduction, p.16).
[21] Note that order for alternative forms was not counterbalanced.

(because biomarkers only were measured on that occasion). Given the nature of the tests and the age of the participants, a break was suggested at the half way point or taken at other times upon request of the participant.

Considering this dissertation proposed the construct of working memory to be a cognitive biomarker of ageing, the validity of the working memory measures presented here and their relationship with age require further exploration. Chapter 4 presents these results.

CHAPTER 4: WORKING MEMORY TASKS
FACTOR STRUCTURE AND AGE DIFFERENCES

4.1 Overview: Analysis of working memory tasks

Prior to presenting cross-sectional results analysing whether various biomarkers correlate with functional outcome measures in elderly people, two aspects of the working memory tasks are first explored; construct validity and age differences. The first consideration is whether the tasks selected as working memory tasks actually measure working memory. Exploratory factor analysis tests whether the working memory tasks used in the current study (Reading Span, Dot Matrix and Digit Ordering) load on one factor, which is independent from factors loaded by other types of memory tasks. The second issue addressed in this Chapter is that of establishing age differences in performance on the working memory tasks. The extent of age differences will be discussed, in addition to investigating which components of the Baddeley and Hitch (1974) working memory model are responsible for age-related deterioration.

4.2 Working memory: Construct validity

The construct validity of the working memory tasks was established using initial test session data via examination of inter-correlations and by factor analysis[22] with other memory measures. The logic behind this approach was that, if the three working memory tasks correlated significantly and highly with one another, then this would be accepted as demonstration that all three tasks were measuring the same construct. Secondly, exploratory factor analysis was carried out on the three working memory tasks, together with two subtests from the Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog); word recall and recognition. These memory tasks measure

[22] All statistical analyses were conducted using the statistical computer program SPSS, version 12.1

short-term memory, or primary memory according to Craik's (2000) short-term memory processing continuum. If the ADAS-Cog subtests and working memory tasks loaded on different factors then this would support the proposition that the three working memory tasks were assessing a separate memory construct.

All working memory tasks correlated significantly with each other (Table 20). However, Reading Span and Dot Matrix showed a moderate correlation whereas Digit Ordering showed weaker correlations with both of these tasks. Park et al. (2002) reported a similar correlation (i.e. r = .64) between a reading span task and a visuo-spatial working memory task (letter rotation). Moreover, many studies using reading span-type tasks have reported similar inter-correlations with other working memory tasks, ranging from r = .50 to r = .66 (Kirasic et al., 1996; Park et al., 1996; Salthouse & Babcock, 1991).

Table 20
Inter-correlations of the working memory tasks (initial test session data)

	N	1.	2.	3.
1. Reading Span	149	-	.64	.29
2. Dot Matrix	63	.64	-	.38
3. Digit Ordering	149	.37	.38	-

All correlations are significant; $p < .01$
Values below the diagonal are correlations when sample is listwise for Dot Matrix (N = 63)

The correlations between Digit Ordering (a verbal, numeric task) and Reading Span (r = .29) and Digit Ordering and Dot Matrix (r = .38) were substantially lower than the correlation between Reading Span and Dot Matrix. This may be due to Reading Span and Dot Matrix tasks having very similar administration. The less complicated instructions of Digit Ordering, compared to the other tasks (as discussed in Chapter 2), may have influenced performance, with evidence that participants generally found the Digit Ordering task less difficult (see Table 21, p.115; measure is number correct with higher scores denoting better performance).

Table 21
Descriptive statistics for working memory tasks (initial test session data)

Task	N	Median	Mean	SD	Range	Maximum
Reading Span	149	13.00	13.91	8.81	0 - 30	30
Dot Matrix	63^	7.00	9.03	5.43	1 - 25	30
Digit Ordering	150	15.00	14.90	3.10	9 - 21	21

^The reasons behind this smaller sample size are discussed in Appendix K.

Participant feedback about the Digit Ordering task suggested that some participants used a 'strategy' that may have mitigated dependence on working memory. For example, for trials including seven or eight digits arranged in ascending order, some participants simply remembered the digit(s) omitted and ordered the sequence accordingly, rather than remembering the complete sequence and re-ordering the numbers, as per previous trials. In other words, because only the digits 1 through 9 were used as stimuli, if a trial involved listening to and re-arranging eight digits, then only one digit was not included in that trial. Therefore, by remembering the digit not mentioned, re-arranging the other eight digits could be completed without a significant reliance on working memory. Waters and Caplan (2003) made a similar observation with a working memory task ('missing digit') used in their study. Moreover, Waters and Caplan reported poor inter-correlations between the 'missing digit' task and other tasks of working memory (e.g. sentence span). It is also possible that the difference in magnitude among correlations reflected the different content domains of the tasks. For example, Digit Ordering involves numeric stimuli whereas Reading Span and Dot Matrix both incorporate verbal stimuli[23]. In any event, correlations of lower magnitudes (i.e. $r < .40$) are not uncommon among working memory tasks (Oberauer et al., 2000; Waters & Caplan). Therefore, on balance, these results were accepted that the three tasks were converging on a common construct of working memory.

For the exploratory factor analysis, the three working memory tasks and two

[23] For the memory storage component of Dot Matrix (i.e. the position of the dot in the grid), the visuo-spatial material could also have been encoded verbally, in the sense of 'row 1, column 5', producing retrieval of information that was not strictly based on visuo-spatial ability/memory.

short-term memory tasks were subjected to principal components analysis (PCA). Before PCA was conducted, the suitability of the data for factor analysis was assessed. Three aspects were satisfied that supported a factor solution; firstly, the correlation matrix contained coefficients of .3 and above; secondly, the Kaiser-Meyer-Oklin value was above .6 (.66); and thirdly, Bartlett's Test of Sphericity was significant ($p < .01$). PCA revealed the presence of two components with eigenvalues greater than one, explaining 45.8% and 22.1% of the variance, respectively. Moreover, inspection of the scree plot (see Figure 12) supported retaining two components for further analysis[24].

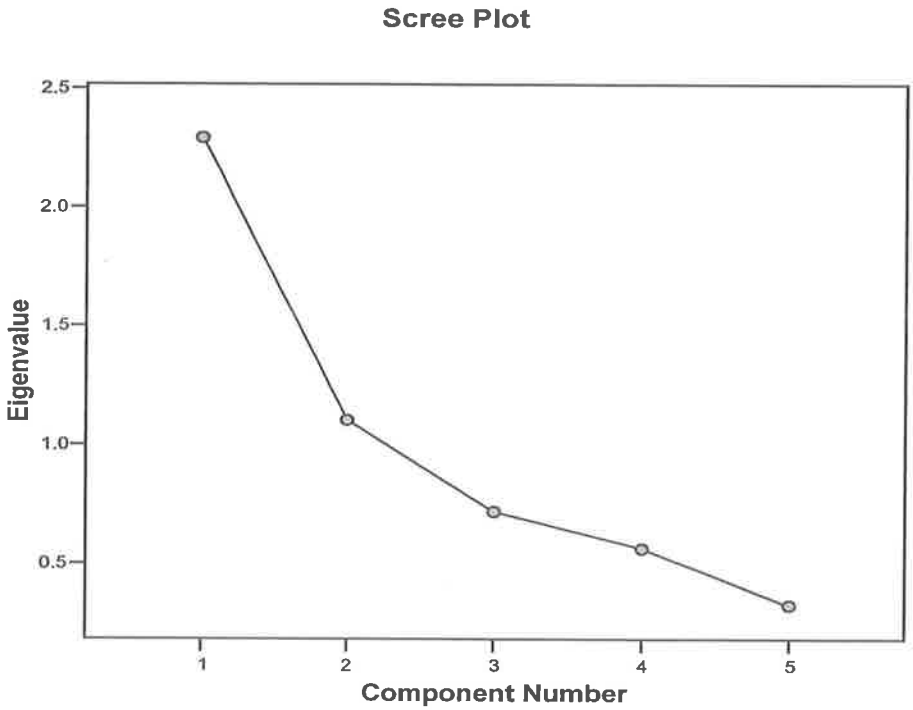


Figure 12 Evidence of two memory components from exploratory factor analysis

[24] In recent times, commonly used techniques for determining the number of components in factor and principal component analyses (PCA), such as eigenvalues-greater-than-one rule and the scree test, have been considered 'flawed' (O'Connor, 2000, p.396). Techniques deemed more reliable, such as parallel analysis and Velicer's minimum average partial (O'Connor) were considered for the current PCA. However, upon discussion with a statistician knowledgeable in this area (Dr. Nick Burns), these more reliable techniques were judged unnecessary, based on the simplicity of the PCA here (i.e. five variables and an easily-interpretable scree plot).

Varimax rotation[25] was performed to assist interpretation of the components. The rotated solution (Table 22) revealed strong loadings on both components, with variables loading substantially on one component. This two-factor solution explained 67.9% of the variance, with Component 1 contributing 42.3% and Component 2 contributing 25.6%. Consistent with previous research, recall and recognition memory tasks represented short-term memory; and Reading Span, Dot Matrix and Digit Ordering represented working memory. Thus, results from both the inter-correlations and factor analysis therefore supported an interpretation that Reading Span, Dot Matrix and Digit Ordering tasks were assessing working memory ability.

Table 22
Principal components analysis: Component loadings from the varimax rotation of the two-factor solution for memory tasks

Memory Item	Component 1 Working Memory	Component 2 Short-term Memory
Reading Span	**.82**	-.22
Dot Matrix	**.83**	-.02
Digit Ordering	**.70**	-.00
Recognition (mean no. of errors)	.09	**.92**
Recall (mean no. of errors)	-.49	**.62**

4.3 Working memory: Age differences

Working memory task selection can be based on content domain or structure and function. From a theoretical perspective, Reading Span, Dot Matrix and Digit Ordering can be considered to reflect different content domains of Baddeley and Hitch's (1974) tripartite model (verbal, visuo-spatial and attentional, respectively; see Chapter 2 for details). In contrast, from a concrete standpoint, these tasks reflect general working memory capacity (as shown by the principal components analysis)[26]. Therefore, investigation of the relationship between each content domain of working memory and

[25] This analysis was repeated using oblique rotation. The results were substantially the same, although the version reported here provided a better fit.
[26] In order to demonstrate components from factor analysis that represent each content domain, multiple tasks from each domain would be required. This was not achievable in the current project due to time constraints.

age was limited. As a result, investigation into distinguishing storage (i.e. structure) and processing (i.e. function) aspects of working memory was carried out.

Before exploring which of these two working memory aspects explains age-related deterioration on working memory tasks, it is pertinent to establish the relationship between age and the working memory tasks used in the current study. Tables 23 to 25 (p.119) show correlations between each working memory task and age, across all three assessment occasions. Correlations are also provided with the influence of sex and years of education have statistically removed (partialed out) from the relationship between age and working memory.

As reported in Chapter 2, age-related deterioration on tasks of working memory can be considered a robust finding. Results of the current study are generally in line with this position. Both Reading Span and Digit Ordering performance demonstrated significant, negative correlations with age, even when statistically controlling for the influence of sex and years of education on working memory performance (see Tables 23 and 24). Dot Matrix failed to correlate significantly with age on any of the measurement occasions (despite an increase in sample size on subsequent sessions), ruling out examination of the differential decline hypothesis (described in Chapter 2, p.49). This lack of correlation with age may be due to substantial floor effects on Dot Matrix (see descriptive statistics provided earlier in Table 21). It is interesting to note that on the final test occasion, none of the working memory tasks correlated significantly with age. It is plausible that this result reflected selective attrition (i.e. older and less physically and mentally able participants are more likely to drop out), a possibility that will be discussed in more detail in Chapter 6[27]. Independent of significance, all correlations

[27] Similarly, it is possible that practice effects could have masked age effects on Reading Span; see Table 44, p.167 for the means and SDs for Reading Span across all three measurement occasions. Digit Ordering and Dot Matrix tasks were not included in subsequent analyses (see pp.131-132 for this rationale).

Table 23
Working memory task performance and its relationship with age on the initial assessment occasion (Pearson's r)

Working memory task	N	Age	N	Partial Ed	N	Partial Sex	N	Partial Ed & Sex
Reading Span	149	-.28**	136	-.26**	146	-.28**	135	-.26**
Dot Matrix	63	-.23	57	-.21	60	-.23	56	-.20
Digit Ordering	150	-.18*	137	-.18*	147	-.17*	136	-.18*

Ed = Education (years)
** p < .01; * p < .05

Table 24
Working memory task performance and its relationship with age on the intermediate assessment occasion (Pearson's r)

Working memory task	N	Age	N	Partial Ed	N	Partial Sex	N	Partial Ed & Sex
Reading Span	136	-.24**	124	-.22*	133	-.24**	123	-.22*
Dot Matrix	128	-.01	115	.10	125	-.01	114	.09
Digit Ordering	133	-.20*	121	-.19*	130	-.20*	120	-.19*

Ed = Education (years)
** p < .01; * p < .05

Table 25
Working memory task performance and its relationship with age on the final assessment occasion (Pearson's r)

Working memory task	N	Age	N	Partial Ed	N	Partial Sex	N	Partial Ed & Sex
Reading Span	127	-.11	117	-.09	124	-.11	116	-.09
Dot Matrix	120	-.14	110	-.07	117	-.13	109	-.07
Digit Ordering	127	-.13	117	-.12	124	-.12	116	-.11

Ed = Education (years)

were small in magnitude ($r < -.30$).

Low magnitude and non-significant correlations are likely to be the result of having a sample with a restricted range. The current sample was restricted in several ways, not only in task performance distributions but also in terms of the age range sampled. Many studies that have investigated the effects of ageing have involved participants spanning a range of 40 or more years (e.g. 20 to 65 years and older) and have, for example, reported correlations of $r = -.47$ ($p < .01$) and $r = -.52$ ($p < .01$) between age and computation span and listening span, respectively (Salthouse & Babcock, 1991). The current study spanned approximately 20 years and therefore the magnitude of differences was not likely to be as large. Moreover, the current sample could be described as well-educated (mean years of education = 11.68, SD = 4.13) and cognitively intact (scores below 17 on the Alzheimer's Disease Assessment Scale – Cognitive). Such characteristics are likely to make the effects of age on performance minimal, compared to less cognitively intact elderly samples (Waters & Caplan, 2003). Waters and Caplan also suggested that an absence of age differences on working memory task performance may be the result of poor test reliability. It is less likely that this would have been the main cause in the current study because Reading Span achieved a moderate test-retest reliability of $r = .58$ over 18 months.

4.3.1 Storage or processing: Which element of working memory is more sensitive to the ageing process?

Reading span-type tasks have commonly been used as measures of overall working memory performance (e.g. Baddeley & Hitch, 2000; Fisk & Warr, 1996). However, these types of tasks are complex (Lustig et al., 2001; Meguro et al., 2000; McCabe & Hartman, 2003) because they draw simultaneously upon various components of the working memory system. Baddeley (1990b) acknowledged that

reading span-type tasks are likely to involve strategy selection (central executive), the phonological loop, knowledge of vocabulary and coordination of these aspects (see also Fisk & Warr; Gathercole & Baddeley, 1993, Just & Carpenter, 1992; Lehto, 1996). Therefore, in terms of the Baddeley and Hitch (1974) tripartite model, age-related deterioration on reading span performance, or other measures of overall working memory ability, could stem from limited capacity of the phonological loop (i.e. storage/structure), impaired functioning of the central executive (i.e. processing/function), or both. Indeed, previous studies have indicated that all three of these scenarios may be responsible for age-related deficits on working memory performance.

Considering studies that have implicated the phonological loop as fundamental to age-related deficits on working memory performance, it is possible that central executive functioning was also involved. Thus, Hester, Kinsella and Ong (2004) compared the effect of age on the Wechsler digit span forwards and backwards tests. Hester et al. found similar rates of age-related deficits on both digit spans forwards and backwards, concluding that decline in phonological loop capacity explains age-related deterioration on working memory performance. Similar results were found by Gregoire and Van der Linden (1997). However, although there is general agreement that digit span forwards measures the storage capacity of the phonological loop (e.g. Baddeley, 2000a; Baddeley 2002), it has also been recognised that digit span backwards confounds this and central executive functioning (e.g. Groeger, Field & Hammond, 1999; Lezak, 1995). Consistent with this, both Hester et al. and Gregoire and Van der Linden acknowledged the alternate conclusion from their results; that these could also implicate the central executive.

Evidence for reduced central executive functioning as responsible for age-

related deterioration on working memory performance comes from studies that have involved Wechsler digit span tasks and from studies that have compared normal ageing to disease processes in ageing. For example, a study by Babcock and Salthouse (1990) and the results of a meta-analysis by Verhaeghan, Marcoen and Goossens (1993) found that, in general, age-related differences on digit span backwards (interpreted as predominantly representing central executive functioning) were larger than differences for digit span forwards. Dysfunction of the central executive component has been proposed as responsible for the memory deficits of patients with dementia of the Alzheimer's type – a dementia the incidence of which increases with age (Baddeley, Logie, Bressi, Della Sala & Spinnler, 1986). Moreover, Belleville, Perez and Malenfant (1996) have suggested that diminished central executive functioning characterises normal ageing. Alternatively, however, the results from Meguro et al.'s (2000) study (which also employed Wechsler digit spans) suggested that both central executive functioning and phonological loop capacity can be the causes of age-related deterioration on working memory performance, with the component responsible being dependent on the age groups involved.

In an attempt to decouple the operation of two theoretically distinguishable constructs (executive processes and structural components) in the current study, tasks similar to the Wechsler digit span tasks were added to the final assessment occasion (see pp.124-125 for a task description). The rationale for this procedural modification derived from the study by Meguro et al. (2000), who sought to distinguish phonological loop capacity from central executive functioning as responsible for age-related deterioration on working memory performance. This separation was achieved statistically. The study by Meguro et al. will now be described in detail.

4.3.1.1 Meguro et al. (2000)

Meguro et al. (2000) found negative age effects on a Japanese version of a reading span task. They considered Baddeley and Hitch's (1974) tripartite model "helpful for analysing the age-related decline of the reading span task performance" (p.392). Meguro et al. postulated that if age-related deficits on the reading span task were solely due to decreased capacity of the phonological loop, then this age difference should disappear when the score of a test measuring phonological loop capacity was used as a covariate in analysis of covariance (ANCOVA). Conversely, if age-related deficits on the reading span task were due only to central executive dysfunction, then employing a measure of phonological loop capacity as a covariate in ANCOVA should have no effect on reading span age differences. Alternatively, if age-related deficits on the reading span task were due to both decreased phonological loop capacity and central executive dysfunction, then age differences on the reading span task should disappear when a measure of both of these components was employed as a covariate in ANCOVA. Their study employed digit span forwards as the measure of phonological loop capacity and digit span backwards as the measure of both phonological loop capacity and central executive functioning.

Meguro et al. (2000) compared reading span performance across three age groups, using one-way analysis of variance. There was a significant, negative effect of age on reading span performance, as measured by the traditional span score (see current Table 26, p.124). ANCOVA was then applied to investigate which specific working memory component explained age differences on the reading span task. For this statistical test, reading span performance was the dependent variable, age group was the independent variable, and digit span forwards or backwards was the covariate. With digit span forwards (measuring phonological loop capacity) as the covariate, Meguro et

al. found that there was no longer a significant age difference between the middle-aged and elderly group. They therefore concluded that this initial difference in reading span performance was mainly due to reduced capacity of the phonological loop from middle to old age. However, when employing digit span backwards (measuring central executive functioning) as the covariate, they found that there was no longer a significant age difference between the young and middle-age group; and they therefore concluded that this difference in reading span performance was mainly due to a difference in the capacity of the central executive. The age difference between the young and elderly group remained significant with the use of each covariate. Meguro et al. did not provide any conclusions about this matter but, taken together, these results could mean that decline in executive functioning commences prior to middle-aged, followed by subsequent deterioration of the phonological loop with no further appreciable decline in central executive capacity. Alternatively, an untested construct might explain these age differences.

Table 26
Performance on a reading span task by age group (data from Meguro et al., 2000)

Age group	N	Mean Age	SD	Mean Span	SD
Young	21	28.8	6.07	3.50	1.70
Middle-aged	20	49.4	5.68	2.53	1.50
Elderly	21	68.3	6.47	2.02	1.00

One-way ANOVA: $F(2, 59) = 23.66$, $p < .01$

4.3.1.2 Application of Meguro et al.'s (2000) study here

Given the context of the current study, it was appropriate to adopt Meguro et al.'s (2000) procedure. To achieve this, data for digit span forwards and backwards were collected on the final measurement occasion[28]. The digit span tasks are similar to Digit Ordering; participants were verbally presented with sequences of digits to re-order. The digit-sequences increased in length by one digit after two trials at the

[28]Other data used in this dissertation were co-collected with another PhD student. However, data for the digit span tasks were collected by the author alone, from only half the total sample.

same sequence length had been completed. In the case of digit span forwards, sequences needed to be repeated in the order delivered; for digit span backwards, digits were repeated in the reverse order to presentation (i.e. recall the last digit heard first and work 'backwards'). Digit span forwards ranged from two to nine digits to order and digit span backwards ranged from two to eight digits. Instructions and stimuli for these tasks closely resembled the Digit Span tasks from the Wechsler Adult Intelligence Scale (Wechsler, 1997). The scoring method employed was number correct, with one point allocated for every digit-sequence recalled entirely correctly. Higher scores therefore indicated better performance.

Before ANCOVAs were carried out involving the digit span tasks as covariates, it was necessary to establish the presence of age differences on the working memory tasks (i.e. determine a dependent variable). For the sample who completed the digit span tasks (demographic details to be provided in Table 27, p.126), Digit Ordering was the only task to correlate significantly with age ($r = -.33$, $p < .05$; Reading Span, $r = -.10$, $p > .05$; Dot Matrix, $r = -.21$, $p > .05$). Participants were then divided into two age groups: those younger than 80 years of age ('young-old'); and those 80 years of age and older ('old-old'). An independent samples t-test was carried out to test the significance of the mean difference in Digit Ordering performance between these groups. The results indicated that, on average, the young-old group performed significantly better on the Digit Ordering task than the old-old group (also see Table 27). Consequently, Digit Ordering was used as the dependent variable in the ANCOVAs (rather than Reading Span, as used by Meguro et al., 2000).

The first ANCOVA employed digit span forwards (a measure of phonological loop capacity) as the covariate, age group as the independent variable and Digit Ordering as the dependent variable. Statistically controlling for performance on the digit

span forwards task did not attenuate age differences between the young-old and old-old groups on Digit Ordering ($F = 5.82$, $p < .05$). Therefore, in keeping with the reasoning of Meguro et al. (2000), age differences on a general measure of working memory were not explained by a difference in phonological loop capacity. The second ANCOVA only varied by employing digit span backwards (a measure of phonological loop capacity *and* central executive functioning) as the covariate. Age group still remained a significant influence ($F = 5.41$, $p < .05$) for Digit Ordering performance, despite controlling for performance on digit span backwards. Therefore, the difference between young-old and old-old performance on the Digit Ordering task was not explained by a decrease in phonological loop capacity, together with decline in central executive functioning.

Table 27
Demographic information and performance on Digit Ordering by age group (data from the final measurement occasion)

Age group	N^	Mean Age	SD	Mean Years Education	SD	Mean Number Correct#	SD
Young-old (< 80 years)	35	76.53	2.13	12.53	5.88	16.80	3.15
Old-old (80+ years)	21	83.45	1.85	11.21	2.72	14.90	2.74

Independent t-test for Age: $t = 12.33$, $df = 54$, $p < .01$
Independent t-test for Years of education: $t = 1.12$, $df = 51$, $p > .05$
Independent t-test for Digit Ordering: $t = 2.29$, $df = 54$, $p < .05$
^N = 56 due to the removal of outliers; this total includes 31 females and 25 males.
#Mean Number Correct is equivalent to 6.5 (traditional) span for young-old and 5.5 span for old-old

This inconclusive result prompted a further ANCOVA to be conducted. The covariate on this occasion was a measure of speed of processing[29]. Visual Matching was selected as an appropriate speed measure because, not only does it strongly reflect the speed of processing factor of intelligence (McGrew, 1997), but it also involves numeric stimuli (i.e. digits). Maintaining focus on numeric tasks minimised confounding from stimulus type. Visual Matching is a timed pen and paper task that consists of two

[29] Co-collected data included four measures of speed of processing. Consequently, the author had access to speed of processing data, although these were only ancillary in the current thesis.

columns, each containing 30 rows of digits. Each row contains six digits. Two out of the six digits are the same and the participant must find and circle the two digits that are the same as quickly as possible. Participants are allowed two minutes to complete as many lines as possible. Instructions to this effect were given to participants, in accordance with the manual for the Woodcock-Johnson Psycho-Educational Battery – Revised (Woodcock & Johnson, 1989). Practice was provided. Isolating the matching digits becomes more difficult as the task progresses. That is, the rows of digits start with single digits (e.g. 8, **9**, 5, 2, **9**, 7), then double-digits are introduced (e.g. 85, 32, 74, 90, **61**, **61**), finally progressing to triple-digits (e.g. **968**, 689, 869, **968**, 986, 896). There are 20 lines of each difficulty level before the digit-type changes. One point is allocated for every correct matching-pair identified (maximum score is 60), with higher scores indicating faster speed of processing.

Statistically controlling for performance on Visual Matching (i.e. as a covariate in ANCOVA), resulted in a nonsignificant result for age group on Digit Ordering ($F = 2.51$, $p > .05$). In other words, differences in speed of processing explained age-related deficits on a test of overall working memory ability in the current study. The young-old group could process information significantly faster (mean = 34.71 rows correct, SD = 5.37) than the old-old age group (mean = 31.00 rows correct, SD = 6.21; $t = 2.36$, $df = 54$, $p < .05$). This result was not a product of high inter-correlations among covariates (see Table 28, p.128). Similar outcomes have been found in many other studies. For example, Salthouse and Babcock (1991) found that processing efficiency was the most important determinant of age-related differences in working memory. In turn, they found that age-related declines in processing efficiency were mediated by age-related declines in simple measures of processing speed (age range of their sample = 18 - 87 years). Moreover, Fisk and Warr (1996) found, by way of regression analyses, that

controlling for processing speed attenuated age effects on a reading span test, whereas controlling for central executive functioning did not. The age range included in their studies was 20 to 80 years.

In summary, for the current sample, neither phonological loop capacity nor central executive functioning could account for age differences on a task of general working memory ability (Digit Ordering). In contrast, speed of processing was sensitive to subtle changes in Digit Ordering performance across these elderly age groups. This is not to say that the capacity of the phonological loop and/or functioning of the central executive cannot explain age differences on a task of working memory across different age groups (as shown by the results of Meguro et al., 2000). Rather, these individual working memory components may only differentiate performance up until a certain age, beyond which speed of processing then dominates age changes across working memory and other cognitive tasks. Moreover, this result does not detract from the importance of these working memory components but, instead, highlights the integral role processing speed plays in processes intrinsic to these components.

Table 28
Inter-correlations of tasks used in ANCOVAs to investigate which element of working memory is more sensitive to the ageing process (final test session data)

	1.	2.	3.	4.
1. Digit Span Forwards	-	.49**	-.04	.43**
2. Digit Span Backwards		-	.04	.51**
3. Visual Matching			-	.35**
4. Digit Ordering				-

N = 56 for all tasks; **p < .01

4.3.1.3 Discussion of results about age differences in Digit Ordering

There are two main points to discuss regarding the above results before providing a summary of this Chapter. Firstly, the results of Meguro et al.'s (2000) study, and consequently those of the current study, are dependent on knowing what the tasks employed actually measure. Miyake et al. (2001) considered both digit span tasks to assess phonological loop capacity only. Alternatively, if digit span backwards

incorporates phonological loop capacity *and* central executive functioning (as suggested by Gregoire & Van der Linden, 1997 and Hester et al., 2004) then what differentiates it from other tasks of overall working memory, such as Digit Ordering? In the same way, there is no widely held consensus about what the Digit Ordering task measures, especially in well-educated, cognitively intact elderly populations. In the current study, Digit Ordering was shown to load with other, more commonly used tasks of working memory (see p.117), but it also tended to show weaker correlations with these more complex working memory tasks (e.g. Reading Span; see p.114). Moreover, of all the working memory tasks, Digit Ordering showed the strongest correlations with the digit span tasks (see Table 29). In some circumstances, the correlations between Digit Ordering and the digit span tasks were stronger than the correlation between each of the digit span tasks. In part, the strength of these correlations is likely to reflect similar task procedures and the involvement of the same type of stimuli. These correlations also suggest that Digit Ordering, digit span forwards and digit span backwards draw upon working memory components (phonological loop and central executive) to a similar extent, at least relative to the other tasks of working memory.

Table 29
Inter-correlations of the working memory tasks used in the final test session data

	N	1.	2.	3.	4.	5.
1. Digit Span Forwards	60	-	.52**	.36**	.26*	.55**
2. Digit Span Backwards	60	.49**	-	.36**	.30*	.55**
3. Reading Span	127	.31*	.31*	-	.51**	.31**
4. Dot Matrix	120	.19	.26	.55**	-	.50**
5. Digit Ordering	127	.43**	.51**	.29*	.50**	-

**p < .01; *p < .05
Values below the diagonal are correlations when sample is listwise for Digit Span tasks (N = 56)

The second and final discussion point refers to the presence and absence of significant correlations between age and working memory tasks. That is, when the total sample for the final measurement occasion was analysed, none of the working memory tasks demonstrated significant relationships with age (see Table 25, p.119). In contrast,

when the sub-sample who completed the digit span tasks was investigated, there was a significant, negative relationship between age and Digit Ordering. This difference prompted exploration of what characteristics may have differed between people who completed the digit span tasks and those who did not. Independent samples t-tests were conducted and revealed that participants who completed digit span forwards and digit span backwards performed significantly better on some working memory tasks; in particular, Digit Ordering and Dot Matrix (see Table 30).

Table 30
Various performance comparisons between participants who did and did not complete the digit span tasks on the final measurement occasion

	Digit Span Tasks			
	Completed (N = 60)		Not Completed (N = 67)	
Comparison measures	Mean	SD	Mean	SD
Age	79.12	4.03	78.54	4.25
Years of education	11.89	4.92	11.74	3.62
Digit Ordering**	15.72	3.38	13.37	3.47
Reading Span	14.55	9.76	15.24	7.85
Dot Matrix**	10.58	4.97	8.00	3.79
Raven's	14.41	4.72	13.67	4.07
Basic ADL	6.73	.71	6.71	.76
Instrumental ADL	25.90	1.62	25.53	2.57
Total symptoms	71.60	11.12	72.22	14.45
Total diseases	2.92	1.74	2.74	1.85

ADL = Activities of daily living; Raven's = Raven's Standard Progressive Matrices
Digit Ordering, $t = 3.85$, $df = 125$, $^{**}p < .01$; Dot Matrix, $t = 3.22$, $df = 118$, $^{**}p < .01$

Despite the range of scores on the Digit Ordering task being similar for each group (completed the digit span tasks = 7 - 21; did not complete the digit span tasks = 3 - 20), their performance distributions were different. The performance distribution of those who did not complete digit span forwards and backwards demonstrated more positive kurtosis, illustrated by a median score of 13.00, compared to a median of 16.00 for those who completed the digit span tasks. Chi-square analysis showed that membership to the young-old and old-old group was evenly represented across the samples who did and did not complete the digit span tasks (Yates' Continuity Correction = .02, $df = 1$, $p > .05$). In fact, for both of these samples, N = 23 membership

to the old-old group. Therefore, the presence of a significant and negative correlation between age and Digit Ordering performance for the analyses involving the digit span tasks sub-sample was due only to better Digit Ordering performance in this sample. This is further evidenced by the lack of a significant mean difference in Digit Ordering performance across these age groups for the sample who did *not* complete the digit span tasks (young-old: mean = 13.10, SD = 2.99; old-old: mean = 13.74, SD = 4.07; $t = .73$, $df = 63$, $p > .05$).

4.4 Summary

The construct validity of the three tasks employed to measure working memory was established by factor analysis. The strong loadings of Reading Span, Dot Matrix and Digit Ordering onto one factor, and with a separate factor representing short-term memory, permitted the classification of these tasks as general working memory capacity tasks. Significant and negative correlations between age and these tasks were few and small; the restricted range on some variables in this sample was considered the primary reason behind this. Statistical differentiation of the storage (phonological loop capacity) and processing (central executive functioning) functions of working memory tasks to explain age differences was unsuccessful. Instead, analyses of covariance showed that speed of processing is an important explanatory variable when assessing age differences in working memory performance in a well-educated, cognitively intact, elderly population.

Initially, the utility of each working memory task as a biomarker of independent functioning in an elderly population was to be assessed, with the aim of identifying outcome variables that were strongly related to each of the theoretically proposed content domains of the Baddeley and Hitch (1974) model. However, limited Dot Matrix data on the initial measurement occasion prevented longitudinal analyses

involving this measure (Appendix K cites sample characteristics, the environment of the test administration and the nature of the test battery as reasons for the limited sample size). Moreover, the inadvertent use of nameable visuo-spatial stimuli for Dot Matrix (i.e. the location of a dot in the grid could be remembered by the verbal strategy of 'row 1, column 5'), would have confounded the content domain of this task. In addition to this, results from principal components analysis, inspection of correlation matrices and comparisons of performance on all three working memory tasks, all indicated that Digit Ordering was somewhat different from the other two working memory tasks. Furthermore, feedback from participants about this task highlighted the frequent use of strategy to assist performance[30]. Greater mean and median performance for the overall sample on this task, compared to the other working memory tasks, suggested enhanced performance.

For these reasons, Dot Matrix and Digit Ordering were not included in subsequent analyses to assess the predictive utility of working memory capacity in independent functioning outcomes for an elderly population (and therefore psychometric properties for these tasks were not included in Chapter 3). Instead, Reading Span, representing the construct of general working memory capacity, featured as the sole cognitive biomarker candidate in both cross-sectional (Chapter 5) and longitudinal (Chapter 6) results.

[30] Strategy use is associated with some roles of the central executive component of Baddeley and Hitch's (1974) tripartite model (see Chapter 2, p.32 for more details) and thereby strategy can be considered important to the construct of working memory. However, the 'strategy' reported in the current study appeared less elaborate and more observational in nature and therefore was deemed to confound, rather than reflect, working memory ability.

CHAPTER 5: CROSS-SECTIONAL RESULTS (INITIAL ASSESSMENT)

5.1 Overview

As described in Chapter 1, the primary purpose of this thesis was to evaluate the utility of working memory for predicting independent functioning in an elderly population. Evaluation required comparison with other more commonly used biomarkers (e.g. physiological measures) as well as comparison against chronological age, a measure that often inaccurately reflects functional ability (see Chapter 1, p.9 for an explanation of this). Chapters 2 and 4 have provided an analysis of working memory theory and operationalisation. Chapter 4 concluded that, in the current sample, Reading Span was the most suitable task to use in subsequent analyses because it reflects general working memory capacity, and use of strategy to complete this task appeared to be minimal. Chapter 3 detailed all the measures assessed in this project, including various outcome measures for independent functioning.

Using different types of regression analyses, Chapter 5 presents an initial examination of the predictive utility of Reading Span (i.e. cross-sectional results). As described in Chapter 1, true evaluation of whether working memory capacity can be considered a cognitive *biomarker* can only be made by way of longitudinal results (presented in Chapter 6). Chapter 5 is divided into three main sections; first, descriptive statistics are provided for measures in the test battery; second, the influence of covariates on outcome variables and biomarkers is assessed; and third, regression results are presented for three outcome variables. The Chapter concludes with a discussion of the results and the limitations of method.

5.2 Descriptive statistics

This section provides a description of participant performance on the dementia

screening test (one of the participant selection criteria), followed by outcome variables, then covariates, concluding with biomarkers.

5.2.1 Selection criteria – Dementia Screening

The Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog) was employed to screen participants for dementia. Participants made an average of 4.95 errors (SD = 2.95; median = 4.35) out of 70 items on this test. According to the performance classifications of the ADAS-Cog, 149 of the 150 participants were considered to have normal cognitive functioning (i.e. 0 - 13 errors). One participant had a score of 17.60 errors, meeting the threshold indicating definite impairment (i.e. > 17 errors). However, being so close to the threshold for borderline impairment (i.e. 14 - 17 errors), this participant was not excluded from the study.

5.2.2 Outcome variables

Three outcome variables considered to reflect independent functioning in elderly persons were employed in the initial test session. These included activities of daily living, reasoning ability and life satisfaction. These measures will now be discussed.

5.2.2.1 Activities of daily living

Overall, the elderly participants in this sample tended to be very independent in all types of activities of daily living (ADL). However, although basic ADL were likely to be completed independently (see Table 31, p.135), more dependence was shown for instrumental ADL. This result is consistent with the hierarchical structure proposed by Spector et al. (1987; refer to Chapter 3, p.68 for a description) and with the nature of the sample (i.e. community-dwelling). Table 31 also shows that half of the participants scored the highest value of independence for both types of ADL; and Figure 13 (p.135) clearly shows the marked ceiling effect for this

outcome measure[31].

Table 31
Descriptive statistics for both types of activities of daily living at initial assessment

	N	Median	Mean	SD	Min.	Max.
ADL (total score)	150	34.00	32.62	2.20	23.00	34.00
Basic ADL	150	7.00	6.82	.61	4.00	7.00
Instrumental ADL	150	27.00	25.80	1.87	18.00	27.00

ADL = Activities of daily living

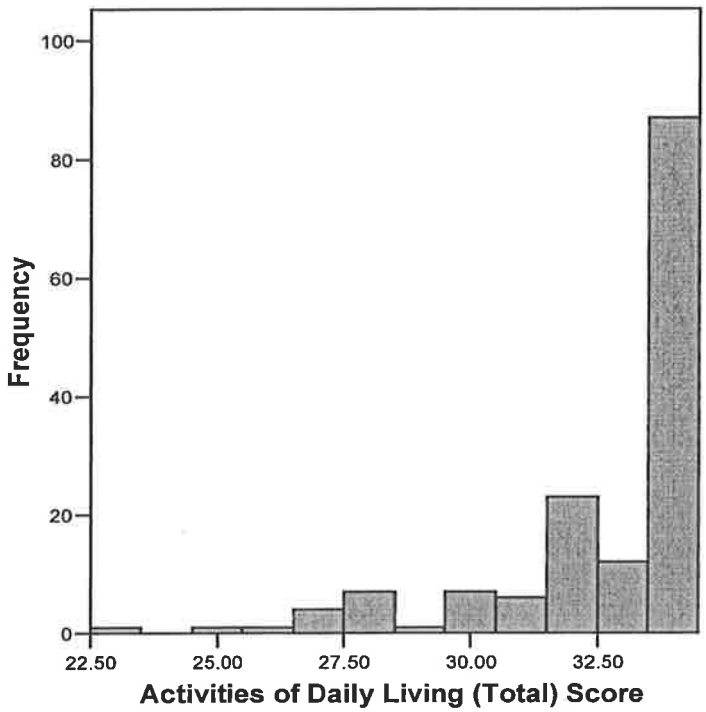


Figure 13 Distribution of activities of daily living scores (total) at the initial test session

This negatively skewed distribution indicates the current elderly sample was independent and high functioning.

5.2.2.2 Reasoning ability

Raven's Standard Progressive Matrices (Raven's) was employed as an indicator of reasoning ability. On average, participants answered 16.61 puzzles correctly (SD = 4.73; median = 17.00), just over half of the total number. Scores ranged from 7 to 26 with no outliers. These scores are higher than those reported by Salthouse (1993). The elderly participants from his sample (60 - 79 years) demonstrated a median score of approximately 10.50 and scores ranged from 3 to 18 (note that the administration of

[31] This is one of several variables that displayed an extremely skewed distribution. Attempts were made to normalise these truncated distributions, but to no avail.

Raven's was different in the current study compared to Salthouse's study). The current sample can be described as cognitively intact. to no avail

5.2.2.3 Life satisfaction

Self-reported life satisfaction was normally distributed among participants. Out of a maximum score of 200, indicating participants were "very satisfied" across all eight facets of life (described in Chapter 3, p.76), mean life satisfaction was rated at 146.62 (SD = 14.77; median = 147.00; range = 101 - 186). The median value here is greater than age norms reported in the Life Satisfaction Scale manual (Salamon & Conte, 1984; 50% quantile = 135.00 for adults aged between 71 and 75 years. Level of satisfaction dropped with each subsequent five-year increment). Consistent with this, Salamon and Conte stated that reported levels of life satisfaction vary according to the characteristics of the sample assessed. For example, individuals with clinical depression, health problems or poor financial status will typically indicate lower life satisfaction across the eight sub-scales. It is clear, therefore, that participants in the current study showed a high level of life satisfaction, indicating lives that are largely absent from the aforementioned negative events.

5.2.3 Covariates

Descriptive statistics for demographic covariates (i.e. age, sex and years of education) were provided in Chapter 3 (see pp.62-63). Other covariates included in the initial test session were depression, health and lifestyle. Descriptive statistics for these measures are provided below.

5.2.3.1 Depression

The Geriatric Depression Scale (GDS) was used to assess depressive symptoms in the current sample. On average, participants indicated they experienced 5 depressive symptoms out of a possible 30 (SD = 4.98; median = 4.00; range = 0 - 25).

This figure is comparable with the mean reported by Anstey and Smith (1999) for a sample of 180 community-dwelling women aged between 60 and 90 years (mean = 6.04, SD = 4.92).

According to the performance classifications of the GDS, 137 of the 150 participants in the current study were in the normal range for non-depressed elderly persons (i.e. 0 - 10 symptoms). Eight participants met the threshold for mild depression (i.e. 11 - 20 symptoms) and five participants met the threshold for moderate to major depression (i.e. 21 - 30 symptoms). In summary, the general absence of depression was reflected by a median score of 4 and a positively skewed distribution (see Figure 14).

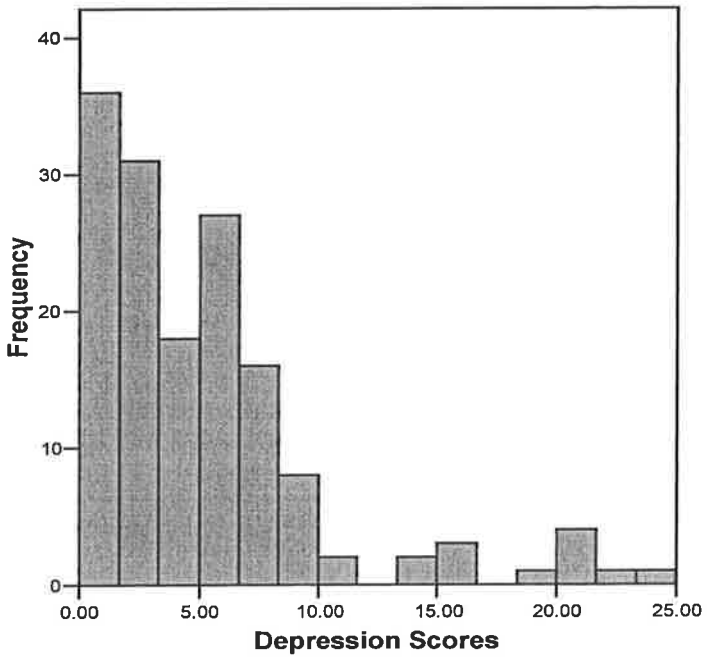


Figure 14 Distribution of depression scores at the initial assessment occasion
The distribution shows marked floor effects with the majority of participants reporting 10 or fewer depressive symptoms. This indicates the participants typically represented a normal, non-depressed elderly sample.

5.2.3.2 Health

Severity and number of chronic diseases experienced comprised the health covariate (see Chapter 3, p.83). Mean symptom severity was 69.53 (SD = 12.36); if all 51 symptoms were experienced severely the maximum score would be 204.

Therefore, participants reported that few, if any, symptoms were suffered severely (illustrated in Figure 15). A range from zero to six chronic diseases was reported by participants (mean = 2.55, SD = 1.58). Of the 17 chronic diseases listed, half the participants reported experiencing 2 diseases only (arthritis and high blood pressure were the most common).

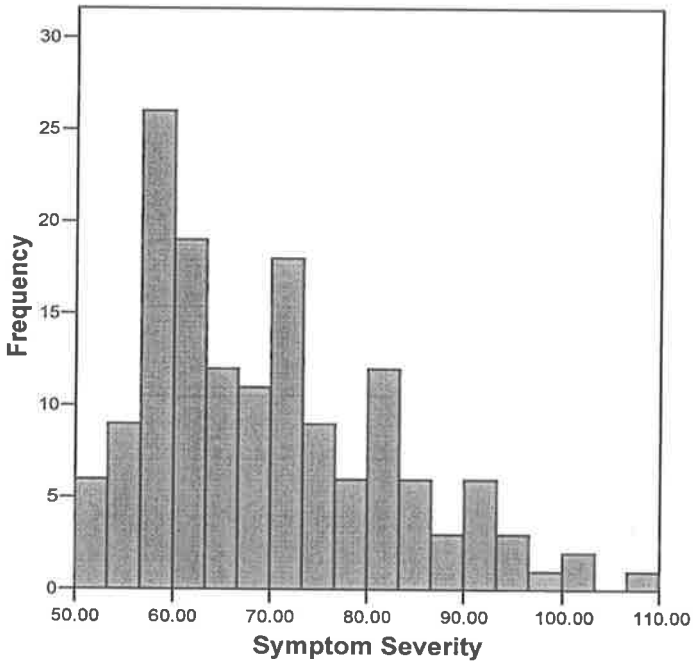


Figure 15 Distribution of symptom severity scores at the initial assessment occasion

The distribution is positively skewed, as suggested by a median score of 67. This reflects the good health of the sample, or at minimum, the asymptomatic stages of disease.

5.2.3.3 Lifestyle

The lifestyle covariate comprised cigarette smoking, exercise, alcohol consumption and diet variables. In regard to smoking status, 64 participants had never smoked, 79 participants were former smokers, and 7 participants currently smoked. Participants exercised an average of four times a week (SD = 2.92; median = 4.5; range = 0 - 14) and each session lasted for an average of 42.78 minutes (SD = 44.93; median = 30.00; range = 0 - 240). These measures were used to produce an 'overall time spent exercising per week' variable; participants exercised for an average of 201.93 minutes

per week (SD = 235.57; median = 157.50; range = 0 - 1680). The most common form of exercise was walking. As described in Chapter 3 (see footnote, p.89), the current sample did not include any heavy drinkers (i.e. individuals consuming > 25 standard alcoholic drinks per week); 63 participants were classified as abstainers (i.e. zero drinks per week); 63 participants were considered light drinkers (i.e. .5 to 7 drinks per week) and 22 participants met the threshold for moderate drinkers (i.e. 8 to 25 drinks per week). Two participants did not describe their alcohol consumption.

In summary, the participants were conscientious about their health, and lifestyle choices were generally consistent with the following widespread health guidelines; don't smoke cigarettes (fewer than 5% currently smoked); exercise for 30 minutes at least three or four times per week (half the participants achieved slightly more than this recommendation); and don't consume more than two (if female) or three (if male) standard drinks of alcohol per day (no participants were classified as heavy drinkers).

The last lifestyle variable assessed was dietary intake. Chapter 3 provided evidence for the influence of certain B-vitamins and anti-oxidants on ageing and cognition. Table 32 (p.140) shows the number of participants who fell below and above the recommended daily intake (RDI) for these nutrients (see Appendix L for the RDI cut-off values). In general, the RDI cut-off values are sex specific and the 'below RDI' category indicates deficient nutrient intake (except for Vitamin E where this indicates sufficient intake). The classification of RDI groups was based on total nutrient intake (i.e. included nutrients obtained from supplements). As can be seen, the majority of participants reported more than sufficient intake of the nutrients assessed.

In addition to the influence of RDI groups, sub-clinical nutrient deficiencies may also influence cognitive performance. Therefore, Table 33 (p.140) provides the descriptive statistics for daily, overall self-reported nutrient intake for the B-Vitamins

and anti-oxidants. Of particular note are the differences between the mean and median levels of nutrient intake. For all nutrients assessed, half of the participants received substantially less nutrients than is reflected by the mean (i.e. intake distributions were markedly positively skewed). The skewed distributions were due to outliers, typically those individuals who took nutrient supplements (N = 69). These individuals were excluded from subsequent analyses (see p.145 and Appendix O).

Table 32
Initial test session: Frequency of participants (N = 135*) in each category of nutrient intake relative to the daily Australian recommended dietary intake (RDI)

	Below RDI	Within RDI	Above RDI
B-Vitamins			
Thiamin	11	N/A	124
Riboflavin	11	N/A	124
Niacin	0	N/A	135
B-6	12	23	100
Folate#	21	N/A	114
B-12#	27	N/A	108
Antix-oxidants			
Vitamin C	11	N/A	124
Vitamin E	79	N/A	56
Vitamin A#^	N/A	133	2

*Not all participants provided sufficient detail about their food consumption to calculate nutrient levels
#measured in micrograms (μg); all others are measured in milligrams (mg)
N/A = Not applicable
^There is only a maximum suggested intake for Vitamin A (i.e. no minimum suggested intake). Therefore, 'above RDI' in this case indicates a negative situation.

Table 33
Initial test session: Descriptive statistics of daily, overall nutrient intake of B-Vitamins and anti-oxidants in the current sample (N=135*)

	Median	Mean	SD	Min.	Max.
B-Vitamins					
Thiamin	1.35	19.59	53.26	.27	228.13
Riboflavin	1.84	21.23	63.36	.41	511.71
Niacin	32.49	41.10	39.20	16.06	317.55
B-6	1.64	22.00	58.98	.49	322.38
Folate#	351.37	490.94	572.98	49.15	3920.28
B-12#	3.41	28.46	108.64	.51	1135.43
Antix-oxidants					
Vitamin C	137.18	381.18	920.13	12.88	7736.45
Vitamin E	6.89	45.56	124.53	1.35	763.00
Vitamin A#	886.29	1361.51	1623.02	159.89	11371.97

*Not all participants provided sufficient detail about their food consumption to calculate nutrient levels
#measured in micrograms (μg); all others are measured in milligrams (mg)

5.2.4 Biomarkers

In the current project, four commonly used biomarkers were included (blood pressure, grip strength, body mass index, visual acuity) in addition to working memory capacity. Descriptive statistics for these variables are provided below.

5.2.4.1 Blood pressure

Systolic and diastolic blood pressures were measured; the mean of three consecutive readings was 147.98 mmHg (SD = 22.70; median = 146.00) and 79.23 mmHg (SD = 11.53; median = 79.00), respectively. Table 34 demonstrates that this blood pressure average can be classified as normal to mild hypertension.

Table 34
Frequency of participants in each blood pressure (BP) category on the initial assessment occasion (N=137*)

BP category	Range^#	Frequency
Optimal	<120/ 80	13
Normal	121-130/ 81-85	18
Normal systolic	131-140/ 86-90	19
Mild hypertension	141-160/ 91-100	50
Moderate hypertension	161-180/ 101-110	25
Severe hypertension	≥ 181/ ≥ 111	12

*Lower sample size is due to the apparatus (i.e. cuff too small and/or too painful when inflated; also see p.161)
^Range = systolic BP/diastolic BP, mmHg
#The range for each category, and subsequent category membership, was based on information in the Omron manual (Omron, 2001).

5.2.4.2 Grip strength

Physiologically, males tend to be stronger than females. Consequently, grip strength is presented for males and females separately. The average grip strength from three trials, using the strongest hand, was used as the variable here. The current data, represented in Figure 16 (p.142), supported the expected physiological difference.

5.2.4.3 Body mass index

Figure 17 (p.143) demonstrates that most participants (73 of the 150) fell within the 'normal' body mass index (BMI) range (i.e. 22 - 26 kg/m²) for elderly persons. The sole outlier was an individual with a 'very underweight' BMI (i.e. < 18

kg/m²). The BMI ranges reported are applicable to older Australians, that is, individuals 74 years of age and older (Victorian Government, 2003).

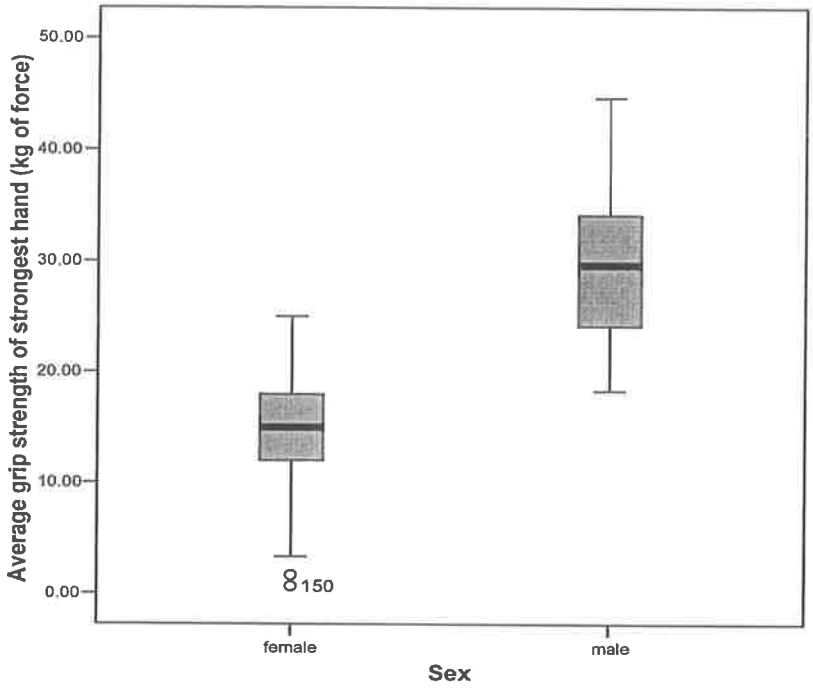


Figure 16 A comparison by sex of average grip strength (strongest hand) on the initial test session (N = 150)

It can be seen that male participants produced almost twice the grip strength (mean = 29.60, SD = 6.47, kg of force) of females (mean = 14.72, SD = 5.11).

5.2.4.4 Visual Acuity

Corrected visual acuity scores were transformed so that higher scores reflected better visual acuity (see Chapter 3, p.103 for a description of and rationale for this process). From 150 participants, 87 (58%) achieved a score of 95, the highest score possible, and this reflects normal distance vision. The lowest score in the current sample was 76 (N = 1), indicating that the smallest size letter read correctly by the participant at a distance of 4 metres, could typically be read at a distance of 24 metres by a person with normal distance vision. The lowest score possible was 40, although no one scored at this level. Therefore, on the whole, participants had very good visual acuity (i.e. the distribution was negatively skewed).

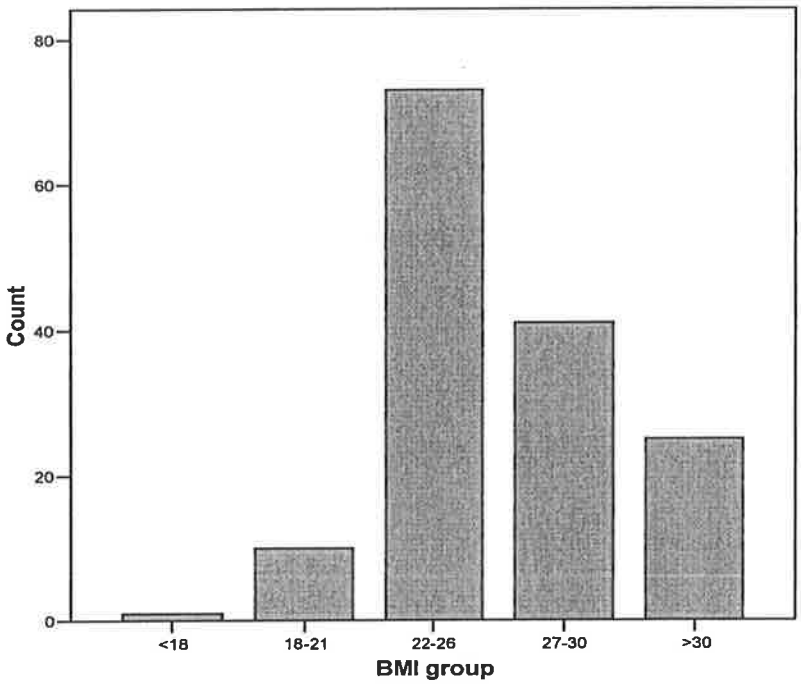


Figure 17 Participant distribution across body mass index categories (initial test session data)

Excluding those in the normal body mass index (BMI) range (i.e. 22 - 26 kg/m²), there was a tendency for participants to be overweight (i.e. BMI = 27 – 30 kg/m²; N = 41) or obese (i.e. BMI = > 30 kg/m²; N = 25).

5.2.4.5 Reading Span

Participants recalled a mean of 13.91 words correctly, out of a possible 30 (SD = 8.81; median = 13.00; range = 0 - 30). This number correct score equates to a traditional span score of 2, meaning the participant can consistently and correctly process and recall the final word from two sentences at a time. This level of working memory capacity is similar to that reported in other studies involving elderly participants (see Chapter 2, pp.47-49 for more details).

5.3 Relationships among variables

Relationships between dependent (i.e. outcome variables) and independent variables (biomarkers) may be explained, at least in part, by associations with other variables (covariates). The associations among these variables were therefore examined, to determine which covariates must be included in subsequent regression analyses. The nature of these associations is presented below.

Of the different types of covariates, demographic variables exerted the most influence. For example, reliable sex differences were found on reasoning ability, grip strength and Reading Span. In all cases, males performed at higher levels than females. The effect sizes for reasoning ability and Reading Span were small, whereas, as expected, it was large for grip strength (see Tables 35 & 36 for eta² and results of independent samples t-tests). Education was also significantly and positively correlated with reasoning ability and Reading Span (see Tables 37 & 38, p.145). Age is another commonly used demographic covariate but in the current study this variable was employed as a biomarker against which to compare the predictive utility of other biomarkers.

Table 35
Performance on outcome variables by sex on the initial assessment occasion

	Sex	
Outcome variables	Female (N=99)	Male (N = 51)
Activities of daily living		
Basic#	6.76	6.94
Instrumental	25.75	25.90
Reasoning ability**	15.55	18.69
Life Satisfaction	146.96	145.96

#This difference is significant (df = 148, t = 2.03, p < .05, eta² = .03). However, only 13 people achieved less than the maximum score and 11 of these participants were female. Thus, excluding these outliers, all males and females achieved the highest score possible (7.00).
Instrumental ADL: df = 148, t = .48; **Raven's: df = 148, t = 4.05, p < .01, eta² = .10;
Life satisfaction: df = 148, t = .42.

Table 36
Performance on biomarkers by sex on the initial assessment occasion

		Sex	
Biomarkers		Female (N = 99)	Male (N=51)
Age	(years)	77.71	77.35
Systolic BP	(mmHg)	149.13	146.03
Diastolic BP	(mmHg)	79.16	79.35
Grip Strength	(kg of force)**	15.13	29.60
BMI	(kg/m²)	26.84	25.74
Visual Acuity	(at 4 metres)	93.32	93.78
Reading Span	(no. correct)*	12.81	16.04

BP = blood pressure; BMI = body mass index
Age: df = 148, t = .51; Systolic BP: df = 135, t = .77; Diastolic BP: df = 135, t = .09;
**Grip strength: df = 78, t = 14.15, p < .01, eta² = .3; BMI: df = 148, t = 1.82;
Visual Acuity: df = 148, t = .96; *Reading Span: df = 147, t = 2.02, p < .05, eta² = .03.

There were no reliable relationships between biomarkers and depression. The

only reliable relationship between depression and outcome variables was for Life Satisfaction ($r = -.45$, $p < .01$, $N = 134$), an expected outcome. Thus, within the scoring range considered to reflect normal non-depressed elderly (i.e. excluding outliers for depression), more depressive symptoms were associated with less life satisfaction. Other outcome measures initially showed significant relationships with depression but these disappeared once outliers were excluded (see Appendix M for details).

Table 37
Correlations between education and outcome variables on the initial assessment occasion

Biomarkers	N	Pearson's r
Activities of daily living		
Basic	136	.02
Instrumental	136	.09
Reasoning ability	136	.28**
Life Satisfaction	135	.11

**p < .01

Table 38
Correlations between education and biomarkers on the initial assessment occasion

Biomarkers	N	Pearson's r
Age	136	-.07
Systolic BP	123	.08
Diastolic BP	123	.05
Grip Strength	136	-.02
BMI	136	-.12
Visual Acuity	136	.06
Reading Span	135	.19*

BP = Blood pressure; BMI = Body mass index
*p < .05

On the whole, health and lifestyle covariates did not have significant relationships with either biomarker or outcome variables. Of the significant associations, these tended to disappear once sex and outliers were taken into consideration (see Appendices N & O for details); the exception to this was life satisfaction. As might be expected, life satisfaction showed significant and negative associations with symptom severity ($r = -.39$, $p < .01$, $N = 134$) and total number of diseases ($r = -.27$, $p < .01$, $N = 134$). Based on the limited associations between health and other salient variables, health was not further assessed by different body systems or

by symptom severity threshold (see Chapter 3, pp.83-84 and Appendix G for the rationale of this analysis).

In summary, sex and education were covaried in all regression analyses. Depression, symptom severity and total number of diseases were *not* covaried in regression analyses involving life satisfaction because none of these covariates was significantly associated with biomarker variables. The key purpose of the current project was to evaluate the predictive utility of selected biomarkers, not that of covariates. The limited influence of covariates is likely to be due to the particular characteristics of this sample; participants functioned independently and were healthy, well-educated and cognitively intact.

5.4 Regression analyses: Predicting independent functioning outcome variables

The initial assessment occasion included three outcome variables: activities of daily living (consisting of basic and instrumental activities), reasoning ability and life satisfaction. Regression analyses were employed to identify which biomarkers (independent variables) could predict each of these outcome measures (dependent variables), after controlling for sex and education. Hierarchical linear regression (using the enter method)[32] was employed for reasoning ability and life satisfaction whereas logistic regression was used for activities of daily living[33]. The results of the regression analyses are presented below. To assist in their interpretation, bivariate correlations between dependent and independent variables are provided in Appendix P.

5.4.1 Activities of daily living

Logistic regression involves the prediction of group membership, meaning the

[32] In general, covariates (i.e. sex and education) were entered simultaneously on the first step; all biomarkers were entered simultaneously on the second step.

[33] An assumption of linear regression is that the dependent variable be approximately normally distributed. The distribution of the activities of daily living was very skewed (see p.135), so that logistic regression was more appropriate.

dependent variable is dichotomous. Therefore, activities of daily living (ADL) scores were recoded as either 'independent' (i.e. participants indicated complete independence on all items by achieving the maximum score) or 'some dependence' (i.e. participants indicated the need for some degree of assistance in at least one of the items; the maximum score was *not* achieved). In logistic regression there are three main statistics to report: Nagelkerke R square, chi-square and the Wald statistic. In comparison with linear regression, the first is similar to the R squared value; the second is a nonparametric statistic that provides an alternative to the F-statistic in analysis of variance; and the third functions as an index of effect size, similar to a Beta value.

Based on previous studies reporting different predictors for basic and instrumental ADL (see Chapter 3, p.68 for details), the current ADL items were also analysed as separate dependent variables. For basic ADL, 137 participants (91.3%) were classified in the 'independent' group and 13 (8.7%) were classified in the "some dependence" group. Table 39 (p.148) shows that (neither sex nor education were significant predictors), grip strength had a significant impact on basic ADL group membership, explaining 23% of basic ADL functional ability. As expected, the 'independent' group was stronger (mean = 20.57 kg, SD = 8.88) than the 'some dependence' group (mean = 11.48 kg, SD = 5.80). Chronological age was also a significant predictor, explaining a significant 13% of basic ADL functional ability. As expected the 'some dependence' group was older (mean = 81.18 years, SD = 4.66) than the 'independent' group (mean = 77.25 years, SD = 4.23). The effect size of these mean differences was moderate (grip strength: $eta^2 = .08$; age: $eta^2 = .06$).

In terms of predicting 'independent' basic ADL group membership, age was able to classify 100%, and grip strength 99.3%, of participants correctly. In contrast, age was not able to classify any participants to the 'some dependence' group and grip

strength could only classify 7.7% correctly. The latter result is likely to be due to the imbalance of participants across the two functional ability groups; less than 10% of the sample was classified as having 'some dependence' in basic ADL. Nonetheless, overall classification accuracy was 91.3% for both biomarker measures.

Table 39
Logistic regression statistics when dependent variable was *basic* activities of daily living group (initial assessment occasion)

Independent Variable	N	NRS	Model chi-square (df)	Wald
Covariates				
Sex#	150	.08	5.46 (1)*	3.35
Education	140	.00	.25 (1)	.24
Biomarkers				
Age	150	.13	9.11 (1)**	8.35**
Systolic BP	137	.00	.18 (1)	.19
Diastolic BP	137	.04	2.61 (1)	2.40
Grip strength	150	.23	15.83 (1)**	10.65**
BMI	150	.05	3.18 (1)	3.38
Visual acuity	150	.00	.11 (1)	.11
Reading Span	149	.06	3.80 (1)	3.16

**p < .01; *p < .05
NRS = Nagelkerke R square; df = degrees of freedom; Wald = Wald statistic
#this variable significantly improved the model (chi-square) but was not itself a significant predictor (i.e. the Wald statistic was not significant)
BP = blood pressure; BMI = body mass index

For instrumental ADL, 88 participants (58.7%) were classified in the 'independent' group and 62 (41.3%) were classified in the 'some dependence' group. Table 40 (p.149) demonstrates that chronological age was the only independent variable to have a significant impact on instrumental ADL group membership, explaining 13% of instrumental ADL functional ability. Similar to basic ADL, the 'some dependence' group was older (mean = 79.22 years, SD = 4.85) than the 'independent' group (mean = 76.43 years, SD = 3.64). The effect size of this mean difference was moderate to large (eta^2 = .09). Age correctly classified 84.3% of participants as 'independent' and 47.4% as having 'some dependence'. Overall, classification accuracy was 69.3%.

5.4.2 Reasoning ability

The second outcome measure, reasoning ability, was measured by Raven's

Standard Progressive Matrices (Raven's). Hierarchical linear regression was conducted in order to determine how much variance each biomarker could explain once the effects of sex and education (i.e. covariates) had been accounted for (Section 5.3 demonstrated that males performed significantly better than females on Raven's and so did participants with more education). All biomarkers, except for systolic blood pressure, accounted for significant, *additional* variance in Raven's performance (see Table 41, p.150, for regression statistics). As shown by the Beta values in Table 41, better reasoning ability was significantly associated with greater working memory capacity (Reading Span), younger chronological age, higher diastolic blood pressure, stronger grip strength, higher body mass index and better corrected visual acuity.

Table 40
Logistic regression statistics when dependent variable was *instrumental* activities of daily living group (initial assessment occasion)

Independent Variable	N	NRS	Model chi-square (df)	Wald
Covariates				
Sex	150	.00	.14 (1)	.14
Education	140	.01	.75 (1)	.74
Biomarkers				
Age	150	.13	15.15 (1)**	13.50**
Systolic BP	137	.00	.01 (1)	.01
Diastolic BP	137	.01	.64 (1)	.63
Grip strength	150	.03	2.86 (1)	2.76
BMI	150	.01	1.03 (1)	1.01
Visual acuity	150	.03	3.63 (1)	3.25
Reading Span	149	.03	2.91 (1)	2.84

**p < .01
NRS = Nagelkerke R square; df = degrees of freedom; Wald = Wald statistic
BP = blood pressure; BMI = body mass index

After taking the variance explained by sex and education (18%) into consideration, Reading Span explained a further 10% of variance in Raven's performance. Figure 18 (p.150) illustrates that this is more than double the variance explained by chronological age and more than the variance explained by any individual physiological biomarker.

Table 41
Linear regression statistics when dependent variable was Raven's (initial assessment occasion)

Independent Variable	R square change	ANOVA (Overall Model)	Beta
Covariates			
Sex	.10	F(1, 148) = 16.41**	.32**
Education	.08	F(1, 138) = 12.73**	.29**
Biomarkers			
Age	.04	F(3, 136) = 12.74**	-.20**
Systolic BP	.02	F(3, 123) = 9.98**	.13
Diastolic BP	.06	F(3, 123) = 12.64**	.24**
Grip strength	.06	F(3, 136) = 13.99**	.39**
BMI	.07	F(3, 136) = 14.65**	.26**
Visual acuity	.05	F(3, 136) = 13.29**	.22**
Reading Span	.10	F(3, 135) = 17.73**	.34**

**p < .01
BP = blood pressure; BMI = body mass index

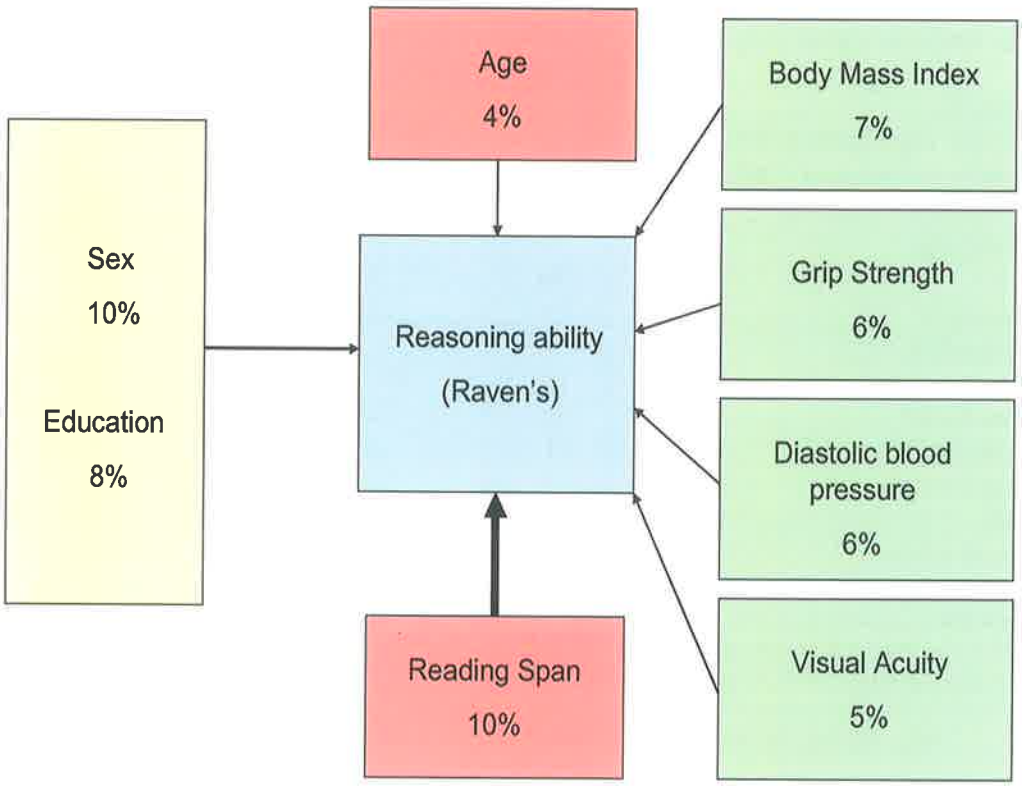


Figure 18 Initial assessment occasion: Representation of how much variance individual biomarkers explained of reasoning ability
The variance explained by each biomarker is in addition to the variance explained by the covariates (i.e. sex and education).

5.4.3 Life satisfaction

The regression procedure described for reasoning ability was also conducted with life satisfaction as the dependent variable (i.e. independent variables and covariates were the same as aforementioned). Results showed that age was the sole significant predictor of life satisfaction (see Table 42). Chronological age explained a significant, but small, 4% of variance in life satisfaction; older age was associated with increased life satisfaction. Neither covariate explained a significant amount of variance. These results suggest that other factors, not employed in the current project as independent variables, influence life satisfaction. For example, health measures (covariates) included in the current study were shown to have significant negative correlations with life satisfaction (see earlier, p.145).

Table 42
Linear regression statistics when dependent variable was life satisfaction (initial assessment occasion)

Independent Variable	R square	ANOVA (Overall Model)	Beta
Covariates			
Sex	.00	F(1, 146) = .15	-.03
Education	.02	F(1, 136) = 2.89	.14
Biomarkers			
Age	.04	F(1, 146) = 6.04*	.20*
Systolic BP	.00	F(1, 133) = .19	-.04
Diastolic BP	.00	F(1, 133) = .22	-.04
Grip strength	.02	F(1, 146) = .61	.07
BMI	.00	F(1, 146) = .32	-.04
Visual acuity	.00	F(1, 146) = .03	-.02
Reading Span	.01	F(1, 145) = 1.05	.09

*p < .05
BP = blood pressure; BMI = body mass index

5.5 Result summary and discussion

For the most part, and relative to other biomarkers, chronological age was the most reliable correlate of a range of dependent variables considered to reflect independent functioning in an elderly sample. For example, age was a significant negative predictor of basic and instrumental activities of daily living (ADL) and

reasoning ability but a positive predictor of life satisfaction. However, the results also indicated that age has limited predictive utility. For example, for basic and instrumental ADL, the classification of participants into the 'some dependence' group was poor compared to the 'independent' group; all biomarkers, except for systolic blood pressure, explained more variance than age in reasoning ability; and age explained less than 5% of variance in life satisfaction. These results support the conclusion that chronological age is not a particularly good predictor of functional outcomes in elderly persons (as discussed in Chapter 1). Finally, Reading Span, a proposed cognitive biomarker, explained the most variance in reasoning ability but did not explain significant variance in any other outcome variable.

The initial limited success of current biomarkers to explain variability in the outcome variables employed is likely to be the result of this sample having restricted ranges. Not only was the current sample restricted in terms of the age-range of participants (as discussed in Chapter 4), but there was only a small degree of variance in key outcome measures such as ADL, because of marked ceiling effects. Consequently, these biomarkers have had the role of differentiating very subtle differences in task performances. At this stage, it appears that the current biomarkers (age, physiological and cognitive) are not particularly sensitive in a high-functioning, independent, healthy and cognitively-intact sample. However, longitudinal results will provide more information about each biomarker's prognostic utility.

The limited variance in performance on outcome variables, predominantly ADL, and other measures (e.g. health), could possibly be the result of the self-report nature of these measures. Despite reports of reasonable correlations between these measures and more objective measures (see Chapter 3, pp.82-83), it is possible that self-report does not accurately reflect a participant's functioning. That is, self-reports are clearly

subjective. For example, two participants may report that they are able to shop independently but the quality of the transaction and degree of difficulty experienced may vary greatly. Moreover, neither participant may realise that they are experiencing difficulties; they may assume their experience as being universal. Consequently, including a more objective assessment of ADL would enhance this outcome measure. For example, other studies have employed the following practical tasks, which are rated by an administrator: walking endurance and efficiency, ability to pour a jug of water, climb and descend tram or bus steps (e.g. Fone & Lundgren-Lindquist, 2003).

In terms of covariates, health and lifestyle measures had limited influence on outcome variables and biomarkers. Other studies have reported similar results and attributed this to the narrow range of health status (as exists in the current sample) or to insensitive health measures (e.g. Salthouse et al., 1990). Perhaps more objective health assessments, such as tests from a general practitioner, would have improved the sensitivity of the current health measure. Similarly, assessment of nutritional intake (a lifestyle measure) may have been improved by employing blood tests rather than relying on self-report techniques. Even if participants accurately recorded their dietary intake, it is difficult to ascertain how much of each nutrient was actually absorbed by their bodies. For example, different cooking and storage processes can influence the amount of active nutrients found in foods; and other health conditions (e.g. atrophic gastritis; Hassing et al., 1999) can physically limit the absorption of nutrients. Exercise is another lifestyle variable that would benefit from objective assessment; it is difficult to know how strenuous the exercise undertaken by participants was or whether the exercise activities could be classified as cardiovascular, resistance work, or both. For example, housework may raise a sweat but not be typically considered as exercise, whereas walking, a more common form of exercise, may not raise a sweat at all.

Similarly, if participants walked whilst carrying something (e.g. a bag of shopping), this introduces resistance into the exercise and produces a more complete form of exercise.

Of the biomarkers, corrected visual acuity had a particularly restricted range. An increased distance between the participant and the Snellen chart may have improved the variability of performance. However, as indicated in Chapter 3 (p.102), this was not feasible in the current test environment. It should be acknowledged that auditory acuity and lung function (i.e. forced expiratory volume) were other biomarkers considered for inclusion at the outset of this project; and several studies have reported reliable results using these measures to predict cognitive and functional outcomes (e.g. Anstey, 1999; Anstey et al., 1997; Baltes & Lindenberger, 1997; Lin et al., 2004; Fozard, Metter & Brant, 1990; Marsiske et al., 1997). However, due to financial and practical constraints, these biomarkers were not included.

Other observations regarding biomarkers involved blood pressure and grip strength. Blood pressure readings were reported for 137 from 150 participants. There were two reasons for this smaller sample size; firstly, the cuff was not large enough to fit around some participants' arms adequately; and secondly, some participants found that when the cuff inflated, it became too tight and they were unwilling to continue with the measurement. Moreover, blood pressure readings were confounded by blood pressure medication. That is, some individuals with lower ('normal') blood pressure readings may have been on blood pressure medication and some individuals with higher ('mild hypertension') blood pressure readings may not. Therefore, it is difficult to assess the real effects of blood pressure in this context. Lastly, it is possible that the particular grip strength device employed here put females at a disadvantage. Due to the dimensions of the device, the (typically) larger hand size of males meant they could get a better grip, and consequently produce a stronger score. This advantage would perhaps

be minimised with more recent models of dynamometer.

In summary, the biomarkers employed in the current project have been found to have only limited success in predicting concurrent functional outcomes. This is most likely due to the high-functioning elderly sample assessed. Given that the study design was longitudinal, the opportunity to make changes to the test battery in order to improve the sensitivity of measures was limited. Nonetheless, in an attempt to improve the sensitivity of the outcome measures, an additional dependent variable was included on the final assessment occasion (and life satisfaction was no longer included as an outcome measure, due to the absence of significant predictors). Longitudinal assessment relies on the calculation of change scores, which in turn, requires the same form of measurement on each test occasion. Moreover, sex and years of education were the only covariates included in the longitudinal analyses due to the limited influence of the other covariates in the cross-sectional results. Chapter 6 presents the longitudinal results of this study.

CHAPTER 6: LONGITUDINAL RESULTS

THE CALCULATION AND USE OF CHANGE SCORES

6.1 Overview

As described in Chapter 3, the design of the current study was longitudinal; if the true prognostic utility of working memory is to be demonstrated (a criterion for definition of a biomarker) and compared with that of other, more commonly used measures, then the predictive ability of these measures must be assessed over time. In light of this, Chapter 6 is presented in six main parts; first, a discussion of confounds to longitudinal analyses is provided; second, the calculation of change scores is discussed; third, descriptive statistics from the selected change score method are presented; fourth, the results of the regression analyses employing change scores are shown; fifth, participants who continued with the study are compared to those who discontinued in order to assess for selective attrition; and sixth, longitudinal results are summarised and compared to cross-sectional results.

6.2 Confounds to longitudinal analyses

A number of researchers have commented that longitudinal analysis, or more specifically, the analysis of change, warrants the use of 'special' statistical methods (e.g. Collins, 1996; Diggle et al., 1994; Roberts & Chapman, 2000). However, there is little consensus on which statistical method is the most appropriate (Alder et al., 1990; Cronbach & Furby, 1970). Regardless of the specific method selected, it has been stated that the "interpretation of results of repeated testing is a complicated endeavor" (Dikman, Heaton, Grant & Temkin, 1999, p.354). This is largely due to the fact that researchers must distinguish between 'true', systematic or statistically reliable change and random change. Sources of random change include measurement error or

regression-to-the-mean, initial test performance, test-retest reliability, practice effects and demographic measures such as age, sex and education (Rabbitt et al., 2001; Sherman et al., 2003; Tombaugh, 2005). These issues will now be discussed in more detail, before presenting the most common statistical methods used to calculate change. Needless to say, superior methods for measuring change will take such sources of error into consideration.

6.2.1 Measurement error/Regression-to-the-mean

Generally, this issue refers to the unreliability of test measurement. That is, at each measurement occasion, it is possible that the observed score achieved by an individual is not a true reflection of his/her ability. As a result,

> Individuals who score lower than their 'true score' at Time 1 and higher then their 'true score' at Time 2, will appear to have changed, even if they have not. Likewise, some individuals will appear to have remained stable because of measurement error, even when their true scores have changed (Anstey & Hofer, 2004, p.99).

Follow-up studies (involving two measurement occasions only) are particularly susceptible to measurement error, an occurrence that has also been referred to as regression-to-the-mean.

> Regression-to-the-mean is manifested by the tendency for baseline scores at either extreme of the distribution to move towards the mean upon retesting. The net effect is to amplify measured change (negative or positive) for individuals initially performing at the upper or lower end of the distribution (Sawrie, Marson, Boothe & Harrell, 1999, p.P117).

The susceptibility of change scores to regression-to-the-mean is dependent on the test-retest reliability of the measures employed and the initial level of performance. The

poorer the test-retest reliability and the more extreme the initial scores, the more probable that regression-to-the-mean effects will influence change scores (Raykov, 1992; Sawrie et al.).

6.2.2 Initial test performance

A feature intrinsic to longitudinal data is autocorrelation (Anstey & Hofer, 2004; Diggle et al., 1994; Rabbitt et al., 2001). That is, for each individual, performance at one measurement occasion is correlated with performance at another measurement occasion. In other words, performance across measurement occasions tends to be positively correlated within individuals. However, correlations (Pearson's r) tend to be less than 1 between test and retest measures. This is because some individuals typically change more than others (Steyer, Eid & Schwenkmezger, 1997). Additionally, the degree of change shown by individuals is not generally independent of their initial performance (Bond, 1979; Raykov, 1992). This association is not necessarily a product of measurement error; it can reflect the systematic changes of interest.

6.2.3 Test-retest reliability

It has been argued that small test-retest correlations are indicative of different constructs being measured at each measurement occasion, implying that the measure used has poor reliability (Collins, 1996). However, this is not necessarily the case because test-retest correlations are also dependent upon the heterogeneity of change within the sample tested. Providing error measurement is low, small test-retest correlations can suggest large interindividual variability in change and good measure reliability (Collins). Conversely, large test-retest correlations suggest interindividual variability in change is limited to several forms that do not greatly disrupt the rank order of individuals over time (e.g. all individuals change approximately the same amount) and this may be associated with high or low reliability of difference scores (Collins).

6.2.4 Practice effects

Another potentially confounding influence when determining longitudinal change is practice effects. That is, when the same, or a similar measure ("parallel form"), is administered repeatedly, individuals typically demonstrate improvement on the task due to practice or familiarity (Rabbitt et al., 2001). In particular, this improvement can mask or underestimate age-related decline (as is relevant in the current study). Parallel forms of psychometric tests were created to minimise these effects, yet they do not necessarily prevent them from occurring altogether. The magnitude of practice effects is dependent on the time interval between test occasions, the nature of the test itself, and the characteristics of the sample tested, such as their age and health (e.g. greater practice effects have been found for older and more frail individuals; Rabbitt et al.). It has been suggested that practice effects can influence performance on test-retest intervals separated by as much as 6 years (Zelinski & Burnight, 1997).

6.2.5 Demographic measures

In psychology, particularly in cognitive psychology, demographic measures such as age, sex and education can influence initial performance on a test. If these measures are not somehow accounted for in the calculation of change in test performance, change scores may not accurately reflect the degree of change experienced by each individual, and subsequently by the sample as a whole. For example, both receiving more years of formal education and maintaining cognitive activity with age have been reported to have protective effects on cognitive decline that is typically experienced with increasing age (e.g. Wilson, Barnes & Bennett, 2003). If education level (or number of years) is not considered, some individuals may be misrepresented as demonstrating little or lots of difference between test measurements.

6.3 Measuring change: Different types of change scores

There are two commonly used methods for analysing longitudinal change that have been used in the field of psychology; simple difference change scores and regression-based change scores (Gupta, Srivastava & Sharma, 1988; Roberts & Chapman, 2000; Williams & Zimmerman, 1982; Zimmerman & Williams, 1982). The attributes and related components of these methods are discussed in the following sections. Growth modelling (e.g. latent growth curves, structural equation modelling) is another method for analysing longitudinal change. However, given that valid interpretation of some of these models requires equal spacing between retest intervals (McArdle, 1988); and that, "the individual growth curve conceptualisation of change has received only limited attention and use in the gerontological literature" (Alder et al., 1990, p.561); growth modelling will not be discussed here. In any case, recent studies have suggested that assessing change via growth modelling and residualised methods produce effectively indistinguishable results (Roberts & Chapman).

6.3.1 The simple difference score

The simple difference score (D) or raw change score is calculated by D = T1 – T2, where T1 represents the first measurement occasion and T2 represents the second/final measurement occasion. This formula is commonly reversed such that D = T2 – T1. Conceptually, this score is the simplest and most easily interpretable change measure (Collins, 1996; Williams & Zimmerman, 1982). However, some disadvantages to this method have been documented. There is no control for measurement error, and as such this score has been classified as *not* 'base-free' because it does not consider the influence of initial performance; and these scores usually have low reliability (Cronbach & Furby, 1970; Irving & Meyer, 1999; Willett, 1988; Williams & Zimmerman). As a result, conclusions made regarding the change in performance based on the simple

difference score can be misleading.

6.3.1.1 Mean difference score

This is an extension of the simple difference score and relies on mean performance of the sample studied to indicate the change in the sample's performance. Whether there has been a significant change in performance is assessed by conducting a repeated-measures analysis of variance. This analysis provides information about the characteristics of the population in question and can often show an increase in mean performance, at least in part indicating the presence of practice effects (Roberts & Chapman, 2000; Tombaugh, 2005). Whilst such an analysis can be considered as a first attempt at measuring change, it does not provide "a generally viable approach to longitudinal data analysis" (Diggle et al., 1994, p.130). This is likely to be due to the limitations mentioned in Section 6.2.1 and that mean performance obscures a whole range of change trajectories within a sample, leading to an incomplete understanding of the process under investigation (Roberts & Chapman). For example, in relation to cognitive performance in elderly persons, mean performance would typically suggest deterioration whereas some elderly individuals would not change at all or may even improve (Christensen et al., 1999).

6.3.2 Regression-based/Residualised change scores

With this approach, linear regression is used to predict retest scores from observed initial scores. The predicted retest scores (Yp) are then subtracted from the observed retest scores (Yo) to produce an unstandardised regression-based change (RBC) score. That is, RBC = Yo – Yp. Demographic (e.g. sex and years of education) and other potentially confounding variables (e.g. retest interval) can also be included in the regression equation used to predict retest scores. The RBC score can be standardised (SRBC) into a z-score by dividing the RBC score by the standard error of the estimate

(SEest) from the regression analysis, so SRBC = (Yo – Yp)/SEest. This score (a standardised residual) permits the degree of change shown by an individual on one test measure to be comparable to the degree of change shown by this individual on another test measure (Sawrie et al., 1999).

In other words, a score is residualised by "expressing the posttest score as a deviation from the posttest-on-pretest regression line" (Cronbach & Furby, 1970, p.68). The residualised score provides a means of identifying individuals who have changed more or less than expected, based on the trend for that sample (Cronbach & Furby; Sherman et al., 2003). According to Irving and Meyer (1999), "...residual scores attempt to assess what change would have occurred had all individuals started out with identical initial scores" (p.88). As such, this measure is referred to as 'base-free', or independent of initial performance. In contrast to the simple difference scores, residualised difference scores always have a Pearson correlation of zero with the corresponding pretest scores (Williams & Zimmerman, 1982). On the whole, regression-based change scores have been reported to be advantageous over other change scores because they take the effects of random change confounds (as discussed in Section 6.2) into consideration (Dikman et al., 1999; Sawrie et al., 1999; Williams, Zimmerman, Rich & Steed, 1984a; Williams, Zimmerman, Rich & Steed, 1984b).

6.3.2.1 Reliable Change Index

A principal limitation to the regression-based change scores is that they do not necessarily suggest when a significant change has occurred from the initial performance (Hermann et al., 1996; Sherman et al., 2003), only that the psychometric properties of the change are reliable. In order for *significant change* to be established, a reliable change index (RCI) is used. A RCI is a means of estimating whether a change score has been obtained by chance (Iverson & Green, 2001). "It defines the range in

which an individual score is likely to fluctuate because of the imprecision of a measuring instrument" (Jacobson & Truax, 1991, p.16). The index is formed by using a z-distribution. In the case of a simple difference change score, this score "can be multiplied by a z-score to provide a confidence interval for possible measurement error" (Iverson & Green, p.458). Multiplying by a z-score of 1.64 suggests a confidence interval of 90%, or that there is a 10% chance of that change score falling within the range of measurement error that indicates no real change in performance.

6.4 Summary: Selection of a change score method

In regard to longitudinal studies focusing on independent functioning in elderly people, there has been little discussion as to the specific method used to calculate change in level of independent functioning. However, what is evident is that many researchers have employed the simple difference (or mean difference) change score, despite the aforementioned confounds not being taken into consideration in this change measurement (e.g. Femia et al., 1997; Njegovan et al., 2001). In contrast, standardised regression-based change scores have been found to "be useful for classifying normal variability in healthy older adults and were strongly associated with diagnostic change (from normal to dementia)" (Frerichs & Tuokko, 2006, p.110). Based on this information and the fact that standardised regression-based methods take various sources of error into consideration, this longitudinal statistical method was considered the most appropriate for use in the current study.

6.5 Change Scores: Reliability assessment and descriptive statistics

This section of the longitudinal analysis involves three main steps. The first step examines the presence of true change confounds in the current test variables (e.g. practice effects, regression-to-the-mean), the second step involves developing standardised regression-based regression equations to calculate change for the test

variables, and the third step assesses which of the change scores generated is clinically relevant (incorporating the notion of a reliable change index). Each of these steps is carried out for outcome and biomarker variables. A detailed description of these variables has been provided in Chapter 3.

6.5.1 Presence of true change confounds

This analysis identifies the presence of normal decline, practice effects and regression-to-the-mean. To assess these first two confounds to true change, mean change scores were calculated and then compared using paired-sample t-tests for the outcome variables (because there were only two measurement occasions; 0 & 18 months), and repeated measures analysis of variance for the biomarker variables (because these were measured on three occasions; 0, 6 & 18 months). This analysis permits the determination of significant mean change across measurement occasions.

As can be seen in Table 43 (p.166), only two outcome variables showed significant mean change. Basic activities of daily living (ADL) scores declined slightly but significantly from initial to final assessment occasions ($t = 2.68$, $df = 124$), indicating that the sample became more dependent in these activities with time (i.e. over approximately 18 months). The effect size of this decline is moderate ($eta^2 = .05$). Performance on Raven's Standard Progressive Matrices (Raven's) also declined significantly over the same time period ($t = 9.85$, $df = 124$), suggesting poorer reasoning ability with increasing age (effect size is large, $eta^2 = .44$). These results are consistent with the attributes of the sample. That is, decline in both fluid ability and independence in daily activities is anticipated in an elderly population.

In terms of the biomarker variables employed to predict these outcome measures, three demonstrated a significant effect of time on performance, as shown by Wilks' Lambda (Table 44, p.167). These variables included a measure of working

memory capacity, Reading Span [$F(2,121) = 5.20$, partial $eta^2 = .08$], systolic blood pressure [$F(2, 104) = 12.73$, partial $eta^2 = .20$], and diastolic blood pressure [$F(2, 104) = 12.62$, partial $eta^2 = .20$]. The partial eta^2 statistic suggests a moderate effect size for change in Reading Span performance and a large effect size for change in both of the blood pressure measures. Systolic and diastolic blood pressure decreased from initial to intermediate assessment and then again from intermediate to final assessment. In terms of its physiological implications, this result is relatively meaningless without a classification of hypertension. However, it is plausible that this result, on the whole, is a reflection of participants becoming more relaxed with the testing procedure (stress has been shown to increase blood pressure, see Chapter 3 for a discussion on this). As for Reading Span, there was a significant increase in working memory performance from initial to intermediate measurement occasions. This is likely to be due to practice effects. From intermediate to final measurement occasions, Reading Span performance declined to about the level obtained at the initial assessment. This suggests that working memory performance, on average, did not decline. This last result is surprising given that decline on tasks of working memory with increasing age has been well documented; reasons for this limited age-association have been discussed in Chapter 4.

Table 43
Mean change statistics for measures of independent functioning (i.e. outcome variables)

Outcome	Initial Assessment			Final Assessment		
Variable	N(a)	Mean	SD	N(b)	Mean	SD
Basic ADL**	150	6.82	.61	126	6.72	.73
Instrumental ADL	150	25.80	1.87	126	25.71	2.18
Raven's**	150	16.61	4.73	125	14.02	4.38
Crystallised ability^	146	34.45	4.16	123	35.08	3.49
Memory-based ADL #				125	81.41	9.90

**p < .01
ADL = Activities of Daily Living; Raven's = Raven's Standard Progressive Matrices
N(a) = 150 participants completed the initial assessment occasion
N(b) = 127 participants completed the final assessment occasion
^This is a composite measure of three tasks; see Chapter 3 for a description of the individual tasks
#Memory-based ADL was administered only at the final assessment occasion and therefore there is no initial or change score data for this measure

Table 44

Mean change scores for biomarker variables employed to predict independent functioning measures

Biomarker	Initial Assessment			Intermediate Assessment			Final Assessment		
Variable	N(a)	Mean	SD	N(b)	Mean	SD	N(c)	Mean	SD
Systolic BP**	137#	147.98	22.70	124	141.34	20.97	114	140.90	20.45
Diastolic BP**	137#	79.23	11.53	124	76.23	10.75	114	75.86	9.90
Grip strength	150	19.78	9.01	136	20.28	9.48	126	20.27	9.69
BMI	150	26.47	4.13	137	26.31	4.22	127	26.59	4.28
Visual Acuity	150	93.48	2.78	137	93.66	2.58	126	93.41	3.59
Reading Span**	149	13.91	8.81	136	16.26	8.50	127	14.91	8.77

**Wilks' Lambda, $p < .01$

BP = Blood Pressure; BMI = Body Mass Index

N(a) = 150 participants completed the initial assessment occasion

N(b) = 137 participants completed the intermediate assessment occasion

N(c) = 127 participants completed the final assessment occasion

#Decreased sample size for blood pressure was due to technical issues with the equipment and participant ethics (a discussion of these matters was provided in Chapter 5)

Two indicators were used to assess for the presence of regression-to-the-mean. First, test-retest reliability coefficients (Pearson's r) were calculated; for the independent functioning outcome measures, reliability coefficients were based on initial and final measurements; for biomarker variables, reliability coefficients were calculated for both initial-intermediate measurements and initial-final measurements. Regression-to-the-mean occurs only when the correlation between two measurement occasions is less than perfect, that is, when $r < 1$ (Nesselroade, Stigler & Baltes, 1980). Second, initial test scores were correlated (Pearson's r) with simple change scores for each measure (change score = intermediate or final assessment – initial assessment). This was carried out because regression-to-the mean occurs only when this correlation is negative (Speer, 1992).

The test-retest reliabilities for the outcome measures ranged from .66 to .85 (see Table 45, p.169). At first glance, these correlation coefficients may appear low to moderate but, due to the restricted ranges of the sample on several measures (refer to Chapters 4 & 5 for an illustration of this), it is likely that these correlations have been suppressed to some extent. In any event, they provide evidence for regression-to-the-mean. Raven's Standard Progressive Matrices (Raven's), crystallised ability and instrumental activities of daily living (ADL) were particularly susceptible to regression-to-the-mean because they showed significant and negative initial and change score correlations.

The test-retest reliabilities for the biomarker variables ranged from .56 to .97 between initial and intermediate measurement occasions; and from .47 to .95 between initial and final measurement occasions (see Table 46, p.170). The influence of restricted samples on the magnitude of correlations also applies here. However, once again, all variables showed evidence of regression-to-the-mean. Salient variables in this

regard, as evidenced by significant and negative initial and change score correlations for both reliability coefficient calculations, include Reading Span and systolic and diastolic blood pressure. Small decreases in test-retest reliability coefficients from the initial-intermediate measurements to initial-final measurements, for most biomarkers, supports evidence that longer retest intervals produce less stable scores (Tombaugh, 2005). In summary, the longitudinal data were exposed to the influences of random change. Therefore, controlling for these confounds was necessary in subsequent change score analysis.

Table 45
Indicators of test-retest psychometric confounds for measures of independent functioning (i.e. outcome variables)

Outcome Variable	N(a)	Mean Change	Test-retest^ Reliability	Initial & Change Score Correlation
Activities of daily living				
Basic	126	-0.10	.74**	-.12
Instrumental	126	-0.09	.66**	-.22*
Raven's	125	-2.59	.69**	-.44**
Crystallised ability	119	+0.63	.85**	-.32**
Memory-based ADL#	125			

**p < .01;*p < .05
N(a) = 127 participants completed the final assessment occasion
ADL = Activities of Daily Living; Raven's = Raven's Standard Progressive Matrices
^In the case of Raven's this refers to parallel test reliability
#these details are not available because this questionnaire was administered on the final measurement occasion only

6.5.2 Developing standardised regression-based regression equations

The purpose of the standardised regression-based change score method is to use regression procedures to predict the final performance on a particular measure. This predicted final score is based on initial performance and other information (e.g. demographic variables). The difference between the predicted final score and the observed final score is what constitutes the change score. The current study employed linear regression using the enter method (p = .05 as the criterion for entrance and p = .10 as the criterion for removal) to produce the change scores. A separate regression analysis was carried out for each independent functioning outcome measure; the

Table 46

Indicators of test-retest psychometric confounds for biomarker variables employed to predict independent functioning measures

		Initial vs Intermediate				Initial vs Final Assessment		
Biomarker Variable	N(a)	Mean Change	Test-retest Reliability	Initial & Change Score Correlation	N(b)	Mean Change	Test-retest Reliability	Initial & Change Score Correlation
Systolic BP	121	-6.64	.69**	-.49**	110	-7.08	.69**	-.53**
Diastolic BP	121	-3.00	.71**	-.45**	110	-3.37	.67**	-.56**
Grip strength	136	+0.50	.96**	-.03	126	+0.49	.94**	-.03
BMI	137	-0.16	.97**	-.06	127	+0.12	.95**	-.11
Visual Acuity	137	+0.04	.65**	-.33**	126	-0.40	.47**	-.13
Reading Span	132	+1.79	.56**	-.51**	127	+0.14	.58**	-.47**

**p < .01

BP = Blood Pressure; BMI = Body Mass Index

N(a) = 137 participants completed the initial and intermediate assessment occasions

N(b) = 127 participants completed the initial and final assessment occasions

observed final assessment score for the outcome measure was entered as the dependent variable and the initial assessment score for the outcome measure was entered as an independent variable (block one). After the block one independent variable had been entered, sex, years of education and the time interval between measurement occasions (in months), were entered (block two independent variables). The *predicted* final scores (i.e. generated by the regression analyses) were then subtracted from the *observed* final scores. This change score was then standardised by dividing it by the standard error of the estimate from the regression analysis. The resulting change score, the standardised regression-based change score, provided an indication of who changed more or less than expected. It should be noted that the change in activities of daily living measures over time was not analysed in this way due to its non-normal distribution, as demonstrated in Chapter 5. The longitudinal method employed for this independent functioning outcome measure is discussed in Section 6.6.1 (pp.179-180).

Similar regression analyses were carried out for the biomarker variables. However, two separate regression equations were formed for each biomarker variable; one where the observed biomarker score at the intermediate assessment was the dependent variable (to generate a 6 month change score) and another where the observed biomarker score at the final assessment was the dependent variable (to generate an 18 month change score). For all regression equations, the biomarker score at the initial assessment formed the independent variable in block one and the independent variables in the second block were the same as those mentioned for the independent functioning outcome variables. Corrected visual acuity was not included in these analyses because of the measure's negatively skewed distribution; at minimum, dependent variables are expected to have near-normal distributions in linear regression.

Tables 47 to 49 (pp.173-175) show the details of the regression equations

described above. From these Tables it is clear that the initial measurement score is the most significant predictor for all final measurement scores (as indicated by highly significant unstandardised Beta coefficients). Both sex and education made significant contributions to the prediction of the final measurement score for some variables (e.g. diastolic blood pressure, and Reading Span, respectively). In contrast, the time interval between assessment occasions was not a significant predictor of any final assessment score (as indicated by non-significant unstandardised Beta coefficients); this presumably reflects the consistency with which participants were assessed in subsequent test sessions. That is, it was endeavoured to maintain relatively similar test-retest intervals for all participants; where possible, the retest order was the same as that established in the initial test session[34].

6.5.3 Clinically relevant change scores

The standardised regression-based change (SRBC) scores were assessed in terms of confidence intervals such that individuals with SRBC scores greater than z = 1.64 (90% confidence interval) were considered to have improved significantly; individuals with SRBC scores less than z = -1.64 were considered to have declined significantly; and individuals with SRBC scores between z = -1.64 and z = 1.64 were considered to have remained unchanged. Examination of Tables 50 and 51 (p.176) clearly shows that the majority of participants remained stable across all variables and across both the 6 month retest interval (relevant to biomarkers only) and the 18 month retest interval. Of participants whose ability remained stable, the highest percentage was found for body mass index (BMI) and grip strength over 6 months and Reading Span over 18 months. Of participants whose ability improved, the highest percentage was found for Reading

[34]This is supported by the low standard deviations of the test intervals. Initial to intermediate, mean = 5.58 months, SD = .79 (range = 3.25 - 8.46 months); Intermediate to final, mean = 13.42 months, SD = 1.08 (range = 10.33 - 16.67 months); and initial to final, mean = 19.03 months, SD = .83 (range = 17.80 - 21.90 months).

Table 47

Details of the regression equations for predicting the final assessment score of independent functioning outcome variables

Outcome Variable	N(a)	R	SEest	Constant	Unstandardised Beta Coefficients			
					Initial Score	Sex	Education	Time Interval
Activities of daily living								
Basic^								
Instrumental^								
Raven's	118	.71	3.15	3.10	.60**	-.03	.17*	-.05
Crystallised ability	114	.86	1.81	9.74	.66**	.47	.10*	.04
Memory-based ADL #	125							

**p < .01

N(a) = 127 participants completed the initial and final assessment occasions

R = Multivariate correlation coefficient; SEest = standard error of the estimate

Raven's = Raven's Standard Progressive Matrices

^these details are not available because standardised regression-based change scores were not calculated; see p.171 for more details

#these details are not available because this questionnaire was administered on the final measurement occasion only

Table 48
Details of the regression equations for predicting the intermediate assessment biomarker scores

Biomarker Variable	N(a)	R	SEest	Constant	Unstandardised Beta Coefficients			
					Initial Score	Sex	Education	Time Interval
Systolic BP	112	.69	15.43	44.35	.63**	-3.48	.18	1.09
Diastolic BP	112	.73	7.50	20.41	.68**	-3.05*	.04	.97
Grip strength	126	.97	2.48	1.81	.89**	2.84**	.01	.24
BMI	127	.97	.97	.68	.99**	.09	-.04*	-.04
Visual Acuity#								
Reading Span	127	.60	6.94	-.88	.50**	.53	.38*	.90

**p < .01; *p < .05
N(a) = 137 participants completed the initial and intermediate assessment occasions
R = Multivariate correlation coefficient; SEest = standard error of the estimate
BP = Blood Pressure; BMI = Body Mass Index
#these details are not available because standardised regression-based change scores were not calculated; see p. 171 for more details

Table 49

Details of the regression equations for predicting the final assessment biomarker scores

Biomarker Variable	N(a)	R	SEest	Constant	Unstandardised Beta Coefficients			
					Initial Score	Sex	Education	Time Interval
Systolic BP	103	.70	14.94	-5.15	.61**	-2.00	.01	3.04
Diastolic BP	103	.68	7.39	11.67	.56**	-1.81	.13	1.08
Grip strength	119	.95	3.10	-2.41	.88**	3.13**	-.16*	.15
BMI	120	.95	1.34	-.35	.98**	-.15	-.04	.09
Visual Acuity#								
Reading Span	120	.61	7.06	-5.88	.54**	.13	.41*	.44

**p < .01; *p < .05

N(a) = 127 participants completed the initial and final assessment occasions

R = Multivariate correlation coefficient; SEest = standard error of the estimate

BP = Blood Pressure; BMI = Body Mass Index

#these details are not available because standardised regression-based change scores were not calculated; see p. 171 for more details

Table 50

Proportion of participants in each clinically relevant change category (over 18 months) for independent functioning outcome measures

Variable	Total N	Improved (%)	Stable (%)	Declined (%)
Activities of daily living				
Basic^				
Instrumental^				
Raven's	118	4.2	89.0	6.8
Crystallised ability	114	2.6	88.6	8.8
Memory-based ADL #				

Raven's = Raven's Standard Progressive Matrices

^these details are not available because standardised regression-based change scores were not calculated; see p.171 for more details

#these details are not available because this questionnaire was administered at the final measurement occasion only

Table 51

Proportion of participants in each clinically relevant change category (over 6 and 18 months) for biomarkers

Variable	Total N(a)	Total N(b)	Improved (%)		Stable (%)		Declined (%)	
			6 mths	18 mths	6 mths	18 mths	6 mths	18 mths
Systolic BP	112	103	5.4	4.9	91.1	89.3	3.6	4.6
Diastolic BP	112	103	4.5	5.8	91.1	88.4	4.5	5.8
Grip strength	126	119	3.2	5.0	92.1	89.9	4.8	5.0
BMI	127	120	1.6	4.2	92.1	87.5	6.3	8.3
Visual Acuity#								
Reading Span	127	120	8.7	7.5	89.8	90.8	1.6	1.7

Mths =months; BP = Blood Pressure; BMI = Body Mass Index

N(a) = 137 participants completed the initial and intermediate assessment occasions

N(b) = 127 participants completed the initial and final assessment occasions

#these details are not available because standardised regression-based change scores were not calculated; see p.171 for more details

Span over both the 6 and 18 month intervals; this apparent anomaly has been discussed earlier in this Chapter. Of participants whose ability declined, the highest percentage was found for crystallised ability and BMI over the 18 month retest interval; BMI also showed the highest percentage of decline over the 6 month retest interval. Declines in performance on tasks of crystallised ability hold particular importance because deterioration in this type of ability has been associated with the terminal drop, the time that closely precedes death (see Chapter 3 for more of a discussion on this topic).

In conclusion, for all variables, a very small proportion of participants showed significant amounts of decline over the 18 month assessment period (i.e. less than 10% of participants). Moreover, about the same, or smaller proportions of participants improved significantly over the same test interval (Reading Span being the main exception). These results further support conclusions that the current sample was high-functioning, well-educated and healthy and that ageing need not always be associated with deterioration or functional dependence (e.g. Gill et al., 1996).

6.6 Regression analyses: Predicting independent functioning outcome measures over time

The longitudinal data were analysed in a similar way to the cross-sectional; different regression methods were employed to identify significant predictors of independent functioning outcome measures. However, compared to the cross-sectional analyses, there were two main differences in the longitudinal results that follow; firstly, there were more independent variables in the regressions (i.e. change scores were incorporated); and secondly, there were more dependent variables in the regression (i.e. including change scores and new outcome measures). These differences will now be discussed and then the regression results will be presented. To assist in the interpretation of regression results, bivariate correlations between dependent and

independent variables are provided in Appendix Q.

Change scores, as calculated by standardised regression methods, were included in regression analyses as both independent and dependent variables. For example, in addition to the use of initial biomarker scores, independent variables also included the amount of change found in each biomarker over 6 and 18 months. Similarly, in addition to the use of final scores on independent functioning measures, dependent variables also included, where possible, the change in performance over 18 months (i.e. change scores require the same measure to be assessed over a minimum of two occasions). Figure 19 (p.179) best illustrates how change scores were used in longitudinal regression analyses. In order to predict *final* independent functioning scores, separate regression analyses were conducted employing: (1) initial performance, (2) 6 month change, and (3) 18 month change, biomarker scores as independent variables. In order to predict the *change* in independent functioning scores over *18 months*, only the initial scores of the independent variables were employed (denoted by the dashed lines in Figure 19). The decision was made not to use change scores to predict change scores because the results generated from such analyses would be difficult to interpret; the 6 month change would also be represented in the 18 month change score, making it difficult to distinguish which component of the 18 month score the independent variables were explaining.

Four independent functioning outcome measures were included in the longitudinal results; activities of daily living (represented by basic and instrumental tasks); reasoning ability; crystallised ability; and memory-based activities of daily living. Basic and instrumental activities of daily living (ADL) and reasoning ability were assessed on both the initial and final measurement occasions. Therefore, in addition to final scores, 18 month change scores could be calculated and used as dependent variables for these measures. Although crystallised ability was not employed

as a dependent variable in the cross-sectional results, it was assessed on both initial and final measurement occasions; this permitted the change in crystallised ability over 18 months to be employed as a dependent variable in the longitudinal results. Initial crystallised ability was not of particular interest whereas change in this ability was. Finally, the memory-based ADL measure was only administered on the final measurement occasion, preventing the use of an 18 month change score as a dependent variable; reasons for its inclusion will be discussed in the following sections.

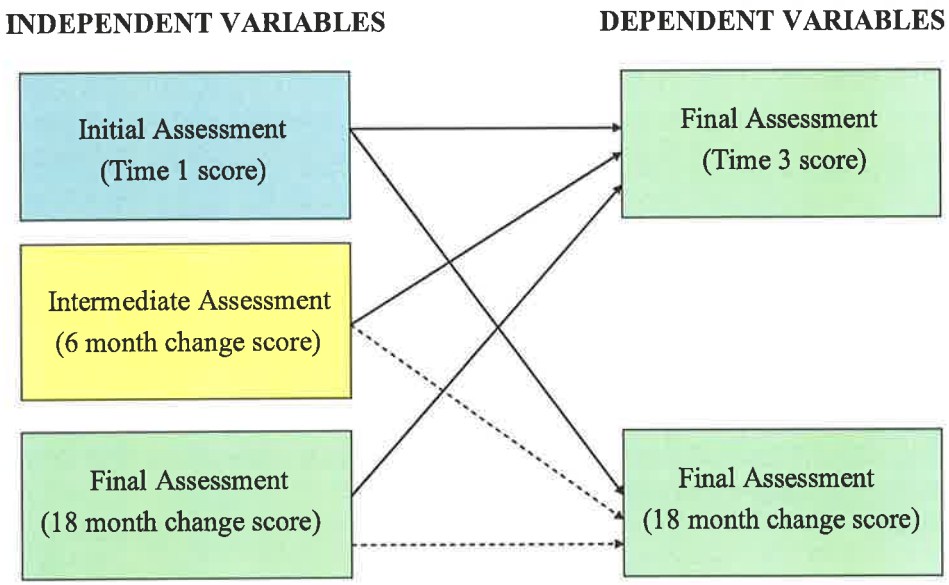


Figure 19 How change scores will be used in longitudinal regression analyses

6.6.1 Activities of daily living

Standardised regression-based change scores could not be calculated for the two types of activities of daily living (ADL) measures because both of these scores showed strongly truncated distributions; dependent variables require near normal distributions for linear regression (as discussed in Chapter 5). Instead, logistic regression, which involves the prediction of group membership for a dichotomous variable, will be used. For congruence with the relationships between independent and dependent variables depicted in Figure 19, participants' self-reported ability in basic and instrumental ADL

were categorised into 'independent' or 'some dependence' (i.e. the same groups employed in cross-sectional results) for final ADL scores and into 'improve or no change' and 'decline' for change in ADL ability over 18 months. Sex and years of education were included as covariates.

Table 52 shows descriptive statistics on whether participants reported an increase in, the same, or a decrease in independence level for both basic and instrumental ADL. Similar to other independent functioning outcome measures (see p.176), the majority of participants reported no change in the ability to carry out basic and instrumental ADL over 18 months and a higher proportion of participants reported a loss of independence, rather than a gain (this may be due, in part, to the ceiling effects on this measure, as reported in Chapter 5). Longitudinal results were also consistent with the hierarchical relationship reported previously between basic and instrumental ADL (see Chapter 5); more participants experienced a loss of independence in instrumental ADL than basic ADL. This result reflects the community-dwelling nature of the sample.

Table 52
Change in self-reported independence in basic and instrumental activities of daily living (ADL) over 18 months

Variable	Increased independence (%)	Same degree of independence (%)	Decreased independence (%)
Basic ADL	1.6	90.5	7.9
Instrumental ADL	11.9	65.9	22.2

N=126, listwise for participants who completed both initial and final test sessions

6.6.1.1 Basic activities of daily living

Of the 126 participants to complete the final assessment occasion, 108 (85.7%) were categorised as 'independent' in basic ADL and 18 (14.3%) were considered dependent to some extent (i.e. maximum score was not achieved). As shown in Table 53 (p.181), sex, a demographic variable, best predicted final basic ADL group membership, explaining a significant 14% of this functional ability; females (N=17)

were more likely to be in the 'some dependence' group (continuity correction chi-square = 6.53, $p < .05$). However, after statistically controlling for the influence of sex, initial chronological age explained a further (and significant) 7% of variance in final basic ADL ability. In other words, participants who were in the 'independent' basic ADL group at the final test session were significantly younger at the start of this study (initial mean age = 76.87 years, SD = 3.77) compared to individuals who belonged to the 'some dependence' group at the final test session (initial mean age = 79.49 years, SD = 5.51). It is interesting to note that, after covarying for sex, initial Reading Span performance was of borderline significance for predicting final basic ADL group membership (explaining an additional 6% of the variance, $p = .052$). Participants who were 'independent' in basic ADL at the final test session demonstrated better Reading Span scores (i.e. greater working memory capacity) at the initial test session (mean = 15.52 words correct, SD = 8.95) than those who showed 'some dependence' in basic ADL at the final test session (initial mean = 10.28 words correct, SD = 7.60). Effect size of these mean differences was small to moderate (age: $eta^2 = .03$; Reading Span: $eta^2 = .04$).

Table 53
Logistic regression statistics: Initial biomarker performance to predict final *basic* activities of daily living group membership

Variable	N	NRS change	Model chi-square (df)	Wald
Covariates				
Sex	126	.14	10.10 (1)**	5.34*
Education	119	.01	.85 (1)	.82
Biomarkers				
Age	126	.07	15.31 (2)**	4.89*
Systolic BP	114	.00	11.24 (2)**	.01
Diastolic BP	114	.02	12.83 (2)**	1.51
Grip strength	126	.03	12.30 (2)**	2.13
BMI	126	.04	13.36 (2)**	3.21
Visual acuity	125	.00	9.97 (2)**	.12
Reading Span	126	.06	14.44 (2)**	3.77^

**p < .01; *p < .05
^p = 0.052
NRS = Nagelkerke R square; df = Degrees of freedom; Wald = Wald statistic
BP = Blood pressure; BMI = Body mass index

Similar to the cross-sectional results, correct group classification was poor for the 'some dependence' basic ADL group compared to the 'independent' group; both sex and Reading Span could predict 100% of 'independent' participants correctly but none of the 'some dependence' (overall classification accuracy was 97.7%). Again, this is likely to be due to the imbalance of participants across these groups. In contrast, age at initial testing could predict 100% of 'independent' group membership and 5.6% of 'some dependence' group membership (overall classification accuracy was 86.5%).

When biomarker change scores (i.e. over 6 and 18 months) were employed to predict final basic ADL group membership, sex was the only significant predictor (see Table 54 & Table 55, p.183). This was also the case (see Table 56, p.184) when initial biomarker scores were used to predict 18 month change in basic ADL group membership (see p.186 for descriptive statistics on change in group membership). In summary, sex is a strong predictor of basic ADL ability over time, whereas 'biomarker' variables were limited in this capacity.

Table 54
Logistic regression statistics: 6 month change in biomarker performance to predict final *basic* activities of daily living group membership

Variable	N	NRS change	Model chi-square (df)	Wald
Covariates				
Sex	126	.14	10.10 (1)**	5.34*
Education	119	.01	.85 (1)	.82
Biomarkers				
Age	116	.03	10.77 (2)**	2.03
Systolic BP	101	.00	7.20 (2)*	.13
Diastolic BP	101	.00	7.07 (2)*	.00
Grip strength	115	.01	7.93 (2)*	.96
BMI	116	.01	7.73 (2)*	.81
Visual acuity#				
Reading Span	116	.01	7.23 (2)*	.35

**p < .01; *p < .05
NRS = Nagelkerke R square; df = Degrees of freedom; Wald = Wald statistic
BP = Blood pressure; BMI = Body mass index
#standardised regression-based change scores were not calculated; see p.171 for more details

6.6.1.2 Instrumental activities of daily living

For instrumental ADL ability, 75 (59.5%) of the 126 participants to

complete the final assessment occasion were categorised as 'independent' and 51 (40.5%) were classified as having 'some dependence'. Of the significant predictors for final instrumental ADL group membership, chronological age explained the most variance (11%), followed by Reading Span (8%) and grip strength (6%). Note, unlike basic ADL, sex was not a significant predictor (see Table 57, p.184). Compared to participants in the 'independent' instrumental ADL group (mean initial age = 76.24 years, SD = 3.45; mean initial Reading Span = 16.27 words correct, SD = 8.68; and mean initial grip strength = 21.53 kg, SD = 8.94), membership to the 'some dependence' group at the final measurement occasion was associated with older age (mean = 78.72 years, SD = 4.64), lesser working memory capacity (mean = 12.57 words correct, SD = 8.93) and weaker grip strength (18.31 kg, SD = 9.10) at the outset of the study. The effect size of these mean differences was moderate (age: $eta^2 = .07$; Reading Span: $eta^2 = .06$; grip strength: $eta^2 = .05$).

Table 55
Logistic regression statistics: 18 month change in biomarker performance to predict final *basic* activities of daily living group membership

Variable	N	NRS change	Model chi-square (df)	Wald
Covariates				
Sex	126	.14	10.10 (1)**	5.34*
Education	119	.01	.85 (1)	.82
Biomarkers				
Age	126	.01	11.11 (2)**	1.04
Systolic BP	102	.01	9.27 (2)*	.29
Diastolic BP	102	.00	8.99 (2)*	.00
Grip strength	118	.00	7.99 (2)*	.05
BMI	119	.03	10.49 (2)**	2.23
Visual acuity#				
Reading Span	119	.00	8.19 (2)*	.01

**p < .01; *p < .05
NRS = Nagelkerke R square; df = Degrees of freedom; Wald = Wald statistic
BP = Blood pressure; BMI = Body mass index
#standardised regression-based change scores were not calculated; see p.171 for more details

Age and grip strength were both able to achieve 85.3% correct group classification for the 'independent' instrumental ADL group, compared to Reading Span's 73.3%. In contrast, Reading Span was most able to classify the 'some

dependence' group correctly (47.1%; age = 43.1% and grip strength = 29.4%). Overall classification accuracy was greatest for age (68.3%) and Reading Span and grip strength were comparable (62.7%).

Table 56
Logistic regression statistics: Initial biomarker performance to predict change over 18 months in *basic* activities of daily living

Variable	N	NRS change	Model chi-square (df)	Wald
Covariates				
Sex	126	.14	10.10 (1)**	5.34*
Education	119	.01	.85 (1)	.82
Biomarkers				
Age	118	.04	10.37 (2)**	2.23
Systolic BP	106	.02	9.14 (2)*	.86
Diastolic BP	106	.01	8.58 (2)*	.33
Grip strength	118	.00	8.12 (2)*	.49
BMI	118	.02	9.28 (2)*	1.24
Visual acuity	117	.00	7.93 (2)*	.02
Reading Span	118	.03	9.69 (2)**	1.43

**p < .01; *p < .05
NRS = Nagelkerke R square; df = Degrees of freedom; Wald = Wald statistic
BP = blood pressure; BMI = body mass index

Table 57
Logistic regression statistics: Initial biomarker performance to predict final *instrumental* activities of daily living group membership

Variable	N	NRS	Model chi-square (df)	Wald
Covariates				
Sex	126	.01	1.16 (1)	1.14
Education	119	.00	.03 (1)	.03
Biomarkers				
Age	126	.11	10.45 (1)**	9.52**
Systolic BP	114	.00	.33 (1)	.33
Diastolic BP	114	.00	.02 (1)	.02
Grip strength	126	.06	6.16 (1)*	5.68*
BMI	126	.00	.95 (1)	.10
Visual acuity	125	.00	.06 (1)	.06
Reading Span	126	.08	8.08 (1)**	7.51**

**p < .01; *p < .05
NRS = Nagelkerke R square; df = Degrees of freedom; Wald = Wald statistic
BP = blood pressure; BMI = body mass index

Sex and education were not significant predictors of final instrumental ADL, and therefore they were not covaried in further regression analyses involving this dependent variable. In terms of employing biomarker change scores to predict final group membership, physiological variables were the only significant predictors, over

both 6 and 18 months (see Tables 58 & 59); grip strength and body mass index (BMI) each explained 5% of variance in instrumental ADL functional ability, respectively.

Table 58
Logistic regression statistics: 6 month change in biomarker performance to predict final *instrumental* activities of daily living group membership

Variable	N	NRS	Model chi-square (df)	Wald
Biomarkers				
Age	123	.00	.24 (1)	.24
Systolic BP	101	.01	.82 (1)	.81
Diastolic BP	101	.00	.17 (1)	.17
Grip strength	115	.05	4.53 (1)*	4.12*
BMI	116	.01	.01 (1)	.01
Visual acuity#				
Reading Span	116	.01	.57 (1)	.56

*p < .05
NRS = Nagelkerke R square; df = Degrees of freedom; Wald = Wald statistic
BP = blood pressure; BMI = body mass index
#standardised regression-based change scores were not calculated; see p.171 for more details

Table 59
Logistic regression statistics: 18 month change in biomarker performance to predict final *instrumental* activities of daily living group membership

Variable	N	NRS	Model chi-square (df)	Wald
Biomarkers				
Age	126	.00	.14 (1)	.14
Systolic BP	102	.01	1.04 (1)	1.02
Diastolic BP	102	.00	.01 (1)	.01
Grip strength	118	.02	1.98 (1)	1.91
BMI	119	.05	4.47 (1)*	4.18*
Visual acuity#				
Reading Span	119	.02	1.53 (1)	1.48

*p < .05
NRS = Nagelkerke R square; df = Degrees of freedom; Wald = Wald statistic;
BP = blood pressure; BMI = body mass index
#standardised regression-based change scores were not calculated; see p.171 for more details

Counter-intuitively, participants whose grip strength decreased slightly over 6 months (mean = -.29, SD = .95) were associated with 'independent' instrumental ADL group membership and 'some dependence' group membership was associated with a slight increase in grip strength over the same length of time (mean = .17, SD = 1.08). The effect size of this mean difference was small to moderate (eta^2 = .04). For BMI, stability over 18 months (mean = .07, SD = .88) was associated with 'independent' final instrumental ADL group membership whereas participants whose BMI dropped over

this period (mean = -.30, SD = 1.13) were more likely to be functionally dependent to some extent. The effect size for this mean difference was also small to moderate (eta^2 = .04).

Grip strength and body mass index were able to successfully classify similar numbers of participants as 'independent' (94.2% and 93.0%, respectively) but BMI classified the 'some dependence' group more accurately (25.0% compared to 17.4% for grip strength). Overall classification accuracy was slightly greater for BMI (65.5%) than grip strength (63.5%).

In contrast to the change scores, initial performance on any of the physiological biomarkers was not able to predict change in instrumental ADL over 18 months (see Table 60, p.187). To carry out logistic regression, participants were grouped as either 'improve or no change' or 'decline' in instrumental functional ability (Table 52, p.180 shows that participants who became more dependent, or 'declined' were in the minority). Initial chronological age and Reading Span performance each explained a significant 6% of variance in change in the extent of independence in instrumental ADL. Older chronological age at the outset of the project (mean = 78.82 years, SD = 4.87) was associated with decline in instrumental functional ability over 18 months (compared to the 'improve/no change' group: mean initial age = 76.80 years, SD = 3.81). Similarly, lesser initial working memory capacity was associated with decline in instrumental ADL (mean = 12.14 words correct, SD = 9.16; 'improve/no change' group: mean = 15.52 words correct, SD = 8.77). The effect size of both of these mean differences was small (age: eta^2 = .04; Reading Span: eta^2 = .02). Classification accuracy was essentially the same for age and Reading Span ('improve/no change' accuracy was 100% for both and overall accuracy was 78.8% and 78.0%, respectively). However, chronological age had better classification accuracy for the 'decline' group

(3.8% compared to 0% for Reading Span).

Table 60
Logistic regression statistics: Initial biomarker performance to predict change over 18 months in *instrumental* activities of daily living

Variable	N	NRS	Model chi-square (df)	Wald
Biomarkers				
Age	118	.06	4.89 (1)*	4.72*
Systolic BP	106	.00	.12 (1)	.12
Diastolic BP	106	.01	.51 (1)	.50
Grip strength	118	.03	2.62 (1)	2.44
BMI	118	.01	.39 (1)	.39
Visual acuity	117	.03	2.47 (1)	2.20
Reading Span	118	.06	5.00 (1)*	4.52*

*p < .05
NRS = Nagelkerke R square; df = Degrees of freedom; Wald = Wald statistic;
BP = blood pressure; BMI = body mass index

6.6.2 Reasoning ability

Reasoning ability in the current study was measured using Raven's Standard Progressive Matrices (Raven's). Consistent with the cross-sectional results and based on its normal distribution, hierarchical linear regression was conducted in order to determine how much variance each biomarker (i.e. independent variable) explained once the effects of sex and education (i.e. covariates) had been accounted for; Chapter 5 demonstrated the influence of covariates on Raven's performance. Tables 61 to 63 (pp.188-189) present regression results where initial, 6 month change, and 18 month change in biomarker scores, respectively, were employed as independent variables. For these Tables, sex and education explained a significant 16% of variance in final Raven's performance. Males, and those with more years of formal education, achieved significantly higher final Raven's scores. Table 64 (p.190) presents regression results where initial biomarker scores were employed to predict change in Raven's performance over 18 months.

After controlling for the effects of sex and education, initial Reading Span performance explained a highly significant 6% of variance in final Raven's performance; this was twice the amount of other significant predictors (initial grip

strength and body mass index each accounted for 3% variance). Greater working memory capacity, stronger grip strength and a higher body mass index at the outset of this project predicted better reasoning ability 18 months later (see Table 61).

Table 61
Linear regression statistics: Initial biomarker performance to predict final Raven's score

Variable	R square change	ANOVA (Overall Model)	Beta
Covariates			
Sex	.04	F(1, 123) = 5.38*	.21*
Education	.12	F(1, 116) = 16.14**	.35**
Biomarkers			
Age	.02	F(3, 114) = 8.03**	-.12
Systolic BP	.00	F(3, 109) = 6.95**	-.03
Diastolic BP	.00	F(3, 109) = 7.10**	.06
Grip strength	.03	F(3, 114) = 8.86**	.28*
BMI	.03	F(3, 114) = 8.90**	.18*
Visual acuity#	.00	F(3, 114) = 7.25**	.03
Reading Span	.06	F(3, 114) = 10.83**	.26**

**p < .01; * p < .05
BP = blood pressure; BMI = body mass index
#this measure was included here because independent regression variables have less stringent distribution assumptions

Concerning 6 and 18 month change in biomarker scores as independent variables, Reading Span was the only significant predictor of final Raven's performance (see Tables 62 and 63, p.189). The rate of change in Reading Span over 6 and 18 months explained 5% and 6% of variance in final Raven's performance, respectively. Participants who improved on Reading Span performance over 6 and 18 months achieved higher Raven's scores on the final measurement occasion. Rate of change in chronological age and physiological measures were not significant predictors of final Raven's performance.

As for predicting 18 month change (i.e. typically decline) in Raven's performance, none of the initial biomarker scores was a significant predictor (see Table 64, p.190). Moreover, sex and education did not show a significant relationship with change in Raven's performance.

Table 62
Linear regression statistics: 6 month change in biomarker performance to predict final Raven's score

Variable	R square change	ANOVA (Overall Model)	Beta
Covariates			
Sex	.04	F(1, 123) = 5.38*	.21*
Education	.12	F(1, 116) = 16.14**	.35**
Biomarkers			
Age	.00	F(3,114) = 7.24**	.02
Systolic BP	.00	F(3, 96) = 6.19**	-.05
Diastolic BP	.02	F(3, 96) = 6.78**	-.12
Grip strength	.01	F(3, 110) = 7.48**	-.10
BMI	.00	F(3, 111) = 7.11**	-.04
Visual acuity#			
Reading Span	.05	F(3,111) = 10.07**	.23**

**p < .01; *p < .05
BP = blood pressure; BMI = body mass index
#standardised regression-based change scores were not calculated; see p.171 for more details

Table 63
Linear regression statistics: 18 month change in biomarker performance to predict final Raven's score

Variable	R square change	ANOVA (Overall Model)	Beta
Covariates			
Sex	.04	F(1, 123) = 5.38*	.21*
Education	.12	F(1, 116) = 16.14**	.35**
Biomarkers			
Age	.00	F(3, 114) = 7.22**	.00
Systolic BP	.02	F(3, 97) = 6.87**	-.13
Diastolic BP	.00	F(3, 97) = 6.18**	-.03
Grip strength	.02	F(3, 113) = 8.35**	.15
BMI	.00	F(3, 114) = 7.44**	.06
Visual acuity#			
Reading Span	.06	F(3, 114) = 10.38**	.23**

**p < .01; * p < .05
BP = blood pressure; BMI = body mass index
#standardised regression-based change scores were not calculated; see p.171 for more details

Figure 20 (p.191) clearly illustrates that Reading Span was the best predictor of final Raven's performance and that the rate of change in, and initial performance on, Reading Span accounted for similar amounts of variance.

6.6.3 Crystallised ability

As discussed in Chapter 3, declines in crystallised ability are believed to be an indicator of proximity to death, or 'terminal drop'. Despite the independent functioning

nature of the current sample, the highest proportion of clinically relevant decline (8.8%) was found for performance on a composite measure of crystallised ability. Therefore, change in crystallised ability over 18 months was included as a dependent variable in the longitudinal analyses. Prior to presenting the regression results, some descriptive statistics and relationships with covariates are presented.

Table 64
Linear regression statistics: Initial biomarker performance to predict 18 month change in Raven's performance

Variable	R square change	ANOVA (Overall Model)	Beta
Covariates			
Sex	.00	F(1, 116) = .33	.05
Education	.00	F(1, 116) = .34	.05
Biomarkers			
Age	.00	F(1, 116) = .10	-.03
Systolic BP	.03	F(1, 104) = 3.03	-.17
Diastolic BP	.01	F(1, 104) = .48	-.07
Grip strength	.00	F(1, 116) = .34	.05
BMI	.00	F(1, 116) = .01	-.01
Visual acuity#	.01	F(1, 116) = 1.05	-.10
Reading Span	.02	F(1, 116) = 2.36	.14

BP = blood pressure; BMI = body mass index
#this measure was included here because independent regression variables have less stringent distribution assumptions

One hundred and nineteen participants completed all crystallised ability tasks (the composite measure included three tasks) on both the initial (mean = 35.11, SD = 3.63; range = 26.00 - 41.33) and final measurement occasions (mean = 35.06, SD = 3.52; range = 25.67 - 42.00). On both occasions, education was positively and significantly correlated with crystallised ability ($r = .44$ and $r = .46$, respectively). Males (mean = 36.20, SD = 3.15) tended to perform better than females (mean = 34.45, SD = 3.57); this difference was significant on the final measurement occasion ($t = 2.64$, $df = 117$, $p < .01$). Therefore, sex and education were covaried in the regression analyses. Linear regression was conducted due to the variable's near-normal distribution.

As shown by Table 65 (p.192), initial Reading Span performance was the only significant predictor of change in crystallised ability over 18 months; explaining 5% of

the variance in performance. Greater working memory capacity at the outset of this project was associated with an absence of decline on crystallised ability 18 months later.

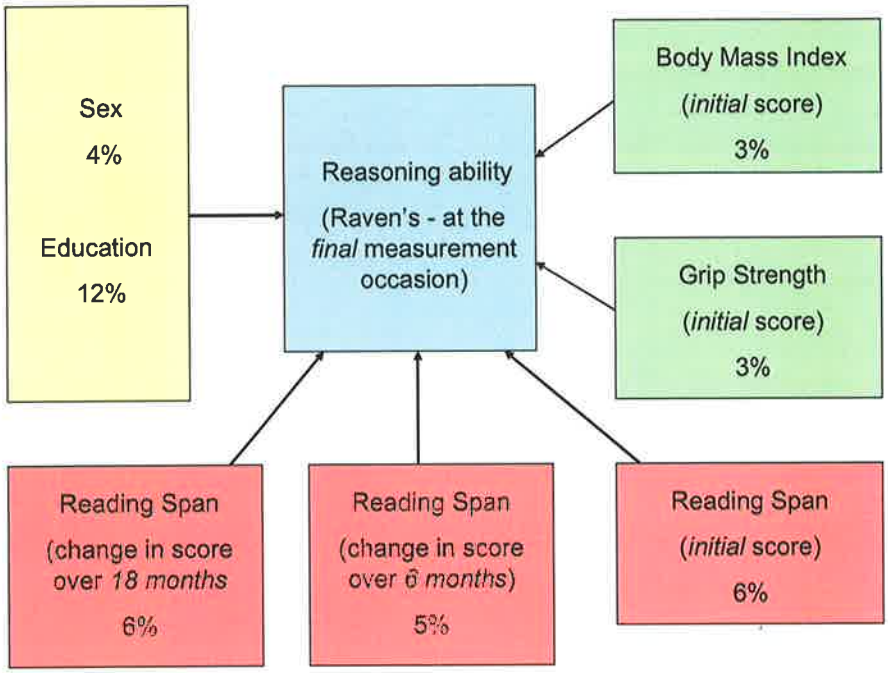


Figure 20 Longitudinal results: Representation of how much variance different biomarker scores explained of reasoning ability on the final measurement occasion

The variance explained by each biomarker is in addition to the variance explained by the covariates (i.e. sex and education).

6.6.4 Memory-based activities of daily living

The memory-based activities of daily living (ADL) questionnaire provided a subjective appraisal of the day-to-day memory problems encountered by participants. This measure was administered on the final measurement occasion only[35] and therefore no analyses could be conducted employing the 18 month change score as a dependent variable. Similarly, descriptive statistics and covariate associations are also presented here before reporting the results of linear regression (the distribution was normal).

In general, participants reported few memory problems relating to daily

[35] This was included due to the ceiling effects found on the composite activities of daily living questionnaire (i.e. limited variance due to high levels of independence). It was anticipated that this memory questionnaire would provide a more sensitive outcome measure of independent functioning in this population.

activities; the mean score was 81.41 (N = 125; SD = 9.90) out of a total possible score of 105, where higher scores denote fewer memory problems. The maximum score achieved was 104, so all participants reported at least one memory problem (minimum score achieved = 47). In terms of relationships with covariates, the memory-based ADL measure was not significantly associated with education (r = .13); and an independent samples t-test showed no significant mean difference (t = .55, df = 123, p > .05) between males (mean = 80.75, SD = 10.73) and females (mean = 81.77, SD = 9.46). Therefore, the influence of sex and education were not controlled for in regression analyses.

Table 65
Linear regression statistics: Initial biomarker performance to predict 18 month change in crystallised ability

Variable	R square change	ANOVA (Overall Model)	Beta
Covariates			
Sex	.03	F(1, 112) = 2.83	.16
Education	.00	F(1, 112) = .39	.06
Biomarkers			
Age	.00	F(1, 112) = .10	-.03
Systolic BP	.01	F(1, 101) = 1.33	-.11
Diastolic BP	.00	F(1, 101) = .44	-.07
Grip strength	.03	F(1, 112) = 3.73	.18
BMI	.00	F(1, 112) = .03	.02
Visual acuity#	.01	F(1, 112) = .99	.09
Reading Span	.05	F(1, 112) = 6.21*	.23*

*p < .05
BP = blood pressure; BMI = body mass index
#this measure was included here because independent regression variables have less stringent distribution assumptions

Of the initial biomarker scores (see Table 66, p.193), Reading Span was the only significant predictor of reported memory problems, accounting for a significant 7% of variance. Participants with greater working memory capacity at the outset of this project reported fewer memory problems 18 months later; an anticipated result.

Unlike initial Reading Span performance, Reading Span change scores (over 6 and 18 months) were not significant predictors of memory-based ADL; this result is

consistent with those reported for basic and instrumental ADL. It should be acknowledged that memory-based ADL correlated positively and significantly with instrumental ADL ($r = .25$, $p < .01$) but not with basic ADL ($r = .08$, $p > .05$). Of the other biomarkers, change in systolic blood pressure over 6 months was the sole predictor of memory-based ADL (see Table 67, p.194). Systolic blood pressure explained a small but significant 5% of variance in reported memory problems; a decrease in systolic blood pressure over 6 months was associated with fewer reported memory problems at the final measurement occasion. This finding is consistent with reports that higher blood pressure can impair cognition (see Chapter 3 for a review of this literature). The rate of change in biomarker variables over 18 months was unable to predict subjective memory performance (see Table 68, p.194).

Table 66
Linear regression statistics: Initial biomarker performance to predict final memory-based activities of daily living

Biomarker Variable	R square	ANOVA (Overall Model)	Beta
Age	.01	F(1, 123) = .90	-.09
Systolic BP	.01	F(1, 111) = .60	.07
Diastolic BP	.00	F(1, 111) = .27	.05
Grip strength	.00	F(1, 123) = .20	.04
BMI	.01	F(1, 123) = 1.21	.10
Visual acuity#	.00	F(1, 123) = .36	.05
Reading Span	.07	F(1, 123) = 9.44**	.27**

**p < .01
BP = blood pressure; BMI = body mass index
#this measure was included here because independent regression variables have less stringent distribution assumptions

6.7 Assessment of selective attrition

The attrition rate of the current study was as anticipated (15.3%; see Chapter 3 for comparisons with other studies); 13 participants dropped out from the initial to the intermediate assessment and 23 dropped out between the initial and final assessment. However, further investigation is necessary to assess to what extent this attrition was selective. That is, less motivated and less physically and mentally healthy individuals

typically discontinue participation (Anstey & Hofer, 2004).

Table 67
Linear regression statistics: 6 month change in biomarker performance to predict final memory-based activities of daily living

Biomarker Variable	R square	ANOVA (Overall Model)	Beta
Age	.00	F(1, 120) = .16	-.04
Systolic BP	.05	F(1, 98) = 5.36*	-.23*
Diastolic BP	.03	F(1, 98) = 3.28	-.18
Grip strength	.01	F(1, 112) = 1.49	-.12
BMI	.00	F(1, 113) = .00	.01
Visual acuity#			
Reading Span	.01	F(1, 113) = .81	.08

*p < .05
BP = blood pressure; BMI = body mass index
#standardised regression-based change scores were not calculated; see p.171 for more details

Table 68
Linear regression statistics: 18 month change in biomarker performance to predict final memory-based activities of daily living

Biomarker Variable	R square	ANOVA (Overall Model)	Beta
Age	.00	F(1, 123) = .08	-.03
Systolic BP	.01	F(1, 99) = 1.23	-.11
Diastolic BP	.00	F(1, 99) = .07	.03
Grip strength	.00	F(1, 115) = .12	-.03
BMI	.01	F(1, 116) = .66	.08
Visual acuity#			
Reading Span	.03	F(1, 116) = 3.31	.17

BP = blood pressure; BMI = body mass index
#standardised regression-based change scores were not calculated; see p.171 for more details

On the surface, it appeared that females were more likely to drop out of the study than males (this was approaching significance for initial to intermediate test sessions) but this was not the case for either attrition period (see Table 69, p.195). Across the duration of this project, reasons provided by participants for not continuing included: loss of interest, medical appointments (e.g. having cataract surgery), and too busy or stressed (e.g. family commitments) at that particular point in time. Two participants died.

A more detailed comparison of individuals who dropped out of, and continued with, the current study is provided in Table 70 (p.196). As can be seen, for both test

periods, highly significant differences were found on the following measures between drop outs and those who continued: Reading Span, Raven's Standard Progressive Matrices (Raven's) and a composite measure of crystallised ability. For each of these tasks, drops outs achieved significantly poorer scores on the initial measurement occasion than those individuals who continued. This result illustrates the presence of selective attrition whereby less mentally healthy individuals stopped participating. From the initial to the final test session, individuals who dropped out were also significantly older at the outset of the study and had made more errors on the dementia screening test. Of note is the fact that none of the physical (e.g. activities of daily living, grip strength) or health measures (e.g. symptom severity, number of diseases) demonstrated significant differences between drop outs and those who continued, for either assessment period. This exemplifies the extent to which the sample was healthy and independent functioning.

Table 69
Initial to final measurement occasions: Selective attrition by sex

Sex	Number of dropouts	
	T1 to T2#	T1 to T3^
Females	12	17
Males	1	6

T1 = Initial test session; T2 = Intermediate test session; T3 = Final test session
#Continuity correction chi-square = 3.20, df = 1, p > .05
^Continuity correction chi-square = .40, df = 1, p > .05

In conclusion, the results of this study need to be viewed in light of the fact that participants who had significantly poorer cognitive functioning (across all domains, working memory, fluid and crystallised ability) were more likely to discontinue participation. This has particular relevance for the absence of significant age correlations with working memory measures (see Chapter 4) and how much variance biomarkers could explain in regression analyses (Chapter 6).

6.8 Results summary and discussion

Regardless of the independent variable score employed (i.e. initial, 6 and 18 month

Table 70

Initial to final measurement occasions: Assessment of selective attrition across dependent and independent variables and covariates

Variable (Initial performance)	Initial to Intermediate Session (T1 to T2)				Initial to Final Session (T1 to T3)			
	Drop out (N=13)		Continue (N=137)		Drop out (N=23)		Continue (N=127)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	79.24	5.38	77.43	4.28	79.60*	5.26	77.22	4.13
Education	12.04	2.79	11.64	4.25	10.88	3.03	11.81	4.28
ADAS-Cog	5.44	2.90	4.91	2.96	6.14*	2.74	4.74	2.94
Depression	4.54	5.66	5.07	4.93	5.51	5.70	4.94	4.85
Symptom severity	73.15	11.77	69.19	12.40	71.74	15.90	69.13	11.63
Number of diseases	2.92	1.55	2.51	1.59	3.00	1.62	2.46	1.56
Basic ADL	6.85	.55	6.81	.62	6.70	.70	6.84	.60
Instrumental ADL	25.77	1.74	25.80	1.89	25.39	2.23	25.87	1.80
Grip strength	16.59	6.33	20.09	9.19	17.34	8.46	20.23	9.07
Reading Span	8.17**	4.51	14.42	8.93	8.95**	6.47	14.77	8.90
Raven's	14.00*	3.98	16.86	4.73	13.91**	4.50	17.10	4.62
Crystallised ability	31.62**	5.00	34.72	3.99	30.94**	4.91	35.10	3.67

ADAS-Cog = Alzheimer's Disease Assessment Scale; ADL = Activities of daily living; Raven's = Raven's Standard Progressive Matrices

*Age (T1 to T3): $t = 2.44$, $df = 148$, $p < .05$

*ADAS-Cog (T1 to T3): $t = 2.13$, $df = 148$, $p < .05$

**Reading Span (T1 to T2): $t = 2.39$, $df = 20$, $p < .01$; **Reading Span (T1 to T3): $t = 3.66$, $df = 36$, $p < .01$

*Raven's (T1 to T2): $t = 2.11$, $df = 148$, $p < .05$; **Raven's (T1 to T3): $t = 3.06$, $df = 148$, $p < .01$

**Crystallised ability (T1 To T2): $t = 2.62$, $df = 144$, $p = .01$; **Crystallised ability (T1 to T3): $t = 3.87$, $df = 27$, $p < .01$

change scores), the longitudinal results consistently found that: male sex and younger age were the best predictors of *in*dependence in basic activities of daily living (ADL); stronger grip strength was the best predictor of functional *in*dependence in instrumental ADL; and male sex, more education and greater working memory capacity were the best predictors of good reasoning ability. These results support the findings of previous studies. For example, female sex has often been associated with functional *de*pendence (Omran, Reed & Ferrara, 1999; Parker et al., 1996). This association has been reasoned to be the result of females living longer with their functional disability, compared to males (Guralnik et al., 1997). Younger age has been associated with maintenance of functional *in*dependence (e.g. Ishizaki et al., 2000; Wang et al., 2002) and stronger grip strength has been found to be predictive of independence in instrumental ADL (Ensrud et al., 1994; Ishizaki et al., 2000; Judge et al., 1996). Grip strength has been shown to correlate significantly with other muscle groups (Rantanen, Era, Kauppinen & Heikkinen, 1994) and therefore Ishizaki et al. (2000) suggested that grip strength is a marker of physical activity. However, given that instrumental activities are considered to be more cognitively demanding than basic ADL, it is probable that grip strength (a sensorimotor variable) reflects more than just physical activity and perhaps, like sensory function (i.e. visual and auditory acuity), may indicate integrity of the central nervous system (Baltes & Lindenberger, 1997). Lastly, the close relationship between reasoning ability and working memory is well-documented (e.g. Engle, Tuholski, Laughlin & Conway, 1999; Kyllonen & Christal, 1990; Prabhakaran, Smith, Desmond, Glover & Gabrieli, 1997; Salthouse, 1993). Prabhakaran et al. (1997) suggested that this association was due to recruitment of common neural systems.

Reading Span, a proposed cognitive biomarker, was successful at predicting outcome measures over time. For example, better initial Reading Span performance was

able to predict final 'independent' group membership of basic and instrumental ADL; and poorer initial Reading Span scores significantly predicted an increase in *de*pendence on instrumental ADL over 18 months. Moreover, and as anticipated, Reading Span was also a significant predictor of more cognitively-dependent outcome measures. All Reading Span scores (initial, 6 and 18 month change scores) were able to predict final performance on Raven's Standard Progressive Matrices (Raven's), with better, or improvements on, Reading Span performance associated with better Raven's scores 18 months later. Initial Reading Span performance was the only significant predictor of change in crystallised ability over 18 months; decline in crystallised ability was associated with poorer initial Reading Span scores. Similarly, initial performance on Reading Span was the only significant predictor of final memory-based ADL.

It has been proposed that chronological age is a poor indicator of functional ability (Bauco et al., 1996; Willis et al., 1992). This suggestion has been partially supported in the current study. Age was unable to significantly predict any of the outcome measures involving greater cognitive demands. In contrast, however, age was a significant predictor of both basic and instrumental ADL. Younger initial age predicted final 'independent' group membership in both types of ADL; older initial age was also a significant predictor of increased *de*pendence in instrumental ADL over 18 months. Where both Reading Span and chronological age were significant predictors of an outcome measure, they explained similar amounts of variance.

The physiological and sensorimotor biomarkers were rarely significant predictors of functional ability outcome measures. Exceptions to this included: stronger initial grip strength was a significant predictor of final 'independent' instrumental ADL group membership (although the amount of variance grip strength explained was less than that of age and Reading Span); a small decrease in grip strength over 6 months was

associated with final 'independent' instrumental ADL group membership (but this finding is counter-intuitive); stability of body mass index (BMI) over 18 months was a significant predictor of final 'independent' instrumental ADL group membership (maintenance of BMI is likely to reflect adequate nutrition and the absence of disease); stronger initial grip strength and higher initial BMI predicted better reasoning ability 18 months later (perhaps these variables reflect the condition of the central nervous system); and a small decrease in systolic blood pressure over 6 months predicted fewer reported memory problems (however, this finding appears to be fairly innocuous).

In summary, Reading Span, a measure of working memory capacity, has demonstrated prognostic utility on different functional outcome measures in a community-dwelling, independent functioning, elderly population. Not surprisingly, this predictive validity was greatest on cognitive-based outcome measures, but it was not restricted to these types of measures. Therefore, Reading Span can be considered a strong candidate for a cognitive biomarker of ageing (Chapter 7 will discuss this issue in more detail). Moreover, in terms of predictive utility, Reading Span was at least *as* useful as chronological age and *more* useful than the physiological measures included here. However, it should be noted that the contextual factors of the current study limited the predictive utility of all biomarkers.

Contextual factors primarily refer to the restricted range of the current sample. For example, each biomarker that was a significant predictor of a functional ability outcome measure, typically accounted for 10% or less variance in performance; this figure was even smaller when change scores were significant predictors. Performance variance across several measures was low at the outset of this study (see Chapter 5 for details) and this was compounded by small rates of change over the 18 months of this longitudinal study; for all measures, generally fewer than 10% of participants showed

clinically relevant change (see p.176 for details). What is more, selective attrition further reduced variability of performance on several outcome measures and, in particular, on one key biomarker, Reading Span. The end result was that biomarkers were applied to circumstances that required differentiation in terms of subtle individual differences and changes in performance. From this perspective, Reading Span did particularly well. The employment of alternate and more sensitive physiological and sensorimotor measures (e.g. auditory acuity and lung function) was discussed in Chapter 5.

Lastly, it should be acknowledged that there were some differences between cross-sectional and longitudinal results. In particular, grip strength was a concurrent but not longitudinal predictor of basic ADL; in contrast, Reading Span was a longitudinal but not concurrent predictor of instrumental ADL; and the number of significant predictors of reasoning ability halved from cross-sectional to longitudinal results. Therefore, conducting longitudinal studies is important in order to verify findings. Chapter 7 concludes this thesis and will discuss the implications from the current study and point to future directions.

CHAPTER 7: CONCLUSION

GENERAL DISCUSSION AND FUTURE DIRECTIONS

7.1 Dissertation objectives, achievements and limitations

The aim of this dissertation was to test the prognostic utility of a cognitive measure (working memory capacity) on functional outcome measures in an elderly population and to ascertain whether this cognitive measure could be considered a biological marker of ageing (biomarker). Measures of specific cognitive domains have been previously employed only infrequently in biomarker research and in studies addressing independent functioning in elderly populations. In order to undertake this evaluation, a longitudinal study was designed incorporating commonly used and reliable biomarkers as well as a measure that was not considered a particularly good indicator of functional ability, chronological age. This design allowed working memory capacity to be compared against biomarkers encompassing a range of efficacies. Indeed, working memory capacity was shown to have comparable prognostic utility to other, more commonly used measures. However, none of the biomarker variables included in the current study was a particularly strong predictor of any of the outcome measures.

The absence of a large number and diverse range of significant predictors for functional outcome measures was likely the result of a combination of sample characteristics and sensitivity of dependent and independent measures. The high-functioning, restricted range of the sample meant that biomarkers had to explain very small amounts of variance in task performance; a challenging task for the most reliable of measures. Moreover, the tendency for Reading Span to predict functional outcome measures was assisted by way of most of the outcome measures involving a strong cognitive component. In contrast, some of the physiological and sensorimotor

biomarkers were at a disadvantage due to the good physical health of the sample (e.g. blood pressure) and due to poor measurement reliability (e.g. visual acuity). Therefore, it is probable that in a different population or by employing different outcome measures, a different pattern of significant biomarkers would occur. This emphasises the issue that determining functional capacity is multifactorial and that establishing a solitary explanatory variable (biomarker) for a range of outcome measures is unlikely.

7.2 Criteria for biomarkers and future research

Reading Span has demonstrated prognostic utility for different functional outcome measures in a community-dwelling, cognitively intact elderly population. However, demonstration of this predictive usefulness is only one of several criteria that would permit Reading Span (or working memory capacity more broadly) to be legitimately classified as a biomarker of ageing. Some of the proposed criteria for biomarkers are discussed below, along with findings from the current study and suggestions for future research.

(1) Biomarkers should be crucial to the maintenance of health (Arking, 1991).

This criterion is based on the idea that there is an increased presence of disease with age (Salthouse et al., 1990), and a biomarker should thus reflect this. Reading Span was not significantly related to most of the health measures included in the current study (see Chapter 5 and Appendix F and N). However, the physical health of the current sample was particularly good and all health information provided was in self-report form. Therefore, future research should include assessment of Reading Span in a more physically impaired sample (e.g. elderly people in nursing homes), in additional to using a more diverse battery of health measures (e.g. blood tests, a general practitioner's rating).

(2) Biomarkers should serve as a predictor of life span (i.e. length of life;

Arking, 1991; Birren & Fisher, 1992) and/or biomarkers should serve as a retrospective marker of ageing (Arking).

These two aspects are important if biomarkers are assumed to reflect functional age more accurately than chronological age. A biomarker should be sensitive to mortality; scores on the biomarker should decline significantly (or increase if inversely related to functioning) with proximity of death (Birren & Fisher, 1992). In the current study, Reading Span demonstrated this type of sensitivity; greater initial working memory capacity was associated with an absence of 'terminal drop' – deterioration in crystallised ability signifying proximity to death (see Chapter 6 for details). However, in general, mortality criteria can often be difficult to meet in the human population due to its lengthy life span in comparison to other species (such as mice). For example, over the 18 month period of the current study, two participants died; this small number meant that survival analysis could not be conducted.

(3) Biomarkers should be highly reproducible and measurable (i.e. display change) over a relatively short period of time compared to the life span of the organism (Arking, 1991; Baker & Sprott, 1988).

This involves aspects such as test-retest reliability, which assesses the stability of the biomarker measurement over a very short period of time, for example, 24 hours. High test-retest reliability is desirable because it indicates that the biomarker measurement is likely to be measuring an actual ageing phenomenon, rather than other confounding processes. For example, the latter may include human errors made during the administration and measurement of the biomarker. Blood pressure is a good example of this, because talking or movement by the participant during the measurement can influence the reading (typically to be higher than usual). Unfortunately, due to the nature and size of the test battery and overall time constraints,

test-retest reliability over this short time period was not possible for biomarker measures. Obviously, future studies would incorporate this measurement but previous research has already established reasonable stability over the short-term (see Chapter 3 for details).

High biomarker reproducibility over longer periods of time is also desirable, that is, over time periods greater than those typically used in test-retest reliability (which range from 24 hours to several weeks). Biomarkers displaying change over these 'short' periods of time would reflect ongoing biological ageing processes and also transient biological changes resulting from acute illnesses, such as the common cold (in addition to measurement error). Although Reading Span showed regression-to-the-mean effects (as did systolic and diastolic blood pressure, see Chapter 6), test-retest correlations over 6 and 18 months were acceptable ($r > .50$). Having said this, Reading Span reliability was lower than for most other biomarkers and therefore investigation of prognostic utility and test-retest reliabilities of other working memory tasks would be valuable. However, it should be kept in mind that biomarker assessment over short time intervals makes the implicit assumption that change over the brief interval is representative of longer intervals and the life course (McClearn, 1997). This is likely to lead to generalisations that are not valid, given the heterogeneity involved in human ageing, as discussed in Chapter 1.

(4) Females have a longer life span than males, and, consequently, greater changes in the biomarker measure should be seen in older males than in older females (Birren & Fisher, 1992).

This was not directly assessed in the current study but does emphasise the importance of replicating these results in a sample with more males, or indeed, in individual samples of males and females. The current findings have suggested that

different relationships exist between certain variables for males and females (e.g. nutrition, see Appendix O).

(5) The change in the biomarker should be exacerbated by the presence of age-associated conditions, such as Alzheimer's Disease and cardiovascular disease (Birren & Fisher, 1992).

Individuals with dementia were screened at the outset of the current study and participants with health problems like cardiovascular disease were too few to make a comparison group. However, the sensitivity of Reading Span in these populations would be of great interest and a likely topic of future studies; a study involving Alzheimer's Disease patients was anticipated at the commencement of the current project but time constraints and uncertainty about the number of Alzheimer's Disease patients available to participate prevented this from going ahead (as discussed in Chapter 2).

(6) Biomarkers can show the influence of training and interventions, such as dietary restriction (Piantanelli et al., 1992).

According to Ingram (1991), the purpose of identifying biomarkers is to develop tasks associated with ageing that could be employed to assess interventions in laboratory animals and humans. The more common interventions applied thus far include: caloric restriction, dietary modifications and exercise regimes (Baker & Sprott, 1988; Drachman, 1997). However, the mechanism by which these interventions benefit ageing is unclear. Nonetheless, intervention studies are another likely avenue for future studies involving Reading Span, and other working memory tasks.

(7) Biomarkers should be physically non-lethal, and preferably, not be physically invasive or psychologically traumatic (Arking, 1991; Baker & Sprott, 1988).

Ideally, obtaining biomarker information from participants should be relatively quick and not particularly stressful (physically or mentally) in order to maximize the

biomarker's utility. This allows for frail and fatigued individuals to be assessed and also assists in the collection of reliable results (McClearn, 1997). In the current study, participants did at times find Reading Span (and other tasks of working memory) stressful and frustrating. This is a difficult issue to overcome because, by definition, working memory tasks are complex. However, the current sample was very achievement motivated as well as being well-educated and cognitively intact. This may have exacerbated any frustration or stress. For example, other populations may be less aware of the quality of their performance (i.e. when they are not performing well). This issue also highlights the importance of assessing a range of working memory tasks to assist in finding a balance between measurement sensitivity and maintaining the psychological commitment of participants.

Other areas of future research include employing alternate longitudinal and change score methodologies in order to corroborate the current findings and to incorporate more measures of working memory and cognition. This includes the assessment of more specific measures of working memory that better operationalise the components of Baddeley and Hitch's (1974) tripartite model (and expanding to assessment of the episodic buffer) in addition to investigating different cognitive domains, such as speed of processing and other types of memory. In conclusion, working memory capacity, as measured by Reading Span, is a valuable addition to the assessment of functional ability in an elderly population and highlights the importance of cognition in this context. However, further investigations are required before Reading Span can be classified as a *biomarker* of ageing.

REFERENCES

Agüero-Torres, H., Hillerås, P. K., & Winblad, B. (2001). Disability in activities of daily living among the elderly. *Current Opinion in Psychiatry, 14,* 355-359.

Albert, M. S., Jones, K., Savage, C. R., Berkman, L., Seeman, T., Blazer, D. et al. (1995). Predictors of cognitive change in older persons: MacArthur studies of successful aging. *Psychology and Aging, 10,* 578-589.

Alder, A. G., Adam, J., & Arenberg, D. (1990). Individual-differences assessment of the relationship between change in and initial level of adult cognitive functioning. *Psychology and Aging, 5,* 560-568.

Anderson, N. D. & Craik, F. I. M. (2000). Memory in the aging brain. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford Handbook of Memory*, (pp.411-425). Oxford: Oxford University Press.

Anderson, R. T., James, M. K., Miller, M. E., Worley, A. S. & Longino Jr., C.F. (1998). The timing of change: Patterns in transitions in functional status among elderly persons. *Journal of Gerontology. Series B: Psychological Sciences and Social Sciences, 53B,* S17-S27.

Andrews, K. (2003). *Opening Address. Ageless Workforce Symposium, Exhibition Centre, Sydney*. Retrieved July 13, 2004, from http://www.health.gov.au/mediarel/yr2003/ka/kasp03176.htm

Anstey, K. (1999). Sensorimotor variables and forced expiratory volume as correlates of speed, accuracy, and variability in reaction time performance in late adulthood. *Aging, Neuropsychology, and Cognition, 6,* 84-95.

Anstey, K. & Christensen, H. (2000). Education, activity, health, blood pressure and apolipoprotein as predictors of cognitive change in old age: A review. *Gerontology, 46,* 163-177.

Anstey, K. J., & Hofer, S. M. (2004). Longitudinal designs, methods and analysis in psychiatric research. *Australian and New Zealand Journal of Psychiatry, 38,* 93-104.

Anstey, K. J., Hofer, S. M., & Luszcz, M. A. (2003). Cross-sectional and longitudinal patterns of dedifferentiation in late-life cognitive and sensory function: The effects of age, ability, attrition, and occasion of measurement. *Journal of Experimental Psychology: General, 132,* 470-487.

Anstey, K., Lord, S., & Smith, G. A. (1996). Measuring human functional age: A review of empirical findings. *Experimental Aging Research, 22,* 245-266.

Anstey, K. J., & Smith, G. A. (1999). Interrelationships among biological markers of aging, health, activity, acculturation, and cognitive performance in late adulthood. *Psychology and Aging, 14,* 605-618.

Anstey, K., Smith, G. A., & Lord, S. (1997). Test-retest reliability of a battery of sensory, motor and physiological measures of aging. *Perceptual and Motor Skills, 84,* 831-834.

Anstey, K., Stankov, L., & Lord, S. (1993). Primary aging, secondary aging and intelligence. *Psychology and Aging, 8,* 562-570.

Arking, R. (1991). *Biology of Aging. Observations and principles.* Englewood Cliffs, NJ: Prentice Hall.

Arnett, P. A., Higginson, C. I., Voss, W. D., Bender, W. I., Wurst, J. M., & Tippin, J. M. (1999). Depression in multiple sclerosis: Relationship to working memory capacity. *Neuropsychology, 13,* 546-556.

Atchley, R. C. (1989). The continuity theory of normal aging. *The Gerontologist, 29,* 183-190.

Atkinson, H. H., Cesari, M., Kritchevsky, S. B., Pennix, B. W. J. H., Fried, L. P., Guralnik, J. M., et al. (2005). Predictors of combined cognitive and physical decline. *Journal of the American Geriatrics Society, 53,* 1197-1202.

Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory. A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation. Volume 2,* (pp.89-105). New York: Academic Press.

Australian Bureau of Statistics. (2004). *Population by age and sex, South Australia. Summary. Sex Ratio.* Retrieved January 10, 2004, from www.abs.gov.au/AUSSTATS/abs@.nsf/ProductsbyReleaseDate/142548D794D95D1CA25702F0071F6FE?OpenDocument

Australian Bureau of Statistics. (2005). *Population by age and sex, Australian Sates and Territories. Summary of findings. Older people.* Retrieved October 6, 2006, from http://www.abs.gov.au/Ausstats/abs@.nsf/0e5fa1cc95cd093c4a256811000752b/b52c3903d894336dca2568a9001393c1!OpenDocument

Australian Government (2004). *Descriptions of designated National Research Priorities and associated Priority Goals for 2004-2006.* Retrieved November 30, 2004, from http://www.arc.gov.au/pdf/2004_designated_national_research_priorities_&_associate.pdf

Australian Government Department of Health and Ageing (2006). *Alcohol and older people. Australian alcohol guidelines – fact sheet.* Retrieved October 23, 2006, from http://www.alcohol.gov.au/internet/alcohol/publishing.nsf/Content/guidelines

Babcock, R. L. & Salthouse, T. A. (1990). Effects of increased processing demands on age differences in working memory. *Psychology and Aging, 5,* 421-428.

Bäckman, L. & Hill, R. D. (1996). Cognitive performance and everyday functioning: Patterns in normal aging and dementia. In R. T. Woods (Ed.), *Handbook of the Clinical Psychology of Ageing,* (pp.73-92). Chichester: John Wiley & Sons.

Baddeley, A. (1986). *Working Memory.* Oxford: Clarendon Press.

Baddeley, A. D. (1990a). The development of the concept of working memory. Implications and contributions of neuropsychology. In G. Vallar & T. Shallice (Eds.), *Neuropsychological Impairments of Short-term Memory,* (pp.54-73). Cambridge: Cambridge University Press.

Baddeley, A. (1990b). *Human Memory. Theory and Practice.* Hove: Lawrence Erlbaum.

Baddeley, A. D. (1992). Working memory: The interface between memory and cognition. *Journal of Cognitive Neuroscience, 4,* 281-288.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A,* 5-28.

Baddeley, A. (1998). Recent developments in working memory. *Current Opinion in Neurobiology, 8,* 234-238.

Baddeley, A. (2000a). The new episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4,* 417-423.

Baddeley, A. (2000b). Short-term and working memory. In E. Tulving & F. I. M Craik (Eds.), *The Oxford Handbook of Memory,* (pp. 77-92). Oxford: Oxford University Press.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist, 7,* 85-97.

Baddeley, A., Emslie, H. & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language-Processing Test (SCOLP) Manual.* Suffolk: Thames Valley Test Company.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot-the-Word test: A robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology, 32,* 55-65.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation. Advances in Research and Theory. Volume 8,* (pp.47-90). New York: Academic Press.

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology, 8,* 485-493.

Baddeley, A., & Hitch, G. J. (2000). Development of working memory: Should the Pascual-Leone and the Baddeley and Hitch models be merged? *Journal of Experimental Child Psychology, 77,* 128-137.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory. Mechanisms of Active Maintenance and Executive Control,* (pp.28-61). Cambridge: Cambridge University Press.

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

Baker, G. T. & Sprott, R. L. (1988). Biomarkers of aging. *Experimental Gerontology, 23,* 223-239.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23,* 611-626.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging, 12,* 12-21.

Bashkireva, A. S. & Khavinson, V. K. (2001). Influence of biological age on professional efficiency: Communication I. Biological age and mental efficiency. *Human Physiology, 27,* 353-359.

Bassey, E. J. (2000). The benefits of exercise for the health of older people. *Reviews in Clinical Gerontology, 10,* 17-31.

Bauco, C., Golosio, F., Cinti, A. M., Borriello, C., Raganato, P., Cicconetti, P. et al. (1996). Functional status and well-being of centenarians. *Archives of Gerontology and Geriatrics, suppl. 5,* 363-366.

Bédard, M., Molloy, D. W., Bell, R. & Lever, J. A. (2000). Dementia. Determinants and detection of low body mass index in community-dwelling adults with Alzheimer's Disease. *International Psychogeriatrics, 12,* 87-98.

Belleville, S., Perez, I., & Malenfant, D. (1996). Examination of the working memory components in normal aging and in dementia of the Alzheimer type. *Neuropsychologia, 34,* 195-207.

Benjamin, C. L., Garman, G. R. & Funston, J. H. (1997). *Human Biology.* New York: McGraw-Hill.

Benjamin, H. (1947). Biologic versus chronologic age. *Journal of Gerontology, 2,* 217-227.

Benton, D., Fordy, J., & Haller, J. (1995). The impact of long-term vitamin supplementation on cognitive functioning. *Psychopharmacology, 117,* 298-305.

Binder, E. F., Storandt, M. & Birge, S. J. (1999). The relation between psychometric test performance and physical performance in older adults. *Journal of Gerontology. Series A: Biological Sciences and Medical Sciences, 54A,* M428-M432.

Birren, J. E. (1959). Principles of research on aging. In, J. E. Birren (Ed.), *Handbook of Aging and the Individual: Psychological and Biological Aspects,* (pp.3-42).Chicago: University of Chicago Press.

Birren, J. E., & Fisher, L. M. (1992). Aging and slowing of behaviour: Consequences for cognition and survival. In, T. Sonderegger (Ed.), *Nebraska Symposium on Motivation: 1991. Psychology and Aging,* (pp.1-37). Lincoln, NE: University of Nebraska Press.

Birren, J. E., & Speith, W. (1962). Age, response speed, and cardiovascular functions. *Journal of Gerontology, 17,* 390-391.

Black, S. A. & Rush, R. D. (2002). Cognitive and functional decline in adults aged 75 and older. *Journal of the American Geriatrics Society, 50,* 1978-1986.

Blansjaar, B. A., Thomassen, R. & Van Schaik, H. W. (2000). Prevalence of dementia in centenarians. *International Journal of Geriatric Psychiatry, 15,* 219-225.

Bond, L. (1979). On the base-free measure of change proposed by Tucker, Damarin and Messick. *Psychometrika, 44,* 351-355.

Borkan, G. A. & Norris, A. H. (1980). Assessment of biological age using a profile of physical parameters. *Journal of Gerontology, 35,* 177-184.

Bortz, W. M. (1990). The trajectory of dying. Functional status in the last years of life. *Journal of the American Geriatrics Society, 38,* 146-150.

Boult, C., Kane, R. L., Louis, T. A., Boult, L., & McCaffrey, D. (1994). Chronic conditions that lead to functional limitation in the elderly. *Journal of Gerontology. Series A: Biological Sciences and Medical Sciences, 49,* M28-M36.

Boyd, A. & Dawson, D. R. (2000). The relationship between perceptual impairment and self-care status in a sample of elderly persons. *Physical and Occupational Therapy in Geriatrics, 17,* 1-16.

Branch, L. G., Horowitz, A., & Carr, C. (1989). The implications for everyday life of incident self-reported visual decline among people over age 65 living in the community. *The Gerontologist, 29,* 359-365.

Brink, T. L., Yesavage, J. A., Lum, O., Heersema, P.H., Adey, M., & Rose, T. L. (1982). Screening tests for geriatric depression. *Clinical Gerontologist, 1,* 37-43.

Brink, T. L., Curran, P., Dorr, M. L., Janson, E., McNulty, V. & Messina, M. (1985). Geriatric Depression Scale reliability: Order, examiner and reminiscence effects. *Clinical Gerontologist, 3,* 57-60.

Bruce, M. L., Seeman, T. E., Merrill, S. S., & Blazer, D. G. (1994). The impact of depressive symptomatology on physical disability: MacArthur Studies of Successful Aging. *American Journal of Public Health, 84,* 1796-1799.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101,* 147-158.

Bucks, R. S., Ashworth, D. L., Wilcock, G. K. & Siegfried, K. (1996). Assessment of activities of daily living in dementia: Development of the Bristol Activities of Daily Living Scale. *Age and Ageing, 25,* 113-120.

Byrd, O. E. (1941). *The Byrd Health Attitude Scale.* Standford: Standford University Press.

Camicioli, R., Lea, E., Nutt, J. G., Sexton, G. & Oken, B. S. (2001). Methylphenidate increases the motor effects of L-Dopa in Parkinson's Disease: A pilot study. *Clinical Neuropharmacology, 24,* 208-213.

Cantor, J. & Engle, R. W. (1993). Working memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1101-1114.

Cantor, J., Engle, R. W. & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence, 15,* 229-246.

Carlson, M. C., Fried, L. P., Xue, Q., Bandeen-Roche, K., Zeger, S. L., & Brandt, J. (1999). Association between executive attention and physical functional performance in community-dwelling older women. *Journal of Gerontology. Series B: Psychological Sciences and Social Sciences, 54B,* S262-S270.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor Analytic Studies.* Cambridge: Cambridge University Press.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin, 38,* 592.

Cattell, R. B., & Cattell, A. K. S. (1959). *The Cattell Culture Fair Test.* Illinois: IPAT.

Channon, S., & Baker, J. (1994). Reasoning strategies in depression: Effects of depressed mood on a syllogism task. *Personality and Individual Differences, 17,* 707-711.

Chen, P., Yu, E. S. H., Zhang, M., Liu, W. T., Hill, R., & Katzman, R. (1995). ADL dependence and medical conditions in Chinese older persons: A population based survey in Shanghai, China. *Journal of the American Geriatrics Society, 43,* 378-383.

Chiappe, P., Hasher, L. & Siegel, L. S. (2000). Working memory, inhibitory control, and reading disability. *Memory and Cognition, 28,* 8-17.

Chodzko-Zajko, W. J., Schuler, P., Solomon, J., Heinl, B., & Ellis, N. R. (1992). The influence of physical fitness on automatic and effortful memory changes in aging. *International Journal of Aging and Human Development, 35,* 265-285.

Christensen, H., Korten, A., Jorm, A. F., Henderson, A. S., Scott, R., & Mackinnon, A. J. (1996). Activity levels and cognitive functioning in an elderly community sample. *Age and Ageing, 25,* 72-80.

Christensen, H., & Mackinnon, A. (1993). The association between mental, social, and physical activity and cognitive performance in young and old subjects. *Age and Ageing, 22,* 175-178.

Christensen, H., Mackinnon, A. J., Korten, A. E., Jorm, A. F., Henderson, A. S., Jacomb, P., et al. (1999). An analysis of diversity in the cognitive performance of elderly community dwellers: Individual differences in change scores as a function of age. *Psychology and Aging, 14,* 365-379.

Clark, J. W. (1960). The aging dimension: A factorial analysis of individual differences with age on psychological and physiological measurements. *Journal of Gerontology, 15,* 183-187.

Clark, W. R. (1999). *A means to an end: The biological basis of aging and death.* Oxford: Oxford University Press.

Clarkson-Smith, L., & Hartley, A. A. (1989). Relationships between physical exercise and cognitive abilities in older adults. *Psychology and Aging, 4,* 183-189.

Clarkson-Smith, L., & Hartley, A. A. (1990). Structural equation models of relationships between exercise and cognitive abilities. *Psychology and Aging, 5,* 437-446.

Collette, F., Van der Linden, M., Bechet, S., Belleville, S. & Salmon, E. (1998). Working memory deficits in Alzheimer's Disease. *Brain and Cognition, 37,* 147-149.

Collins, L. M. (1996). Measurement of change in research in aging: Old and new issues from an individual growth perspective. In J. E. Birren & Schaie, K. W. (Eds.), *Handbook of the Psychology of Aging,* (pp.38-56). San Diego: Academic Press.

Compton, D. M., Bachman, L. D., Brand, D. & Avet, T. L. (2000). Age-associated changes in cognitive functioning in highly educated adults: Emerging myths and realities. *International Journal of Geriatric Psychiatry, 15,* 75-85.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30,* 163-183.

Cooper, J. A., Sagar, H. J., Jordan, N., Harvey, N. S. & Sullivan, E. V. (1991). Cognitive impairment in early, untreated Parkinson's Disease and its relationship to motor disability. *Brain, 114,* 2095-2122.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia, 6,* 255-265.

Craik, F. I. M. (1977). Age differences in human memory. In J .E. Birren & K. W. Schaie (Eds.), *Handbook of the Psychology of Aging,* (pp.384-420). New York: Van Nostrand.

Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science, 3,* 155-158.

Craik, F. I. M. (2000). Age-related changes in human memory. In D. C. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer,* (pp.75-92). Philadelphia: Psychology Press.

Craik, F. I. M., Morris, R. G. & Gick, M. L. (1990). Adult age differences in working memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological Impairments of Short-term Memory,* (pp.247-267). Cambridge: Cambridge University Press.

Cronbach, L. J., & Furby, L. (1970). How should we measure 'change' – or should we? *Psychological Bulletin, 74,* 68-80.

Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour, 19,* 450-466.

Dawson, D., Winocur, G. & Moscovitch, M. (1999). The psychosocial environment and cognitive rehabilitation in the elderly. In D. T. Stuss, G. Winocur & I. H. Robertson (Eds.), *Cognitive Neurorehabilitation,* (pp.94-108). Cambridge: Cambridge University Press.

De Lemos, M. M. (1995). *Standard Progressive Matrices: Australian manual.* Melbourne: ACER.

Diggle, P. J., Liang, K., & Zeger, S. L. (1994). *Analysis of longitudinal data.* Oxford: Clarendon.

Dikman, S. S., Heaton, R. K., Grant, I., & Temkin, N. R. (1999). Test-retest reliability and practice effects of Expanded Halstead-Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society, 5,* 346-356.

Dirken, J. M. (1972a). Background and aim of the study. In J. M. Dirken (Ed.), *Functional Age of Industrial Workers,* (pp.1-9). Netherlands: Wolters-Noordhoff.

Dirken, J. M. (1972b). The Design of the Study. In J. M. Dirken, (Ed.), *Functional Age of Industrial Workers,* (pp.10-16). Netherlands: Wolters-Noordhoff.

Dobbs, A. R. & Rule, B. G. (1989). Adult age differences in working memory. *Psychology and Aging, 4,* 500-503.

Drachman, D. A. (1997). Aging and the brain: A new frontier. *Annals of Neurology, 42,* 819-828.

Elsayed, M., Ismail, A. H., & Young, R. J. (1980). Intellectual differences of adult men related to age and physical fitness before and after an exercise program. *Journal of Gerontology, 35,* 383-387.

Engle, R. W. (2001). What is working memory capacity? In H. L. Roediger, J. S. Nairne, I. Neath & A. M. Suprenant (Eds.), *The Nature of Remembering,* (pp. 297-314). Washington D.C: American Psychological Association.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128,* 309-331.

Ensrud, K. E., Nevitt, M. C., Yunis, C., Cauley, J. A., Seeley, D. G., Fox, K. M., et al. (1994). Correlates of impaired function in older women. *Journal of the American Geriatrics Society, 42,* 481-489.

Femia, E. E., Zarit, S. H., & Johansson, B. (1997). Predicting change in activities of daily living: A longitudinal study of the oldest old in Sweden. *Journal of Gerontology: Series B: Psychological Sciences and Social Sciences, 52B,* P294-302.

Fillenbaum, G. G. (1985). Screening the elderly. A brief instrumental activities of daily living measure. *Journal of the American Geriatrics Society, 33,* 698-706.

Fillenbaum, G. G. (1979). Social contest and health assessments of health among the elderly. *Journal of Health and Social Behavior, 20,* 45-51.

Finkel, D., Whitfield, K. & McGue, M. (1995). Genetic and environmental influences on functional age: A twin study. *Journal of Gerontology: Series B: Psychological Sciences and Social Sciences, 50B,* P104-P113.

Fioravanti, M., Ferrario, E., Massaia, M., Cappa, G., Rivolta, G., Grossi, E., et al. (1997). Low folate levels in the cognitive decline of elderly patients and the efficacy of folate as a treatment for improving memory deficits. *Archives of Gerontology and Geriatrics, 26,* 1-13.

Fischer, M. H. (2001). Probing spatial working memory with the Corsi Blocks Task. *Brain and Cognition, 45,* 143-154.

Fisk, J. E. & Warr, P. (1996). Age and working memory: The role of perceptual speed, the central executive, and the phonological loop. *Psychology and Aging, 11,* 316-323.

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

Fone, S., & Lundgren-Linquist, B. (2003). Health status and functional capacity in a group of successfully ageing 65-85 year olds. *Disability and Rehabilitation, 25,* 1044-1051.

Fozard, J. L., Metter, E. J., & Brant, L. J. (1990). Next steps in describing aging and disease in longitudinal studies. *Journal of Gerontology. Series B: Psychological Sciences and Social Sciences, 45,* P116-P127.

Frank, L. K. (1946). Gerontology. *Journal of Gerontology, 1,* 1-11.

Frerichs, R. J., & Tuokko, H. A. (2006). Reliable change scores and their relation to perceived change in memory: Implications for the diagnosis of mild cognitive impairment. *Archives of Clinical Neuropsychology, 21,* 109-115.

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods, Instruments and Computers, 37,* 581-590.

Fry, A. F. & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology, 54,* 1-34.

Furner, S. E., Rudberg, M. A., & Cassel, C. K. (1995). Medical conditions differentially affect the development of IADL disability: Implications for medical care and research. *The Gerontologist, 35,* 444-450.

Furukawa, T., Inoue, M., Kajiya, F., Inada, H., Takasugi, S., Fukui, S., et al. (1975). Assessment of biological age by multiple regression. *Journal of Gerontology, 30,* 422-434.

Galanos, A. N., Pieper, C. F., Cornoni-Huntlwy, J. C., Bales, C. W. & Fillenbaum, G. G. (1994). Nutrition and function: Is there a relationship between body mass index and functional capabilities of community-dwelling elderly? *Journal of the American Geriatrics Society, 42,* 368-373.

Gallassi, R., Morreale, A. & Pagni, P. (2001). The relationship between depression and cognition. *Archives of Gerontology and Geriatrics, suppl. 7,* 163-171.

Gardner, A., & Poelhman, E. (1997). Predictors of the age-related rise in blood pressure in men and women. *Journal of Gerontology. Series A: Biological Sciences and Medical Sciences, 50A,* M1-M6.

Gareri, P., Ruotolo, G., Curcio, M., Giancotti, S., Talarico, F., & Galasso, D. (2001). Prevalence of depression in medically hospitalized elderly patients. *Archives of Gerontology and Geriatrics, suppl.1,* 183-189.

Gathercole, S. E & Baddeley, A. D. (1993). *Working Memory and Language.* East Sussex: Lawrence Erlbaum.

Geva, A. (2002). Working memory deficits in depression. *Dissertation Abstracts International. Section B: The Sciences and Engineering, 63,* 1059.

Gill, T. M., Williams, C. S., Richardson, E. D. & Tinetti, M. E. (1996). Impairments in physical performance and cognitive status as predisposing factors for functional dependence among nondisabled older persons. *Journal of Gerontology. Series A: Biological Sciences and Medical Sciences, 51A,* M283-M288.

Gittings, N. S. & Fozard, J. L. (1986). Age-related changes in visual acuity. *Journal of Experimental Gerontology, 21,* 423-433.

Goldberg, A. P., Dengel, D. R. & Hagberg, J. M. (1995). Exercise physiology and aging. In E. L. Schneider & J. W. Rowe (Eds.), *Handbook of the Biology of Aging,* (pp.331-354). San Diego: Academic Press.

Goodwin, J. S., Goodwin, J. M., & Garry, P. J. (1983). Association between nutritional status and cognitive functioning in a healthy elderly population. *Journal of the American Medical Association, 249,* 2917-2921.

Gregoire, J., & Van der Linden, M. (1997). Effect of age on forward and backward digit spans. *Aging, Neuropsychology, and Cognition, 4,* 140-149.

Greiner, P. A., Snowdon, D. A. & Schmitt, F. A. (1996). The loss of independence in activities of daily living: The role of low normal cognitive function in elderly nuns. *American Journal of Public Health, 86,* 62-66.

Griep, M. I., Mets, T. F., Collys, K., Ponjaert-Kristoffersen, I., & Massart, D. L. (2000). Risk of malnutrition in retirement homes elderly persons measured by the 'Mini-Nutritional Assessment'. *Journal of Gerontology. Series A: Biological Sciences and Medical Sciences, 55A,* M57-M63.

Grigsby, J., Kaye, K., Baxter, J., Shetterly, S. & Hamman, R. F. (1998). Executive cognitive abilities and functional status among community-dwelling older persons in the San Luis Valley Health and Aging Study. *Journal of the American Geriatrics Society, 46,* 590-596.

Groeger, J. A., Field, D., & Hammond, S. M. (1999). Measuring memory span. *International Journal of Psychology, 34,* 359-363.

Guccione, A. A., Felson, D. T., Anderson, J. J., Anthony, J. M., Zhang, Y., Wilson, P. W. F., et al. (1994). The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. *American Journal of Public Health, 84,* 351-358.

Gupta, J. K., Srivastava, A. B. L., & Sharma, K. K. (1988). Estimation of true change with three-wave data. *Journal of Experimental Education, 57,* 59-68.

Guralnik, J. M., & Kaplan, G. A. (1989). Predictors of healthy aging: Prospective evidence from the Alameda County Study. *American Journal of Public Health, 79,* 703-708.

Guralnik, J. M., LaCroix, A. Z., Abbott, R. D., Berkman, L., Satterfield, S., Evans, D. A. et al. (1993). Maintaining mobility in late life. *American Journal of Epidemiology, 137,* 845-857.

Guralnik, J. M., Leveille, S. G., Hirsch, R., Ferrucci, L., & Fried, L. P. (1997). The impact of disability in older women. *Journal of the American Medical Women's Association, 52,* 113-120.

Haavisto, M. & Lehto, J. E. (2004). Fluid/spatial and crystallized intelligence in relation to domain-specific working memory: A latent variable approach. *Learning and Individual Differences, 15,* 1-21.

Harman, D. (1998). Alzheimer's Disease: A hypothesis on pathogenesis. In D. Harman, R. Holliday & M. Meydani (Eds.), *Towards prolongation of the healthy life span: Practical approaches to intervention,* (p.515). New York: New York Academy of Sciences.

Harris, T., Kovar, M. G., Suzman, R., Kleinman J. C. & Feldman, J. J. (1989). Longitudinal study of physical ability in the oldest old. *American Journal of Public Health, 79,* 698-702.

Hassing, L., Wahlin, A., Winblad, B. & Bäckman, L. (1999). Further evidence on the effects of vitamin B12 and folate levels on episodic memory functioning: A population-based study on healthy very old adults. *Biological Psychiatry, 45,* 1472-1480.

Hatsukam, D. K., Benowitz, N. L., Rennard, S. I., Oncken, C., & Hecht, S. S. (2006). Biomarkers to assess the utility of potential reduced exposure tobacco products. *Nicotine and Tobacco Research, 8,* 600-622.

Hébert, R., Brayne, C. & Spiegelhalter, D. (1999). Factors associated with functional decline and improvement in a very elderly community-dwelling population. *American Journal of Epidemiology, 150,* 501-510.

Hendrie, H. C., Gao, S., Hall, K. S., Hui, S. L., & Unverzagt, F. W. (1996). The relationship between alcohol consumption, cognitive performance, and daily functioning in an urban sample of older Black Americans. *Journal of the American Geriatrics Society, 44,* 1158-1165.

Hermann, B. P., Seidenberg, M., Schoenfeld, J., Peterson, J., Leveroni, C., & Wyler, A. R. (1996). Empirical techniques for determining the reliability, magnitude, and pattern of neuropsychological change following epilepsy surgery. *Epilepsia, 37,* 942-950.

Heron, A. (1987). The concept of functional age. In L. Levi (Ed.), *Society, Stress and Disease. Old Age* (pp.31-37). Oxford: Oxford University Press.

Heron, A., & Chown, S. (1967). *Age and Function.* London: J. & A. Churchill.

Hertzog, C., Dixon, R. A., & Hultsch, D. F. (1992). Intraindividual change in text recall of the elderly. *Brain and Language, 42,* 248-269.

Hester, R. L., Kinsella, G. L., & Ong, B. (2004). Effect of age on forward and backward span tasks. *Journal of the International Neuropsychological Society, 10,* 475-481.

Hochschild, R. (1990). Can an index of aging be constructed for evaluating treatments to retard aging rates? A 2,462-person study. *Journal of Gerontology. Series A: Biological Sciences and Medical Sciences, 45,* B187-B214.

Holford, P. (2003). *Optimum Nutrition for the Mind.* London: Judy Piatkus.

Hoppe, C., Müller, U., Werheid, K., Thöne, A. & von Cramon, D. Y. (2000). Digit ordering test: Clinical, psychometric, and experimental evaluation of a verbal working memory test. *The Clinical Neuropsychologist, 14,* 38-55.

Hultsch, D. F., Hertzog, C., Dixon, R. A. & Small, B. J. (1998). *Memory change in the aged.* Cambridge: Cambridge University Press.

Hultsch, D. F., Hertzog, C., Small, B. J., McDonald-Miszczak, L., & Dixon, R. A. (1992). Short-term longitudinal change in cognitive performance in later life. *Psychology and Aging, 7,* 571-584.

Ichikawa, S. (1981). Rated ease of memorization of dot-in-matrix patterns: Multiple regression analysis by physical variables which describe configuration of dots. *Japanese Psychological Research, 23,* 69-78.

Ichikawa, S. (1983). Verbal memory span, visual memory span, and their correlations with cognitive tasks. *Japanese Psychological Research, 25,* 173-180.

Ihl, R., Grass-Kapanke, B., Jänner, M. & Weyer, G. (1999). Neuropsychometric tests in cross-sectional and longitudinal studies – A regression analysis of ADAS-Cog, SKT and MMSE. *Pharmacopsychiatry, 32,* 248-254.

Ingram, D. K. (1991). Is aging measurable? In, F. C. Ludwig (Ed.), *Life Span Extension. Consequences and Open Questions,* (pp.18-42). New York: Springer.

Irving, P. G., & Meyer, J. P. (1999). On using residual difference scores in the measurement of congruence: The case of met expectations research. *Personnel Psychology, 52,* 85-95.

Ishizaki, T., Watanabe, S., Suzuki, T., Shibata, H. & Haga, H. (2000). Predictors for functional decline among nondisabled older Japanese living in a community during a 3-year follow-up. *Journal of the American Geriatrics Society, 48,* 1424-1429.

Iverson, G. L., & Green, P. (2001). Measuring improvement or decline on the WAIS-R in inpatient psychiatry. *Psychological Reports, 89,* 457-462.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12-19.

Jama, J. W., Launer, L. J., Witteman, J. C. M., den Breeijen, J. H., Breteler, M. M. B., Grobbee, D. E., et al. (1996). Dietary antioxidants and cognitive function in a population-based sample of older persons. The Rotterdam Study. *American Journal of Epidemiology, 144,* 275-280.

Jarvik, L. F. (1975). Thoughts on the psychobiology of aging. *American Psychologist, 30,* 576-583.

Jenkins, L., Myerson, J., Joerding, J.A. & Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. *Psychology and Aging, 15,* 157-175.

Jensen, A. R. (1982). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence, Vol. 1,* (pp.255-310). Hillsdale: Erlbaum.

Jensen, M. R., Andersen, K. H. & Jensen, C. H. (1997). Management of undisplaced or minimally displaced Colles' fracture: One or three weeks of immobilisation. *Journal of Orthopaedic Science, 2,* 424-427.

Judge, J. O., Schechtman, K. & Cress, E. (1996). The relationship between physical performance measures and independence in instrumental activities of daily living. *Journal of the American Geriatrics Society, 44,* 1332-1341.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Kail, R. & Hall, L. K. (2001). Distinguishing short-term memory from working memory. *Memory and Cognition, 29,* 1-9.

Kaplan, G. A., & Kotter, P. L. (1985). Self-reports predictive of mortality from ischemic heart disease: A nine-year follow-up of the human population laboratory cohort. *Journal of Chronic Diseases, 38,* 195-201.

Kaplan, G. A., Strawbridge, W. J., Camacho, T. & Cohen, R. D. (1993). Factors associated with change in physical functioning in the elderly: A six year prospective study. *Journal of Aging and Health, 5,* 140-153.

Kaufman, A. S. & Lichtenberger, E. O. (1999). *Essentials of WAIS-III Assessment.* New York: John Wiley & Sons.

Kelly-Hayes, M., Jette, A. M., Wolf, P. A., D'Agostino, R. B. & Odell, P. M. (1992). Functional limitations and disability among elders in the Framingham Study. *American Journal of Public Health, 82,* 841-845.

Kim, Y. S., Nibbelink, D. W. & Overall, J. E. (1994). Factor structure and reliability of the Alzheimer's Disease Assessment Scale in a multicenter trial with linopirdine. *Journal of Geriatric Psychiatry and Neurology, 7,* 74-83.

Kintsch, W., Healy, A. F., Hegarty, M., Pennington, B. F. & Salthouse, T. A. (1999). Models of working memory. Eight questions and some general issues. In A. Miyake & P. Shah (Eds.), *Models of Working Memory. Mechanisms of Active Maintenance and Executive Control,* (pp.412-441). Cambridge: Cambridge University Press.

Kiosses, D. N., Klimstra, S., Murphy, C., & Alexopoulos, G. S. (2001). Executive dysfunction with disability in elderly patients with major depression. *American Journal of Geriatric Psychiatry, 9,* 269-274.

Kirasic, K. C., Allen, G. L., Dobson, S. H. & Binder, K. S. (1996). Aging, cognitive resources, and declarative learning. *Psychology and Aging, 11,* 658-670.

Klatsky, A. L., Armstrong, M., & Friedman, G. D. (1990). Risk of cardiovascular mortality in alcohol drinkers, ex-drinkers, and non-drinkers. *The American Journal of Cardiology, 66,* 1237-1242.

Klunk, W. E. (2002). Blood and CFS biomarkers for AD revisited: What's new, what's good, and is this where we should be looking? *Neurobiology of Aging, 23,* 517-519.

Kyllonen, P. C. & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14,* 389-433.

LaCroix, A. Z., Guralnik, J. M., Berkman, L. F., Wallace, R. B., & Satterfield, S. (1993). Maintaining mobility in late life. II. Smoking, alcohol consumption, physical activity and body mass index. *American Journal of Epidemiology, 137,* 858-869.

Landphair, T. (2006). *Graying US pilots question mandatory retirement.* Retrieved October 6, 2006, from http://www.voanews.com/english/AmericanLife/2006-10-03-voa48.cfm.

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1118-1133.

La Rue, A., Bank, L., Jarvik, L., & Hetland, M. (1979). Health in old age: How do physicians' ratings and self-ratings compare? *Journal of Gerontology, 34,* 687-691.

La Rue, A., Koehler, K. M., Wayne, S. J., Chiulli, S. J., Haaland, K. Y., & Garry, P. J. (1997). Nutritional status and cognitive functioning in a normally aging sample: A 6-y reassessment. *American Journal of Clinical Nutrition, 65,* 20-29.

Launer, L. J., Harris, T., Rumpel, C., & Madans, J. (1994). Body mass index, weight change and risk of mobility disability in middle-aged and older women. *Journal of the American Medical Association, 271,* 1093-1098.

Law, D. J., Morrin, K. A. & Pellegrino, J. W. (1995). Training effects and working memory contributions to skill acquisition in a complex co-ordination task. *Learning and Individual Differences, 7,* 207-234.

Lawton, M. P. & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist, 9,* 179-186.

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *The Quarterly Journal of Experimental Psychology, 49A,* 29-50.

Lezak, M.D. (1995). *Neuropsychological assessment (3rd ed).* Oxford: Oxford University Press.

Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J. & Fischer, J. S. (2004). *Neuropsychological Assessment (4th ed).* Oxford: Oxford University Press.

Li, S., Lindenberger, U. & Sikstrom, S. (2001). Aging cognition: From neuromodulation to representation. *Trends in Cognitive Sciences, 5,* 479-486.

Light, L. L. & Anderson, P. A. (1985). Working-memory capacity, age, and memory for discourse. *Journal of Gerontology, 40,* 737-747.

Lin, M. Y., Gutierrez, P. R., Stone, K. L., Yaffe, K., Ensrud, K. E., Fink, H. A., et al. (2004). Vision impairment and combined vision and hearing impairment predict cognitive and functional decline in older women. *Journal of the American Geriatrics Society, 52,* 1996-2002.

Lindenberger, U. & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong correlation. *Psychology and Aging, 9,* 339-355.

Lockwood, K. A., Alexopoulos, G. S., & van Gorp, W. G. (2002). Executive dysfunction in geriatric depression. *American Journal of Psychiatry, 159,*1119-1126.

Logie, R. H. (1995). *Visuo-spatial Working Memory.* Hove: Lawrence Erlbaum.

Logie, R. H. (1999). Working memory. *The Psychologist, 12,* 174-178.

Loisy, C. & Roulin, J. (2003). Multiple short-term storage: A task to evaluate the coordination function of the central executive. *International Journal of Psychology, 38,* 1-10.

Lord, S. R., Anstey, K. J., Williams, P. & Ward, J. A. (1995). Psychoactive medication use, sensori-motor function and falls in older women. *British Journal of Clinical Pharmacology, 39,* 227-234.

Lustig, C., May, C. P. & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General, 130,* 199-207.

Luszcz, M. A., Bryan, J., & Kent, P. (1997). Predicting episodic memory performance of very old men and women: Contributions from age, depression, activity, cognitive ability, and speed. *Psychology and Aging, 12,* 340-351.

MacDonald, M. C., Almor, A., Henderson, V. W., Kempler, D. & Andersen, E. S. (2001). Assessing working memory and language comprehension in Alzheimer's disease. *Brain and Language, 78,* 17-42.

Marengoni, A., Agüero-Torres, H., Cossi, S., Ghisla, M. K., De Martinis, M., Leonardi, R., et al. (2004). Poor mental and physical health differentially contributes to disability in hospitalized geriatric patients of different ages. *International Journal of Geriatric Psychiatry, 19,* 27-34.

Marsiske, M., Klumb, P., & Baltes, M. M. (1997). Everyday activity patterns and sensory functioning in old age. *Psychology and Aging, 12,* 444-457.

May, C. P., Hasher, L. & Kane, M. J. (1999). The role of interference in memory span. *Memory and Cognition, 27,* 759-767.

McArdle, J. J. (1988). Dynamic but structural equation modelling of repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology, 2nd ed,* (pp.561-614). New York: Plenum Press.

McCabe, J. & Hartman, M. (2003). Examining the locus of age effects on complex span tasks. *Psychology and Aging, 18,* 562-572.

McClearn, G. E. (1997). Biomarkers of age and aging. *Experimental Gerontology, 32,* 87-94.

McClure, J. B. (2002). Are biomarkers useful treatment aids for promoting health behaviour change? An empirical review. *American Journal of Preventative Medicine, 22,* 200-207.

McFarland, R. A. (1953). *Human factors in air transportation: Occupational health and safety.* New York: McGraw-Hill.

McFarland, R. A. (1968). The sensory and perceptual processes in aging. In, K. W. Schaie (Ed.), *Theory and Methods of Research on Aging: Current Topics in the Psychology of Aging: Perception, Learning, Cognition, and Personality,* (pp.9-52). Morgantown: West Virginia University.

McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment. Theories, Tests, and Issues,* (pp.151-179). New York: Guilford.

McGrew, K. S. & Flanagan, D. P. (1999). *An introduction to (CHC) Gf-Gc Theory.* New York: Institute for Applied Psychometrics, St. John's University. Retrieved July 6, 2006, from http://www.iapsych.com/gfgcint.htm.

Meguro, Y., Fujii, T., Yamadori, A., Tsukiura, T., Suzuki, K., Okuda, J., et al. (2000). The nature of age-related decline on the reading span task. *Journal of Clinical and Experimental Neuropsychology, 22,* 391-398.

Miyake, A., Emerson, M. J., & Friedman, N. P. (1999). Good interactions are hard to find. *Behavioral and Brain Sciences, 22,* 108-109.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P. & Hegarty, M. (2001). How are visuospatial working memory, executive functioning , and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General, 130,* 621-640.

Miyake, A. & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of Working Memory. Mechanisms of Active Maintenance and Executive Control,* (pp.442-482). Cambridge: Cambridge University Press.

Mohs, R. C. & Cohen, L. (1988). Alzheimer's Disease Assessment Scale (ADAS). *Psychopharmacology Bulletin, 24,* 627-628.

Mohs, R. C., Rosen, W. G. & Davis, K. L. (1983). The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bulletin, 19,* 448-450.

Mor, V., Murphy, J., Masterson-Allen, S., Willey, C., Razmpour, A., Jackson, M. E., et al. (1989). Risk of functional decline among well elders. *Journal of Clinical Epidemiology, 42,* 895-904.

Moritz, D. J., Kasl, S. V. & Berkman, L. F. (1995). Cognitive functioning and the incidence of limitations in activities of daily living in an elderly community sample. *American Journal of Epidemiology, 141,* 41-49.

Morris, M. C., Evans, D. A., Bienias, J. L., Tangney, C. C., & Wilson, R. S. (2002). Vitamin E and cognitive decline in older persons. *Archives of Neurology, 59,* 1125-1132.

Morris, R. G., Gick, M. L. & Craik, F. I. M. (1988). Processing resources and age differences in working memory. *Memory and Cognition, 16,* 362-366.

Morris, N. & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81,* 111-121.

Müller, U., Werheid, K., Hammerstein, E., Jungmann, S. & Becker, T. (2005). Prefrontal cognitive deficits in patients with schizophrenia treated with atypical or conventional antipsychotics. *European Psychiatry, 20,* 70-73.

Mulrow, C. D., Gerety, M. B., Cornell, J. E., Lawrence, V. A., & Kanten, D. N. (1994). The relationship between disease and function and perceived health in very frail elders. *Journal of the American Geriatrics Society, 42,* 374-380.

Murray, I. M. (1951). Assessment of physiologic age by combination of several criteria – vision, hearing, blood pressure and muscle force. *Journals of Gerontology, 6,* 120-126.

Myerson, J., Emery, L., White, D. A. & Hale, S. (2003). Effects of age, domain, and processing demands on memory span: Evidence for differential decline. *Aging, Neuropsychology, and Cognition, 10,* 20-27.

National Health and Medical Research Council (1991). *Recommended Dietary Intakes for use in Australia.* Canberra: Australian Government.

Nelson, H. D., Nevitt, M .C., Scott, J. C., Stone, K. L., & Cummings, S. R. (1994). Smoking, alcohol, and neuromuscular and physical function of older women. *Journal of the American Medical Association, 272,* 1825-1831.

Nelson, H. E. (1982). *The National Adult Reading Test (NART).* Windsor: NFER NELSON.

Nesselroade, J. R., Stigler, S. M., & Baltes, P. B. (1980). Regression toward the mean and the study of change. *Psychological Bulletin, 88,* 622-637.

Njegovan, V., Man-son-Hing, M., Mitchell, S. L. & Molnar, F. J. (2001). The hierarchy of functional loss associated with cognitive decline in older persons. *Journal of Gerontology. Series A: Biological Sciences and Medical Sciences, 56A,* M638-M643.

Norman, D. A. & Shallice, T. (1980). *Attention to action: Willed and automatic control of behaviour. (Technical report #99).* San Diego: Centre for Human Information Processing, University of California.

Numminen, H., Service, E., Ahonen, T. & Ruoppila, I. (2001). Working memory and everyday cognition in adults with Down's syndrome. *Journal of Intellectual Disability Research, 45,* 157-168.

Oberauer, K., Süss, H. M., Schulze, R., Wilhelm, O. & Wittman, W.W. (2000). Working memory capacity – facets of a cognitive ability construct. *Personality and Individual Differences, 29,* 1017-1045.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments and Computers, 32,* 396-402.

Okada, A. (2002). Deficits of spatial working memory in chronic schizophrenia. *Schizophrenia Research, 53,* 75-82.

O'Leary, U. M., Rusch, K. M., & Guastello, S. J. (1991). Estimating age-stratified WAIS-R IQs from scores on the Raven's Standard Progressive Matrices. *Journal of Clinical Psychology, 47,* 277-284.

Oman, D., Reed, D., & Ferrara, A. (1999). Do elderly women have more physical disability than men do? *American Journal of Epidemiology, 150,* 834-842.

Omron. (2001). *Omron T5 Model Automatic Blood Pressure Monitor Manual.* Singapore: Omron Healthcare.

Ortega, R. M., Requejo, A. M., Andrés, P., López-Sobaler, A. M., Quintas, M. E., Redondo, M. R., et al. (1997). Dietary intake and cognitive function in a group of elderly people. *American Journal of Clinical Nutrition, 66,* 803-809.

Ostir, G. V., Markides, K. S., Black, S. A. & Goodwin, J. S. (2000). Emotional well being predicts subsequent functional independence and survival. *Journal of the American Geriatrics Society, 48,* 473-478.

Paleologos, M., Cumming, R. G., & Lazarus, R. (1998). Cohort study of vitamin C intake and cognitive impairment. *American Journal of Epidemiology, 148,* 45-50.

Palmore, E., & Cleveland, W. (1976). Aging, terminal decline, and terminal drop. *Journal of Gerontology, 31,* 76-81.

Papassotiropoulos, A. & Hock, C. (2002). Biochemical markers of Alzheimer's disease: Wish and reality. *Neurobiology of Aging, 23,* 513-514.

Park, D. C. (1997). Psychological issues related to competence: Cognitive aging and instrumental activities of daily living. In W. Schaie & S. Willis (Eds.), *Social structures and aging,* (pp.66-82). Mahwah: Lawrence Erlbaum.

Park, D. C. (1999). Aging and the controlled and automatic processing of medical information and medical intentions. In D. C. Park, R. W. Morrell, & K. Shifren (Eds.), *Processing of medical information in aging patients: Cognitive and human factors perspectives* (pp.3-22). Mahwah: Lawrence Erlbaum.

Park, D. C. (2000). The basic mechanisms accounting for age-related decline in cognitive function. In D.C. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer,* (pp.3-21). Philadelphia: Psychology Press.

Park, D.C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P.K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17,* 299-320.

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Frieske, D., Zwahr, M., et al. (1996). Mediators of long-term memory performance across the life span. *Psychology and Aging, 11,* 621-637.

Parker, M. G., Thorslund, M., Lundberg, O. & Kareholt, I. (1996). Predictors of physical function among the oldest old. *Journal of Aging and Health, 8,* 444-460.

Parkin, A. J. (1993). *Memory. Phenomena, Experiment and Theory.* Oxford: Blackwell.

Peña-Casanova, J. (1997). Alzheimer's Disease Assessment Scale – Cognitive in clinical practice. *International Psychogeriatrics, 9,* 105-114.

Perkins, A.J., Hendrie, H. C., Callahan, C. M., Gao, S., Unverzagt, F. W., Xu, Y., et Al. (1999). Association of antioxidants with memory in a multiethnic elderly sample using the Third National Health and Nutrition Examination Survey. *American Journal of Epidemiology, 150,* 37-44.

Perlis, R. H., Davidoff, D., Falk, W. E., Round, D. L., Verma, S., & Sederer, L. I. (2002). Outcome measurement in geriatric psychiatry. In W. W. IsHak, T. Burt & L. I. Sederer (Eds.), *Outcome Measurement in Psychiatry. A Critical Review,* (pp.77-96). Washington, D. C: American Psychiatric.

Phillips, L. H., & Hamilton, C. (2001). The working memory model in adult aging research. In J. Andrade (Ed.), *Working memory in perspective,* (pp.101-125). East Sussex: Psychology Press.

Piantanelli, L., Rossolini, G., & Basso, A. (1992). Biomarkers of aging and survival kinetics. In N. Fabris (Ed.), *Physiopathological Processes of Aging: Towards a Multicausal Interpretation,* (pp.9-15). New York: New York Academy of Sciences.

Pitts, D. G. (1982). Visual acuity as a function of age. *Journal of the American Optometric Association, 53,* 117-124.

Pollock, B. G. (2002). Biomarkers. From bedside to bench and back again. *American Journal of Geriatric Psychiatry, 10,* 644-645.

Powell, R. R., & Pohndorf, R. H. (1971). Comparison of adult exercisers and non exercisers on fluid intelligence and selected physiological variables. *Research Quarterly, 42,* 70-77.

Prabhakaran, V., Smith, J., Desmond, J., Glover, G., & Gabrieli, J. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices. *Cognitive Psychology, 33,* 43-63.

Prouteau, A., Verdoux, H., Briand, C., Lesage, A., Lalonde, P., Nicole, L., et al. (2004). Self-assessed cognitive dysfunction and objective performance in outpatients with schizophrenia participating in a rehabilitation program. *Schizophrenia Research, 69,* 85-91.

Qpzm Australia. (2004). *Postcodes.* Retrieved January 10, 2004, from http://postcodes.qpzm.com.au/au/sa/adelaide/eastern-suburbs.

Rabbitt, P., Diggle, P., Smith, D., Holland, F., & Mc Innes, L. (2001). Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. *Neuropsychologia, 39,* 532-543.

Rantanen, T., Era, P., Kauppinen, M., & Heikkinen, E. (1994). Maximal isometric muscle strength and socio-economic status, health and physical activity in 75-year-old persons. *Journal of Aging and Physical Activity, 2,* 206-220.

Raven, J. C. (1958). *Mill Hill Vocabulary Scale.* London: H. K. Lewis.

Raykov, T. (1992). Structural models for studying correlates and predictors of change. *Australian Journal of Psychology, 44,* 101-112.

Riggs, K. M., Spiro, A., Tucker, K., & Rush, D. (1996). Relations of vitamin B-12, vitamin B-6, folate and homocysteine to cognitive performance in the normative aging study. *American Journal of Clinical Nutrition, 63*, 306-314.

Rimm, E. B., Giovannucci, E. L., Willett, W. C., Colditz, G. A., Ascherio, A., Rosner, B., et al. (1991). Prospective study of alcohol consumption and risk of coronary disease in men. *The Lancet, 338,* 464-468.

Roberts, B. W., & Chapman, C. N. (2000). Change in dispositional well-being and its relation to role quality: A 30-year longitudinal study. *Journal of Research in Personality, 34,* 26-41.

Roberts, R. & Gibson, E. (2002). Individual differences in sentence memory. *Journal of Psycholinguistic Research, 31,* 573-598.

Rockey, L. S. (1997). Memory assessment of the older adult. In P. D. Nussbaum (Ed.), *Handbook of Neuropsychology and Aging,* (pp.385-393). New York: Plenum Press.

Rosen, V., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory and Language, 39,* 418-436.

Rosen, W. G., Mohs, R. C. & Davis, K. L. (1984). A new rating scale for Alzheimer's Disease. *American Journal of Psychiatry, 141,* 1356-1364.

Ruchkin, D. S., Grafman, J., Krauss, G. L., Johnson, R., Canoune, H. & Ritter, W. (1994). Event-related brain potential evidence for a verbal working memory deficit in multiple sclerosis. *Brain, 117,* 289-305.

Ryan, J. J., Sattler, J. M. & Lopez, S. J. (2000). Age effects on Wechsler Adult Intelligence Scale-III subtests. *Archives of Clinical Neuropsychology, 15,* 311-317.

Salamon, M. J. (1985). A clinical application for life satisfaction. *Clinical Gerontologist, 3,* 60-61.

Salamon, M. J. (1988). Clinical use of the Life Satisfaction in the Elderly Scale. *Clinical Gerontologist, 8,* 45-55.

Salamon, M. J., & Conte, V. A. (1984). *The Life Satisfaction in the Elderly Scale Manual.* Odessa, Florida: Psychological Assessment Resources.

Salive, M. E., Guralnik, J., Glynn, R. J., Christen, W., Wallace, R. B., & Ostfeld, A. M. (1994). Association of visual impairment with mobility and physical function. *Journal of the American Geriatrics Society, 42,* 287-292.

Salthouse, T. A, (1991a). Cognitive facets of aging well. *Generations: Journal of the American Society on Aging, 15,* 35-38.

Salthouse, T. A. (1991b). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2,* 179-183.

Salthouse, T. A. (1993). Influence of working memory on adult age differences in matrix reasoning. *British Journal of Psychology, 84,* 171-199.

Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology, 8,* 535-543.

Salthouse, T. A. (1995). Differential age-related influences on memory for verbal symbolic information and visual-spatial information. *Journal of Gerontology. Series B: Psychological Sciences and Social Sciences, 50B,* P193-P201.

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science, 13,* 140 – 144.

Salthouse, T. A. & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27,* 763-776.

Salthouse, T. A., Hambrick, D. Z. & McGuthry, K. E. (1998). Shared age-related influences on cognitive and noncognitive variables. *Psychology and Aging, 13,* 486-500.

Salthouse, T. A., Hancock, H. E., Meinz, E. J., & Hambrick, D. Z. (1996). Interrelations of age, visual acuity, and cognitive functioning. *Journal of Gerontology. Series B: Psychological Sciences and Social Sciences, 51B,* P317-P330.

Salthouse, T. A., Kausler, D. H., & Saults, S . J. (1988). Investigation of student status, background variables, and the feasibility of standard tasks in cognitive aging research. *Psychology and Aging, 3,* 29-37.

Salthouse, T. A., Kausler, D. H., & Saults, S . J. (1990). Age, self-assessed health status, and cognition. *Journal of Gerontology. Series B: Psychological Sciences and Social Sciences, 45,* P156-P160.

Salthouse, T. A., Mitchell, D. R. D., Skovronek, E. & Babcock, R. L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 507-516.

Sawrie, S. M., Marson, D. C., Boothe, A. L., & Harrell, L. E. (1999). A method for assessing clinically relevant individual cognitive change in older adult populations. *Journal of Gerontology. Series B: Psychological Sciences and Social Sciences, 54B,* P116-P124.

Schaie, K. W. (1975). Adult development and aging. *Annual Review of Psychology, 26,* 65-96.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist, 49,* 304-313.

Schaie, K. W. & Willis, S. L. (1993). Age difference patterns of psychometric intelligence in adulthood: Generalisability within and across ability domains. *Psychology and Aging, 8,* 44-55.

Schieber, F. (1992). Aging and the senses. In J. E. Birren, R. B. Sloane & G. D., Cohen (Eds.), *Handbook of Mental Health and Aging,* (pp.251-306). San Diego: Academic Press.

Schinka, J. A., Vanderploeg, R. D., Rogish, M., Borenstein, G. A., Mortimer, J. A., & Ordorica, P. I. (2002). Effects of the use of alcohol and cigarettes on cognition in elderly adults. *Journal of the International Neuropsychological Society, 8,* 811-818.

Schulz, R. & Heckhausen, J. (1996). A life span model of successful aging. *American Psychologist, 51,* 702-714.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 298,* 199-209.

Sherman, E. M. S., Slick, D. J., Connolly, M. B., Steinbok, P., Martin, R., Strauss, E., et al. (2003). Re-examining the effects of epilepsy surgery on IQ in children: Use of regression-based change scores. *Journal of the International Neuropsychological Society, 9,* 879-886.

Shock, N. W. (1947). Older people and their potentialities for gainful employment. *Journal of Gerontology, 2,* 93-102.

Silvestri, A., Visconti, G., Fini, M., Pagano, S., Cisternino, D., Mochi, M., et al. (2001). Prevalence of depression and cognitive deficit in a geriatric population. *Archives of Gerontology and Geriatrics, suppl. 7,* 351-355.

Simonsick, E. M., Lafferty, M. E., Phillips, C. L., Mendes de Leon, C. F., Kasl, S. V., Seeman, T. E., et al. (1993). Risk due to inactivity in physically capable older adults. *American Journal of Public Health, 83,* 1443-1450.

Spector, W. D., Katz, S., Murphy, J. B., & Fulton, J. P. (1987). The hierarchical relationship between activities of daily living and instrumental activities of daily living. *Journal of Chronic Diseases, 40,* 481-489.

Speer, D. C. (1992). Methodological developments. Clinically significant change: Jacobson and Truax (1991) revisited. *Journal of Consulting and Clinical Psychology, 60,* 402-408.

Spreen, O. & Strauss, E. (1998). *A Compendium of Neuropsychological Tests. Administration, Norms, and Commentary.* Oxford: Oxford University Press.

Stampfer, M. J., Kang, J. H., Chen, J., Cherry, R., & Grodstein, F. (2005). Effects of moderate alcohol consumption on cognitive function in women. *New England Journal of Medicine, 352,* 245-253.

Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research, 2,* 21-33.

Stine, E. A. L. & Wingfield, A. (1990). How much do working memory deficits contribute to age differences in discourse memory? *European Journal of Cognitive Psychology, 2,* 289-304.

Stip, E., Caron, J., Renaud, S., Pampoulova, T., Lecomte, Y. (2003). Exploring cognitive complaints in schizophrenia: The Subjective Scale to Investigate Cognition in Schizophrenia. *Comprehensive Psychiatry, 44,* 331-340.

Stuck, A. E., Walthert, J. M., Nikolaus, T., Bula, C. J., Hohmann, C., & Beck, J. C. (1999). Risk factors for functional status decline in community-living elderly people: A systematic literature review. *Social Science and Medicine, 48,* 445-469.

Sulander, T., Martelin, T., Rahkonen, O., Nissinen, A. & Uutela, A. (2005). Associations of functional ability with health-related behaviour and body mass index among the elderly. *Archives of Gerontology and Geriatrics, 40,* 185-199.

Süss, H., Oberauer, K., Wittman, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability – and a little bit more. *Intelligence, 30,* 261-288.

Tabbarah, M., Crimmins, E. M. & Seeman, T. E. (2002). The relationship between cognitive and physical performance: MacArthur studies of successful aging. *Journal of Gerontology. Series A: Biological Sciences and Medical Sciences, 57A,* M228-M235.

Tafaro, L., Cicconetti, P., Martella, S., Tedeschi, G., Zannino, G., Troisi, G., et al. (2001). An investigation of behavioural problems in centenarians. *Archives of Gerontology and Geriatrics, suppl. 7,* 375-378.

Tinetti, M. E., Inouye, S. K., Gill, T. M. & Doucette, J. T. (1995). Shared risk factors for falls, incontinence, and functional dependence. *Journal of the American Medical Association, 273,* 1348-1353.

Tirre, W. C., & Peña, C. M. (1992). Investigation of functional working memory in the reading span test. *Journal of Educational Psychology, 84,* 462-472.

Tombaugh, T. N. (2005). Test-retest reliable coefficients and 5-year change scores for the MMSE and 3MS. *Archives of Clinical Neuropsychology, 20,* 485-503.

Towse, J. N., Hitch, G. J. & Hutton, U. (2000). On the interpretation of working memory span in adults. *Memory and Cognition, 28,* 341-348.

Transport South Australia. (2004). *Licences and Certification*. Retrieved July 13, 2004, from http://www.transport.sa.gov.au/licences_certification.

Turner, M. L. & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127-154.

Van der Linden, M., Beerten, A. & Pesenti, M. (1998). Age-related differences in random generation. *Brain and Cognition, 38,* 1-16.

Van der Linden, M., Brédart, S. & Beerten, A. (1994). Age-related differences in updating working memory. *British Journal of Psychology, 85,* 145-152.

Van der Linden, M., Hupet, M., Feyereisen, P., Schelstraete, M., Bestgen, Y., Bruyer, R., et al. (1999). Cognitive mediators of age-related differences in language comprehension and verbal memory performance. *Aging, Neuropsychology, and Cognition, 6,* 32-55.

Vecchi, T., & Cornoldi, C. (1999). Passive storage and active manipulation in visuo spatial working memory: Further evidence from the study of age differences. *European Journal of Cognitive Psychology, 11,* 391-406.

Verhaeghen, P., Marcoen, A., & Goossens, L. (1993). Facts and fiction about memory aging: A quantitative integration of research findings. *Journal of Gerontology. Series B: Psychological Sciences and Social Sciences, 48,* P157-P171.

Verhaeghen, P. & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin, 122,* 231-249.

Victorian Government (2003). *Better Health Channel. Body Mass Index*. Retrieved July 9, 2004 from http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Body_Mass_Index_(BMI)?OpenDocument.

Waldstein, S. R. (2003). The relation of hypertension to cognitive function. *Current Directions in Psychological Science, 12,* 9-12.

Waldstein, S. R., Manuck, S. B., Ryan, C. M., & Muldoon, M. F. (1991). Neuropsychological correlates of hypertension: Review of methodologic considerations. *Psychological Bulletin, 110,* 451-468.

Wang, L., van Belle, G., Kukull, W. B. & Larson, E. B. (2002). Predictors of functional change: A longitudinal study of nondemented people aged 65 and older. *Journal of the American Geriatrics Society, 50,* 1525-1534.

Wareing, M., Fisk, J. E., Murphy, P. & Montgomery, C. (2005). Visuo-spatial working memory deficits in current and former users of MDMA ('ecstasy'). *Human Psychopharmacology: Clinical and Experimental, 20,* 115-123.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology, 49A,* 51-79.

Waters,G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments and Computers, 35,* 550-564.

Webster, I. W. & Logie, A. R. (1976). A relationship between functional age and health status in female subjects. *Journal of Gerontology, 31,* 546-550.

Wechsler, D. (1945). A standardized memory scale for clinical use. *The Journal of Psychology, 19,* 87-95.

Wechsler, D. (1987). *Wechsler Memory Scale – Revised Manual.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale – III.* San Antonio, TX: Psychological Corporation.

Welford, A. T. (1958). *Ageing and Human Skill.* Oxford: Oxford University Press.

Werheid, K., Hoppe, C., Thöne, A., Müller, U., Mungersdorf, M. & von Cramon, D. Y. (2002). The Adaptive Digit Ordering Test: Clinical application, reliability, and validity of a verbal working memory test. *Archives of Clinical Neuropsychology, 17,* 547-565.

Weyer, G., Erzigkeit, H., Kanowski, S., Ihl, R. & Hadler, D. (1997). Alzheimer's Disease Assessment Scale: Reliability and validity in a multicenter clinical trial. *International Psychogeriatrics, 9,* 123-138.

Whitmer, R. A., Gunderson, E. P., Barrett-Connor, E., Quesenberry, C. P., & Yaffe, K. (2005). Obesity in middle age and future risk of dementia: A 27 year longitudinal population based study. *British Medical Journal, 330,* no pagination specified.

Whitney, P., Arnett, P. A. & Driver, A. (2001). Measuring central executive functioning: What's in a reading span? *Brain and Cognition, 45,* 1-14.

Wiederholt, W. C., Cahn, D., Butters, N. M., Salmon, D. P., Kritz-Silverstein, D., & Barrett-Connor, E. (1993). Effects of age, gender and education on selected neuropsychological tests in an elderly community cohort. *Journal of the American Geriatrics Society, 41,* 639-647.

Willet, J. B. (1988). Questions and answers in the measurement of change. *Review of Research in Education, 15,* 345-422.

Williams, R. H., & Zimmerman, D. W. (1982). The comparative reliability of simple and residualised difference scores. *Journal of Experimental Education, 51,* 94-97.

Williams, R. H., Zimmerman, D. W., Rich, J. M., & Steed, J. L. (1984a). Empirical estimates of the validity of four measures of change. *Perceptual and Motor Skills, 58,* 891-896.

Williams, R. H., Zimmerman, D. W., Rich, J. M., & Steed, J. L. (1984b). An empirical study of the relative error magnitude in three measures of change. *Journal of Experimental Education, 53,* 55-57.

Willis, S. L. (1996). Everyday problem solving. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the Psychology of Aging, 4th ed* (pp.287-307). San Diego: Academic Press.

Willis, S. L., Jay, G. M., Diehl, M. & Marsiske, M. (1992). Longitudinal change and prediction of everyday task competence in the elderly. *Research on Aging, 14,* 68-91.

Wilson, R. S., Barnes, L. L., & Bennett, D. A. (2003). Assessment of lifetime participation in cognitively stimulating activities. *Journal of Clinical and Experimental Neuropsychology, 25,* 634-642.

Wilson, R. S., Beckett, L. A., Barnes, L .L., Schneider, J. A., Bach, J., Evans, D. A., et al. (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging, 17,* 179-193.

Wolinsky, F. D., Stump, T. E., & Clark, D. O. (1995). Antecedents and consequences of physical activity and exercise among older adults. *The Gerontologist, 35,* 451-462.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psycho Educational Battery – Revised.* Allen, TX: DLM.

Yesavage, J.A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M. et al. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17,* 37-49.

Yesavage, J. A., Poulsen, S. L., Sheikh, J., & Tanke, E. (1988). Rates of change of common measures of impairment in senile dementia of the Alzheimer's Type. *Psychopharmacology Bulletin, 24,* 531-534.

Zelinski, E. M., & Burnight, K. P. (1997). Sixteen-year longitudinal and time lag changes in memory and cognition in older adults. *Psychology and Aging, 12,* 503-513.

Zimmerman, D. W., & Williams, R. H. (1982). The relative error magnitude in three measures of change. *Psychometrika, 47,* 141-147.

Zimprich, D., & Martin, M. (2002). Can longitudinal changes in processing speed explain longitudinal age changes in fluid intelligence? *Psychology and* Aging, 17, 690-695.

Appendix A: Copy of the Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog)

SCORE SUMMARY SHEET		Score
1 WORD RECALL	(maximum 10)	☐
2 NAMING OBJECTS AND FINGERS	(maximum 5)	☐
3 COMMANDS	(maximum 5)	☐
4 CONSTRUCTIONAL PRAXIS	(maximum 5)	☐
5 IDEATIONAL PRAXIS	(maximum 5)	☐
6 ORIENTATION	(maximum 8)	☐
7 WORD RECOGNITION TASK	(maximum 12)	☐
8 REMEMBERING TEST INSTRUCTIONS	(maximum 5)	☐
9 SPOKEN LANGUAGE ABILITY	(maximum 5)	☐
10 WORD-FINDING DIFFICULTY IN SPONTANEOUS SPEECH	(maximum 5)	☐
11 COMPREHENSION	(maximum 5)	☐
12 CONCENTRATION/DISTRACTIBILITY	(maximum 5)	☐
	TOTAL SCORE (maximum 75)	☐

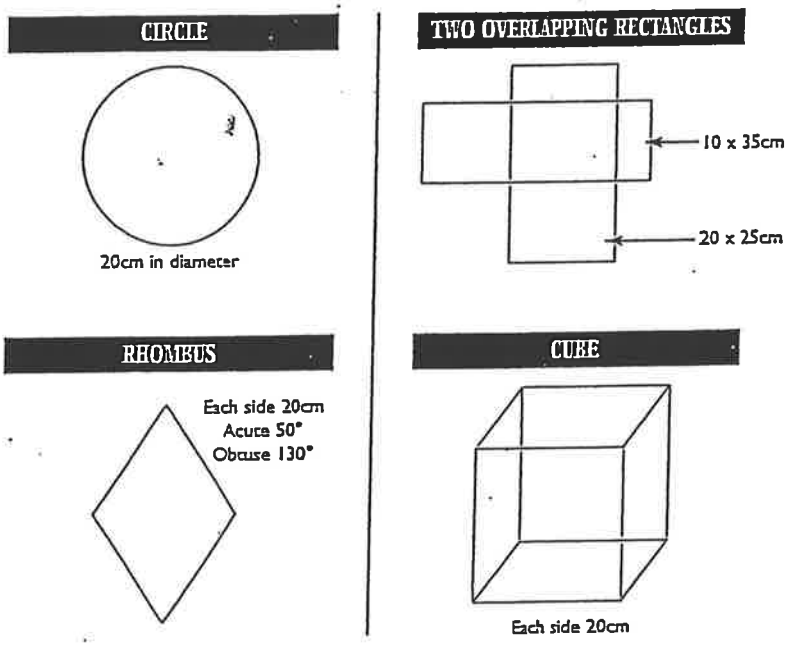
TOOLS NEEDED FOR ADAS-Cog

1 WORD RECALL – Set of cards with the following words on them:
Home, Coin, Railroad, Child, Army, Flag, Skin, Library, Wheat, Ocean.

2 NAMING OBJECTS AND FINGERS – The following objects are needed:
Flower (plastic), Bed (doll's house furniture), Whistle, Pencil, Rattle, Mask, Scissors, Comb, Wallet, Harmonica, Stethoscope, Tweezers.

3 COMMANDS – Pencil, Watch, Card.

4 CONSTRUCTIONAL PRAXIS – Pieces of paper with the following drawings on them.



5 IDEATIONAL PRAXIS – 8½" X 11" Sheet of Paper, Long Envelope, Pencil.

7 WORD RECOGNITION TASK – Set of cards with the following words on them:
Corn, Effort, Party, River, Folly, Locker, Event, Queen, Position, Quality, Sunset, Dove, Belief, Umbrella, Allegory, Hound, Idiom, Hint, Missile, Gem, Proxy, Lobster, Criterion, Deceit, Officer, Thought, Camp, Fate, Golf, Permission, Blister, Concept, Pianist, Gender, Bullet, Intellect, Plant, Amount, Industry, Occasion, Cradle, Banality, Singer, Hypothesis, Noose, Distinction, Tank, Decree.

1 WORD RECALL

For this item the patient is shown 10 cards, each with a different word written on them. The patient is asked t remember the 10 words. At the start of the first trial, the tester gives instructions similar to the following:

"I am going to show you some words, one at a time. Please read each word out loud and try to remember it because later I will ask you to try to remember all of the words I have shown you".

The subject reads aloud 10 words exposed for two seconds each. The subject then recalls the words aloud The 10 cards are shown to the patient again in a different order. Three trials of reading and recalling are given.

2 NAMING OBJECTS AND FINGERS

For this item the patient is asked to name 12 randomly presented real objects, whose frequency values are: high, medium, and low. The patient is also asked to name the fingers of his/her dominant hand, i.e. thumb, little finger, index (pointer, forefinger), middle and ring fingers. The first question about each object should be:

"What is this called?" or "What is the name of this thing?".

If the patient does not respond, then the examiner should give the cue for that item listed. If the patient still doesn't respond or makes an error, go on to the next object. For many of the objects, there is more than one correct response. A response other than the name given should be scored as correct if it is a name that would be used by a nondemented person with the same cultural background as the patient. Descriptions of the object, semantic or phonemic paraphasias, should not be scored as correct. Examples of incorrect responses are "listening thing" for stethoscope, "cutter" for scissors and "prongs" for "tongs".

3 COMMANDS

This task is designed to assess receptive speech. The patient is asked to carry out 1 to 5-step commands. Each command should be read once. If the patient does not respond or makes an error, the tester should then give the entire command one more time. Then go on to the next command. All commands should be given to every patient.

4 CONSTRUCTIONAL PRAXIS

This task assesses the patient's ability to copy four geometric forms ranging from a very simple one, a circle, to a fairly difficult one, a cube. The forms should be presented one at a time, and each form should be on a plain piece of white paper. The tester should give the patient a lead pencil with an eraser along with the drawing. The instructions to the patient should be:

"On this piece of paper is a shape. Try to draw another one that looks just like this, somewhere on the page".

The patient should be allowed two attempts for each shape, and the patient may erase if they need to. If the patient cannot reproduce the figure in two attempts, the tester should go on to the next item.

5 IDEATIONAL PRAXIS

This task is designed to determine whether the patient can perform a familiar but complex sequence of actions. The patient is given a sheet of paper, a long envelope, and a pencil. The tester should give the patient instructions similar to the following:

"I want you to pretend that you have written yourself a letter. Take this piece of paper, fold it so that it will fit into the envelope, and then put it into the envelope. Then seal the envelope, address the envelope to yourself, and show me where the stamp goes".

There are five components to this task, and each one is underlined in the instructions. If the patient forgets part of the task, or is having difficulty, the tester should repeat the instruction for the component of the task where the patient is having difficulty. For example, if the patient stops after folding the paper and putting it into the envelope, the tester should give one reminder on the next component, "Now seal the envelope". If the patient cannot do this part, move on and give one reminder on the next component, "Now address the letter to yourself". After the first complete instruction, only one additional reminder should be given for each component. Impairment in this item should reflect dysfunction in executing an overlearned task only and not recall difficulty. Note that any address which would enable a postal worker to deliver the envelope is counted as correct, even though it might not be the patient's current address. The address should contain the following: name, street, city, and state.

6 ORIENTATION

The components of orientation are person, day of the week, date, month, year, season, time of day, and place. The tester should ask the patient for each of these pieces of information one at a time. Before giving this item the tester should be sure that no clocks, watches or calendars are visible to aid the patient. One point is given for each incorrect response (maximum = 8). Acceptable answers include the following: ± 1 day for the date; within one hour for the hour; partial name for place; naming of upcoming season within one week prior to its onset; and name of previous season for two weeks after its termination. Month, year, day of the week, and the person's first and last name must be exact.

7 WORD RECOGNITION TASK

On this task the patient is given three trials to learn a list of 12 words. The learning part of each trial is similar to the learning part of the word recall task, since the patient is asked to read each word aloud and try to remember it. For each of the three test trials, the 12 studied words are mixed with 12 new words matched to the studied words for frequency and imagery, and the patients is asked to decide for each word whether or not it was one of the studied words. At the start of the first trial, the tester gives instructions to the patient similar to the following:

> *"I am going to show you some words printed on cards. I want you to read each word out loud and try to remember it".*

Some of the words on the word recognition task may not be familiar to the patient, and the patient may have difficulty reading them. If the patient cannot read a word, the tester should say the word out loud. However, it is important for the patient to actually look at each word and try to read it. At the end of the learning portion of a trial, the tester should say to the patient something like the following:

> *"Now I'm going to show you another set of words. Some of the words were on the list I just showed you and others are new. For each word, I want you to tell me whether it is one of the words I just showed you".*

Then the tester shows the first word and says either

> *"Is this one of the words I showed you before, yes or no?" or "Did I show you this word before?"*

The same instruction is given before the second test word. For the remaining test words the tester should say: "How about this one?" If the patient does not remember the task (e.g., reads the word rather than responding yes or no) then the tester should repeat or rephrase the entire question and make a note that the patient had to be reminded of the task instructions. Trials two and three are similar and the tester should

8 REMEMBERING TEST INSTRUCTIONS

This item evaluates the patient's ability to remember the requirements of the recognition task. On each recognition trial, the patient is asked prior to presentation of the first two words, "Did I show you this word before, or is this a new word?" For the third word, the patient is asked, "How about this one?" If the patient responds appropriately, i.e., yes or no, then memory for the instruction is accurate. If the patient fails to respond, this signifies that the instructions have been forgotten, and the instructions are repeated. The procedure used for the third word is repeated for words 4-24. Each instance of memory failure for the test instructions is noted.

9 SPOKEN LANGUAGE ABILITY

This item is a global rating of the quality of speech i.e., clarity, difficulty in making oneself understood. In rating this item the tester should consider all of the speech produced by the patient during the test session. Quantity of speech and word finding difficulty are not rated on this item. It should be noted that the higher scores (4-5) on this item are reserved for patients whose expressive language abilities are impaired to such an extent that they seldom communicate without difficulty.

10 WORD-FINDING DIFFICULTY IN SPONTANEOUS SPEECH

Along with item 9, this item rates impairment in expressive speech, but it rates only word-finding difficulty, whereas item 9 is a more global rating of the extent to which the patient can communicate verbally. To rate this item the tester must determine whether the patient has difficulty in finding the desired word in spontaneous speech. The problem may be overcome by circumlocution i.e., giving explanatory phrases or nearly satisfactory synonyms. Do not include finger and object naming in this rating.

11 COMPREHENSION

This item evaluates the patient's ability to understand speech. To rate this item the tester should consider how well the patient was able to understand the tester's speech during the opening discussion, during the test session and, if applicable, the administration of the noncognitive items. Do not include responses to commands.

12 CONCENTRATION/DISTRACTIBILITY

This item rates the frequency with which the patient is distracted by irrelevant stimuli and/or must be reoriented to the ongoing task because of loss of train of thought or the patient appears to be caught up in his/her own thoughts.

1 WORD RECALL

At the start of the first trial, the tester gives instructions similar to the following:

"I am going to show you some words, one at a time. Please read each word out loud and try to remember it, because later I will ask you to try to remember all of the words I have shown you".

The subject reads aloud 10 words exposed for 2 seconds each. The subject then recalls the words aloud. Three trials of reading and recalling are given.

TRIAL 1			TRIAL 2			TRIAL 3		
	Recalled	Not Recalled		Recalled	Not Recalled		Recalled	Not Recalled
Home	☐	☐	Skin	☐	☐	Railroad	☐	☐
Coin	☐	☐	Child	☐	☐	Ocean	☐	☐
Railroad	☐	☐	Wheat	☐	☐	Flag	☐	☐
Child	☐	☐	Library	☐	☐	Army	☐	☐
Army	☐	☐	Home	☐	☐	Wheat	☐	☐
Flag	☐	☐	Ocean	☐	☐	Child	☐	☐
Skin	☐	☐	Railroad	☐	☐	Coin	☐	☐
Library	☐	☐	Flag	☐	☐	Skin	☐	☐
Wheat	☐	☐	Coin	☐	☐	Home	☐	☐
Ocean	☐	☐	Army	☐	☐	Library	☐	☐
Total not recalled		☐	Total not recalled		☐	Total not recalled		☐

Indicate the total number of words <u>not recalled</u> for each trial.

Score ☐ score = mean number of words <u>not</u> recalled on three trials – (maximum = 10.)

2 NAMING OBJECTS AND FINGERS

The subject names 12 randomly presented real objects. The first question about each object should be:

"What is this called?" or "What is the name of this thing?"

If the subject does not respond, then the examiner should give the cue for that item listed below. If the subject still doesn't respond or makes an error, go on to the next object.

Objects	Standard clue that can be used to assist those subjects having difficulties	Correct	Incorrect (or not named)
Flower	- grows in the garden	☐	☐
Bed	- used for sleeping	☐	☐
Whistle	- makes a sound when you blow it	☐	☐
Pencil	- used for writing	☐	☐
Rattle	- a baby's toy	☐	☐
Mask	- hides your face	☐	☐
Scissors	- cuts paper	☐	☐
Comb	- used on hair	☐	☐
Wallet	- holds your money	☐	☐
Harmonica	- a musical instrument	☐	☐
Stethoscope	- doctor uses it to listen to your heart	☐	☐
Tweezers	- used to pick up things	☐	☐

The subject names the fingers on his/her dominant hand

Finger	Correct	Incorrect (or not named)
Thumb	☐	☐
Index	☐	☐
Middle	☐	☐
Ring	☐	☐
Little Finger	☐	☐

Total incorrect ☐

Score:
- 0 0-2 items named incorrectly (items: objects and fingers named)
- 1 3-5 items named incorrectly
- 2 6-8 items named incorrectly
- 3 9-11 items named incorrectly
- 4 12-14 items named incorrectly
- 5 15-17 items named incorrectly

Score ☐ (maximum 5)

3 COMMANDS

Ask the subject to carry out the following commands. Each command should be read once. If the subject does not respond or makes an error, the tester should then give the entire command one more time. Then go on to the next command. All commands should be given to the subject. Indicate each command performed correctly or incorrectly.

	Correct	Incorrect (or not performed)
Make a fist	☐	☐
Point to the ceiling and then to the floor	☐	☐
Line up a pencil, watch, and card, in that order, on a table in front of the subject.		
Put the pencil on top of the card and then put it back.	☐	☐
Put the watch on the other side of the pencil and then turn over the card.	☐	☐
Tap each shoulder twice, with two fingers, keeping your eyes shut.	☐	☐

Each underlined element represents a single step. Each command is scored as a whole.

Score:

0	all commands correct
1	1 command incorrect, 4 commands correct
2	2 commands incorrect, 3 commands correct
3	3 commands incorrect, 2 commands correct
4	4 commands incorrect, 1 command correct
5	all 5 commands incorrect

Score ☐
(maximum 5)

4 CONSTRUCTIONAL PRAXIS

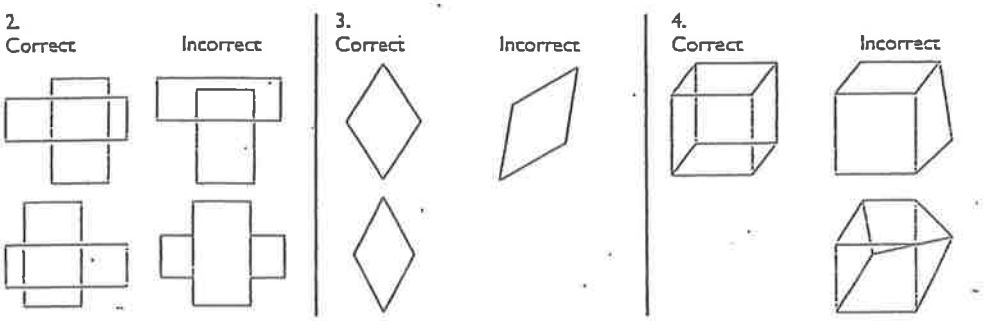
Instruct the subject as follows:

> ***"On this piece of paper is a shape. Try to draw another one that looks just like this, somewhere on the page".***

Allow the subject two attempts for each shape, and permit the subject to erase. If the subject cannot reproduce the figure in two attempts, the tester should go on to the next item.

A drawing should be scored as correct if the subject has reproduced all of the essential geometric features of the original. Changes in size do not count as errors. Small gaps between lines do not indicate an error, as long as the shape has been reproduced. Scoring criteria for each form (examples shown below):

1. **Circle.** A closed curved figure.
2. **Two overlapping rectangles.** Forms must be four-sided, and overlap must be similar to presented form. Changes in size are not scored.
3. **Diamond.** Figure must be four-sided, oriented so that points are at the top and bottom, and the sides are approximately equal length.
4. **Cube.** The form is three-dimensional, with front face in the correct orientation, internal lines drawn correctly between corners. Opposite sides of faces should be approximately parallel.



	Correct	Incorrect (or not drawn)
Circle	☐	☐
Two overlapping rectangles	☐	☐
Diamond	☐	☐
Cube	☐	☐

Score:

0 all 4 drawings correct
1 1 form drawn incorrectly
2 2 forms drawn incorrectly
3 3 forms drawn incorrectly
4 4 forms drawn incorrectly
5 no figures drawn, scribbles; parts of forms; words instead of forms

Score ☐
(maximum 5)

5 IDEATIONAL PRAXIS

Instruct the subject as follows:

> *"I want you to pretend that you have written yourself a letter. Take this piece of paper, fold it so that it will fit into the envelope, and then put it into the envelope. Then seal the envelope, address the envelope to yourself, and show me where the stamp goes".*

Indicate each underlined step completed correctly or incorrectly. If the subject forgets part of the task, or is having difficulty, the tester should repeat the instruction for the component of the task where the subject is having difficulty.

	Correct	Incorrect (or not done)
Fold a letter	☐	☐
Put the letter in an envelope	☐	☐
Seal the envelope	☐	☐
Address the envelope	☐	☐
Indicate where the stamp goes	☐	☐

Score ☐
(maximum 5)

Score:
- 0 all components performed correctly
- 1 failure to perform 1 component
- 2 failure to perform 2 components
- 3 failure to perform 3 components
- 4 failure to perform 4 components
- 5 failure to perform 5 components

6 ORIENTATION

Before testing for orientation, the tester should be sure that no clocks, watches, or calendars are visible to aid the subject. Indicate each item answered correctly or incorrectly.

	Correct	Incorrect (or not answered)		Correct	Incorrect (or not answered)
Full name	☐	☐	Year	☐	☐
Day	☐	☐	Season	☐	☐
Date	☐	☐	Time of day	☐	☐
Month	☐	☐	Place	☐	☐

Score = 1 point is given for each incorrect response

Score ☐
(maximum 8)

Acceptable answers include ± 1 day for the date, naming of upcoming season within 1 week before its onset or name of previous season for 2 weeks after its termination, within 1 hour for the time, and partial name for place. First and last names, day of the week, month, and year must be exact.

7 WORD RECOGNITION

The tester instructs the subject as follows:

> ***"I am going to show you some words printed on cards. I want you to read each word out loud and try to remember it".***

If the subject cannot read a word, the tester says the word out loud. However, it is important for the subject to actually look at each word and try to read it.

At the end of the learning portion of a trial, the tester instructs the subject as follows

> ***"Now I'm going to show you another set of words. Some of the words were on the list I just showed you, and others are new. For each word, I want you to tell me whether it is one of the words I just showed you".***

Then the tester shows the first word and says either

> ***"Is this one of the words I showed you before, yes or no?" or "Did I show you this word before?"***

The same instruction is given before the second test word.

For the remaining test words, the tester should say,

> ***"How about this one?"***

If the subject does not remember the task (e.g., reads the word rather than responding "Yes" or "No"), then the tester should repeat or rephrase the entire question and make a note that the subject had to be reminded of the task instructions. Trials two and three are similar, and the tester should keep track of the number of times the subject has to be reminded of the task instructions.

Score = total number of incorrect responses or "12", whichever is smaller

Trial 1: Score	☐	Reminders	☐
Trial 2: Score	☐	Reminders	☐
Trial 3: Score	☐	Reminders	☐

Score (mean number of incorrect responses for three trials) (maximum 12)	☐	Total Reminders (for scoring item 8)	☐

7 WORD RECOGNITION

Bold words are the words shown before. Italicised words are the words that the subject has not seen. Tick the subject's responses; circles = incorrect responses.

TRIAL 1	Yes/ Shown Before	No/ New	Rem
Corn	○	☐	☐
Effort	⊙	☐	☐
Party	○	☐	☐
River	☐	○	☐
Folly	○	☐	☐
Locker	○	☐	☐
Event	☐	○	☐
Queen	☐	○	☐
Position	☐	○	☐
Quality	○	☐	☐
Sunset	○	☐	☐
Dove	☐	○	☐
Belief	☐	○	☐
Umbrella	☐	○	☐
Allegory	○	☐	☐
Hound	○	☐	☐
Idiom	○	☐	☐
Hint	☐	○	☐
Missile	☐	○	☐
Gem	○	☐	☐
Proxy	☐	○	☐
Lobster	☐	○	☐
Criterion	☐	○	☐
Deceit	○	☐	☐

Total circles ticked (incorrect responses) ☐

TRIAL 2	Yes/ Shown Before	No/ New	Rem
River	☐	○	☐
Officer	○	☐	☐
Thought	○	☐	☐
Event	☐	○	☐
Queen	☐	○	☐
Position	☐	○	☐
Camp	○	☐	☐
Fate	○	☐	☐
Golf	○	☐	☐
Dove	☐	○	☐
Belief	☐	○	☐
Permission	○	☐	☐
Umbrella	☐	○	☐
Hint	☐	○	☐
Missile	☐	○	☐
Blister	○	☐	☐
Concept	○	☐	☐
Proxy	☐	○	☐
Pianist	○	☐	☐
Lobster	☐	○	☐
Gender	○	☐	☐
Criterion	☐	○	☐
Bullet	○	☐	☐
Intellect	○	☐	☐

Total circles ticked (incorrect responses) ☐

TRIAL 3	Yes/ Shown Before	No/ New	Rem
Plant	○	☐	☐
River	☐	○	☐
Amount	○	☐	☐
Event	☐	○	☐
Queen	☐	○	☐
Industry	○	☐	☐
Position	☐	○	☐
Occasion	○	☐	☐
Dove	☐	○	☐
Cradle	○	☐	☐
Banality	○	☐	☐
Singer	○	☐	☐
Belief	☐	○	☐
Umbrella	☐	○	☐
Hypothesis	○	☐	☐
Hint	☐	○	☐
Missile	☐	○	☐
Proxy	☐	○	☐
Noose	○	☐	☐
Distinction	○	☐	☐
Lobster	☐	○	☐
Tank	○	☐	☐
Criterion	☐	○	☐
Decree	○	☐	☐

Total circles ticked (incorrect responses) ☐

Score ☐ (mean number of incorrect responses for three trials.) (maximum = 12, score 12 for more than 12 incorrect responses.)

Total number of reminders (for scoring item 8) ☐

Rem = reminded of instructions

8 REMEMBERING TEST INSTRUCTIONS

Evaluate the subject's ability to remember the requirements of the word recognition task (item 7), based upon noting each instance of failure to remember the test instructions.

Score: 0 = subject never needs extra reminders of instructions

1 = very mild – forgets once

2 = mild – must be reminded 2 times

3 = moderate – must be reminded 3 or 4 times

4 = moderately severe – must be reminded 5 or 6 times

5 = severe – must be reminded 7 or more times

Score ☐

9 SPOKEN LANGUAGE ABILITY

Provide a global rating of the quality of speech, i.e., clarity, difficulty in making oneself understood.

Score: 0 = no instance where it is difficult to understand the subject

1 = very mild – one instance of lack of understandability

2 = mild – subject has difficulty less than 25% of time

3 = moderate – subject has difficulty 25-50% of time

4 = moderately severe – subject has difficulty more than 50% of time

5 = severe – one or two word utterance; fluent, but empty speech; mute Score ☐

10 WORD-FINDING DIFFICULTY IN SPONTANEOUS SPEECH

Rate the subject's difficulty in finding desired words, e.g., circumlocutions.

Score: 0 = no evidence of word-finding difficulty in spontaneous speech

1 = very mild – 1 or 2 instances, not clinically significant

2 = mild – noticeable circumlocution or synonym substitution

3 = moderate – loss of words without compensation on occasion

4 = moderately severe – frequent loss of words without compensation

5 = severe – nearly total loss of content words; speech sounds empty; 1-2 words utterances

☐

11 COMPREHENSION

Rate the subject's ability to understand speech. Do not include responses to commands.

Score: 0 = no evidence of poor comprehension

1 = very mild – 1-2 instances of misunderstanding

2 = mild – 3-5 instances of misunderstanding

3 = moderate – requires several repetitions and rephrasing

4 = moderately severe – subject only occasionally responds correctly; e.g., yes/no questions

5 = severe – subject rarely responds to questions appropriately, not due to poverty of speech

Score ☐

12 CONCENTRATION/DISTRACTIBILITY

Rate the frequency with which the subject is distracted by irrelevant stimuli and/or must be reoriented to the ongoing task because the subject has lost his/her train of thought or appears to be caught up in his/her own thoughts.

Score: 0 = no evidence of poor concentration or distractibility

1 = very mild – one instance of poor concentration

2 = mild – 2-3 instances of poor concentration/distractibility; signs of restlessness and inattentiveness

3 = moderate – 4-5 instances during interview

4 = moderately severe – poor concentration/distractibility throughout much of interview

5=severe – extreme difficulty in concentration and extremely distractible, unable to complete tasks

Score ☐

Appendix B (Part A): Full version of the composite activities of daily living questionnaire

The point (pt) allocation system has been provided next to each item.

This aim of this questionnaire is to assess to what extent your day-to-day activities are independent of another's assistance, or similarly, to what extent your activities have been restricted or limited from what they used to be.

Please tick one box per section next to the sentence that you feel best describes how you go about your daily living activities. Remember, the statements refer to what you are currently able to do, not what you have previously done or would like to do.

Food Preparation

Able to select, plan, prepare and serve meals independently, as required	☐ *4 pts*
Able to prepare food if ingredients supplied/set out	☐ *3 pts*
Unable to cook a meal, but capable of making snacks and reheating food	☐ *2 pts*
Can prepare food if prompted step by step	☐ *1 pt*
Need to have meals prepared and served	☐ *0 pt*

Eating

Able to eat without assistance, using correct cutlery	☐ *3 pts*
Able to eat without assistance provided food is made manageable (that is, food is a particular form, consistency, or size)	☐ *2 pts*
Find it necessary to eat food with fingers	☐ *1 pt*
Need to be fed	☐ *0 pt*

Drink Preparation

Able to select and prepare drinks as required	☐ *3 pts*
Can prepare drinks if ingredients left available	☐ *2 pts*
Can prepare drinks if prompted step by step	☐ *1 pt*
Unable to make a drink even with prompting and supervision	☐ *0 pt*

Drinking

Able to drink without any problems, and from an unmodified glass or mug ☐ *4 pts*

Require aids to drink

[e.g. specific type of cup (***3 pts***), use a straw (***2 pts***)]

Please specify .....

Have difficulty drinking, even with aids ☐ *1 pt*

Require drinks to be administered ☐ *0 pt*

Dressing

Able to select suitable clothing and can dress/undress self unassisted ☐ *3 pts*

Can dress/undress self, but sometimes put clothes on/take clothes off in the wrong order and/or back to front ☐ *2 pts*

Unable to dress/undress self but move limbs to assist ☐ *1 pt*

Require total dressing/undressing (unable to assist) ☐ *0 pt*

Bathing

Able to bathe self (in tub, shower, sponge bath) regularly and without help ☐ *5 pts*

Able to bathe self with help getting in and out of tub/shower ☐ *4 pts*

Need bath/shower to be drawn/turned on, but wash independently ☐ *3 pts*

Can wash face and hands only, cannot bathe rest of body ☐ *2 pts*

Can wash self if prompted and supervised ☐ *1 pt*

Unable to wash self and need full assistance ☐ *0 pt*

Teeth

Able to clean own teeth/dentures regularly and independently ☐ *3 pts*

Can clean teeth/dentures if given appropriate items ☐ *2 pts*

Require some assistance, toothpaste on brush, brush to mouth etc. ☐ *1 pt*

Need full assistance ☐ *0 pt*

Toilet

Can manage independently whilst at toilet, no incontinence ☐ *4 pts*

Need to be reminded, or need help in cleaning self,
or have rare (weekly at most) accidents ☐ *3 pts*

Soil or wet self while asleep (occurs more than once a week) ☐ *2 pts*

Soil or wet self while awake (occurs more than once a week) ☐ *1 pt*

Have no control of bowels or bladder ☐ *0 pt*

Grooming (neatness, hair, nails, hands, face, clothing)

Always well-groomed, without assistance ☐ *3 pts*

Able to groom self with occasional minor assistance (e.g shaving, lipstick) ☐ *2 pts*

Need regular assistance or supervision in grooming ☐ *1 pt*

Require total grooming care, but remain well-groomed thereafter ☐ *0 pt*

Transfers

Able to get in/out of most chairs unaided ☐ *3 pts*

Can get into most chairs but need help to get out ☐ *2 pts*

Need help getting in and out of most chairs ☐ *1 pt*

Totally dependent on being put into and lifted from most chairs ☐ *0 pt*

Mobility

Able to walk independently ☐ *4 pts*

Walk independently with occasional assistance
(e.g. use furniture, or arm for support) ☐ *3 pts*

Use aids to walk
[e.g. a walking frame (***1 pt***), or stick (***2 pts***)]
Please specify .....

Unable to walk ☐ *0 pt*

Mode of Transportation

* Remember, what is it that you are able to do.

Drive a car

- Independently ☐ *7 pts*
- With help from a passenger ☐ *6 pts*

Travel on public transport (eg. bus, train, or tram)

- Independently ☐ *5 pts*
- When accompanied by another ☐ *4 pts*

Travel by taxi

- Travel and organise independently ☐ *3 pts*
- Travel only if organised by someone else ☐ *2 pts*

Travel when transport is provided by a family member or friend ☐ *1 pt*

Do not travel at all ☐ *0 pt*

Shopping

Able to take care of all shopping needs independently ☐ *4 pts*

Able to take car of most shopping needs independently ☐ *3 pts*

Shop independently for 1 or 2 items (small purchases), with or without a list ☐ *2 pts*

Unable to shop alone, but participate when accompanied ☐ *1 pt*

Completely unable to shop ☐ *0 pt*

Communication

Able to hold appropriate conversation (listen and respond at correct times) ☐ *3 pts*

Show understanding and attempt to respond verbally with gestures ☐ *2 pts*

Can make self understood but have difficulty understanding others ☐ *1 pt*

Do not respond to or communicate with others ☐ *0 pt*

Telephone

Able to operate telephone on own initiative (look up and dial numbers etc) ☐ *4 pts*

Dial a few well-known numbers ☐ *3 pts*

Use telephone if number given verbally/visually or predialled ☐ *2 pts*

Answer telephone but do not make calls ☐ *1 pt*

Unable to use telephone at all ☐ *0 pt*

Housekeeping

Maintain house alone or with occasional assistance ☐ *2 pts*

(e.g. 'heavy work domestic help')

Perform light daily tasks such as dish-washing, bedmaking ☐ *1 pt*

Need help with all home maintenance tasks ☐ *0 pt*

Not applicable ☐ -

Gardening

Able to do gardening without assistance ☐ *3 pts*

Can garden but often require assistance ☐ *2 pts*

Garden infrequently, even with lots of assistance ☐ *1 pt*

Do not participate in any gardening ☐ *0 pt*

Not applicable ☐ -

Laundry

Able to do personal laundry completely ☐ *2 pts*

Launder small items (e.g. rinse stockings or socks, etc) ☐ *1 pt*

All laundry must be done by others ☐ *0 pt*

Not applicable ☐ -

Responsibility for own medications

Responsible for taking medication in correct dosages at correct time ☐ *2 pts*

Responsible for medication if it is prepared in advance in separate dosages ☐ *1 pt*

Unable to dispense own medication at all ☐ *0 pt*

Not applicable ☐ -

Ability to handle finances

Able to manage financial matters independently ☐ *4 pts*

Able to manage most financial matters, have problems with specific areas
[e.g. budget, write cheques, pay rent/bills (***3 pts***), go to bank, automatic tellers (***2 pts***)]
Please specify .....

Able to make decisions regarding financial matters, but need someone else to physically arrange the finances. ☐ *1 pt*

Unable to handle money or recognise money values ☐ *0 pt*

Not applicable ☐ -

Games/Hobbies

Able to fully participate in previous pastimes/activities ☐ *2 pts*

Can participate but need instruction/supervision to do so ☐ *1 pt*

No longer able to join in ☐ *0 pt*

Not applicable ☐ -

Appendix B (Part B): Abridged version of the composite activities of daily living questionnaire

The point (pt) allocation system has been provided next to each item.

This aim of this questionnaire is to assess to what extent your day-to-day activities are independent of another's assistance, or similarly, to what extent your activities have been restricted or limited from what they used to be.

Please tick one box per section next to the sentence that you feel best describes how you go about your daily living activities. Remember, the statements refer to what you are currently able to do, not what you have previously done or would like to do.

Food Preparation

Able to select, plan, prepare and serve meals independently, as required	☐ *4 pts*
Able to prepare food if ingredients supplied/set out	☐ *3 pts*
Unable to cook a meal, but capable of making snacks and reheating food	☐ *2 pts*
Can prepare food if prompted step by step	☐ *1 pt*
Need to have meals prepared and served	☐ *0 pt*

Transfers

Able to get in/out of most chairs unaided	☐ *3 pts*
Can get into most chairs but need help to get out	☐ *2 pts*
Need help getting in and out of most chairs	☐ *1 pt*
Totally dependent on being put into and lifted from most chairs	☐ *0 pt*

Mobility

Able to walk independently	☐ *4 pts*
Walk independently with occasional assistance (e.g. use furniture, or arm for support)	☐ *3 pts*
Use aids to walk [e.g. a walking frame (***1 pt***), or stick (***2 pts***)] Please specify	
Unable to walk	☐ *0 pt*

Mode of Transportation

* Remember, what is it that you are able to do.

Drive a car

Independently	☐ *7 pts*
With help from a passenger	☐ *6 pts*

Travel on public transport (eg. bus, train, or tram)

Independently	☐ *5 pts*
When accompanied by another	☐ *4 pts*

Travel by taxi

Travel and organise independently	☐ *3 pts*
Travel only if organised by someone else	☐ *2 pts*

Travel when transport is provided by a family member or friend ☐ *1 pt*

Do not travel at all ☐ *0 pt*

Shopping

Able to take care of all shopping needs independently	☐ *4 pts*
Able to take car of most shopping needs independently	☐ *3 pts*
Shop independently for 1 or 2 items (small purchases), with or without a list	☐ *2 pts*
Unable to shop alone, but participate when accompanied	☐ *1 pt*
Completely unable to shop	☐ *0 pt*

Telephone

Able to operate telephone on own initiative (look up and dial numbers etc)	☐ *4 pts*
Dial a few well-known numbers	☐ *3 pts*
Use telephone if number given verbally/visually or predialled	☐ *2 pts*
Answer telephone but do not make calls	☐ *1 pt*
Unable to use telephone at all	☐ *0 pt*

Housekeeping

Maintain house alone or with occasional assistance ☐ *2 pts*

(e.g. 'heavy work domestic help')

Perform light daily tasks such as dish-washing, bedmaking ☐ *1 pt*

Need help with all home maintenance tasks ☐ *0 pt*

Not applicable ☐ -

Laundry

Able to do personal laundry completely ☐ *2 pts*

Launder small items (e.g. rinse stockings or socks, etc) ☐ *1 pt*

All laundry must be done by others ☐ *0 pt*

Not applicable ☐ -

Ability to handle finances

Able to manage financial matters independently ☐ *4 pts*

Able to manage most financial matters, have problems with specific areas

[e.g. budget, write cheques, pay rent/bills (***3 pts***), go to bank, automatic tellers (***2 pts***)]

Please specify .....

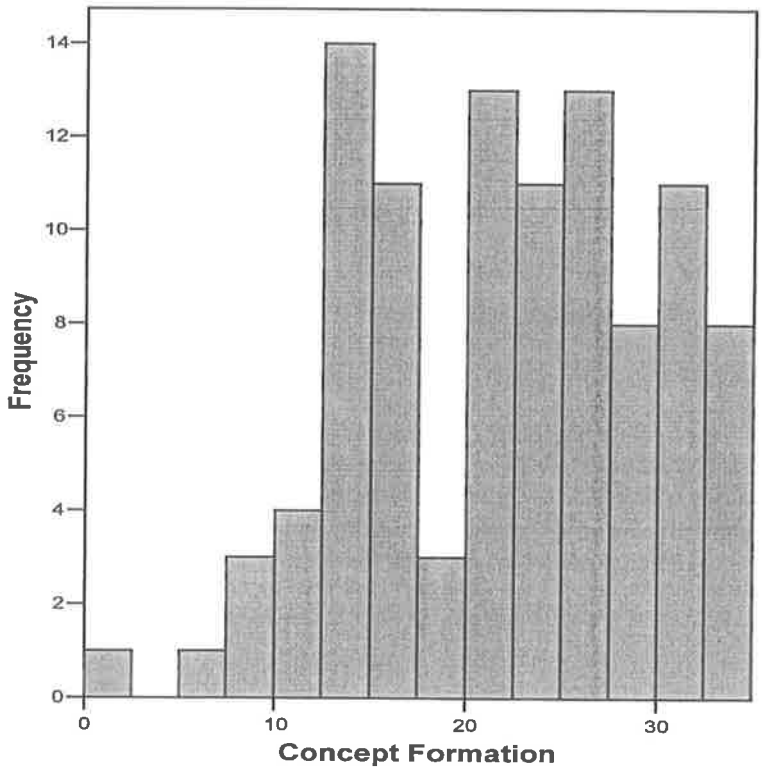
Able to make decisions regarding financial matters, but need someone else to physically arrange the finances. ☐ *1 pt*

Unable to handle money or recognise money values ☐ *0 pt*

Not applicable ☐ -

Appendix C: Performance distribution of Concept Formation on the initial measurement occasion.

The distribution of Concept formation was bi-modal, containing two high points; one at a score of approximately 14, representing individuals who did not have a strong understanding of this task, and the other at a score of around 25, representing individuals who had a better understanding of this task. The maximum score achievable for Concept Formation is 35; two participants achieved this score.



Appendix D: Questions from the Information test used in the current study

There are 40 questions and the *correct answer* is in italics.

1. What do we call a baby cow?
   - A Bull
   - *B Calf*
   - C Foal
   - D Piglet

2. How many things make a dozen?
   - *A Twelve*
   - B Six
   - C Eggs
   - D Ten

3. Who was Captain Cook?
   - A A prime minister
   - *B An explorer*
   - C An inventor
   - D A cook

4. Name two kinds of coins
   - A 5 Dollars and 10 Dollars
   - B Money and Dollars
   - *C 20 cents and 50 cents*
   - D Indian and corn

5. On what continent is China?
   - *A Asia*
   - B South Africa
   - C South America
   - D Europe

6. Which month has one extra day every four years?
   - *A February*
   - B January
   - C May
   - D December

7. What is the capital of Greece?
   - A Rome
   - *B Athens*
   - C Crete
   - D Cairo

8. How is oxygen returned to the air?
   - A By breathing
   - *B By plants*
   - C By the wind
   - D By clouds

9. What is water made of?
   - A Minerals and chemicals
   - B Rain
   - C Helium and Oxygen
   - *D Hydrogen and Oxygen*

10. What are hieroglyphics?
   - A Ancient Greek letters
   - B Roman numerals
   - *C Egyptian picture writing*
   - D Cave drawings

11. What country has the largest population?
   - A India
   - B Russia
   - C North America
   - *D China*

12. What is the main material used to make glass?
   - *A Sand*
   - B Plastic
   - C Hydrogen
   - D Fibreglass

13. In what direction does the sun set?
   - A North
   - B East
   - C South
   - *D West*

14. Who invented the electric light bulb?
   - A Albert Einstein
   - *B Thomas Edison*
   - C Benjamin Franklin
   - D Thomas Jefferson

15. Who wrote Hamlet?
   - A William Tell
   - B Mark Twain
   - C Ernest Hemingway
   - *D William Shakespeare*

16. Who was the prime minister of England during the Second World War?
    - *A* *Winston Churchill*
    - B Stanley Baldwin
    - C Margaret Thatcher
    - D Clement Atlee

17. In what country did the Olympic Games originate?
    - A Egypt
    - *B* *Greece*
    - C Rome
    - D Italy

18. What is a barometer?
    - *A* *It measures air pressure*
    - B It measures wind speed
    - C It measures rainfall
    - D It measures earthquakes

19. On what continent is the Sahara Desert?
    - *A* *Africa*
    - B Europe
    - C Arabia
    - D Asia

20. Who was Anne Frank?
    - A A singer
    - B A pilot
    - *C* *A girl who wrote a diary*
    - D A teacher of deaf and blind

21. Who was Charles Darwin?
    - A He was a poet
    - *B* *He developed the theory of evolution*
    - C He was a character in a Dickens novel
    - D He discovered the structure of DNA

22. Who painted the Sistine Chapel?
    - A Botticelli
    - B Da Vinci
    - C Raphael
    - *D* *Michelangelo*

23. How far is it from London to Sydney (approximately)?
    - A 500 kilometres
    - B 7,000 kilometres
    - *C* *17,000 kilometres*
    - D 40,000 kilometres

24. Who was Mahatma Gandhi?
    - A An Indian prince
    - B A cricket player
    - *C An Indian independence leader*
    - D A Buddhist monk

25. Whose name is usually associated with the theory of relativity?
    - A Planck
    - B Newton
    - C Watson
    - *D Einstein*

26. What causes iron to rust?
    - A Acid
    - B Salt
    - *C Oxygen*
    - D Minerals

27. Name three kinds of blood vessels in the human body
    - A Pulmonary, capillary and aorta
    - *B Artery, vein and capillary*
    - C Artery, aorta and vein
    - D Capillary, jugular and vein

28. Vision problems are most often caused by a deficiency in
    - *A Vitamin A*
    - B Vitamin B
    - C Vitamin C
    - D Vitamin D

29. Who was Catherine the Great?
    - A A Roman Empress
    - B A French Queen
    - *C A Russian Empress*
    - D An Egyptian Queen

30. Which is the closest planet to our Sun?
    - A Mars
    - *B Mercury*
    - C Earth
    - D Venus

31. What is the world population (approximately)?
    - A 4 billion
    - *B 6 billion*
    - C 8 billion
    - D 10 billion

32. What is the capital of Sri Lanka?
    - A Sinhal
    - *B Colombo*
    - C Tamil
    - D Matale

33. What is the speed of light (approximately)?
    - A *300,000 km/sec*
    - B 258,000 km/sec
    - C 362,000 km/sec
    - D 524,000 km/sec

34. What is Marie Curie famous for?
    - *A She was a physicist*
    - B She was a missionary
    - C She was a medical doctor
    - D She was a biologist

35. What does turpentine come from?
    - A Ethyl alcohol
    - B Varnish
    - C Acid
    - *D Pine trees*

36. Who wrote Faust?
    - A Mann
    - B Hesse
    - C Nietzsche
    - *D Goethe*

37. What does the musical term "piano" mean?
    - A To be played evenly
    - B To be played fast
    - *C To be played softly*
    - D To be played smoothly

38. How far above sea level is Mount Everest?
    - A 7,448 metres
    - B 7,984 metres
    - C 9,298 metres
    - *D 8,848 metres*

39. Who was the greek muse of history?
    - A Urania
    - B Thalia
    - C Polymnia
    - *D Clio*

40. From what language does the word "Ombudsman" originate?

A French
B *Swedish*
C German
D Dutch

Appendix E: Copy of the self-report Health Questionnaire

This is a general health questionnaire, which aims to gain an overall picture of you health status. It contains 4 sections that we would like you to complete as accurately as possible.

SECTION 1: GENERAL HEALTH

1. How would you rate your health at the present time compared with other people of your age? (Please tick one)

   □ Poor
   □ Fair
   □ Average
   □ Good
   □ Excellent

2. How would you rate your health at the present time compared with a state of perfect health? (Please tick one)

   □ Poor
   □ Fair
   □ Average
   □ Good
   □ Excellent

3. Do you have problems with your vision? □ Yes No□

4. Do you wear glasses? □ Yes No□

5. Have you been diagnosed with any of the following vision disorders?

Nearsightedness (Myopia)	□ Yes	No/ not sure□
Farsightedness (Hyperopia)	□ Yes	No/ not sure□
Glaucoma	Yes□	No/ not sure□
Cataracts	Yes□	No/ not sure□
Astigmatism (cornea problem)	□ Yes	No/ not sure□

Other (please specify)

.....

6. How many visits to the doctor have you made in the last 6 months?

None	☐
1 - 3	☐
4 – 6	☐
7 – 9	☐
10 +	☐

7. How many hospitalisations have you had in the past 6 months? .....

8. Have you ever had a general anaesthetic? ☐ Yes No☐

9. If Yes, how many have you had? .....

10. Do you have any ongoing medical conditions (eg. Asthma) and do you take any medication for these conditions?

MEDICAL CONDITION	MEDICATION
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....
.....	.....

10. Do you have any acute health problems at the moment (eg. a cold) and are you taking any medication from these problems?

MEDICAL CONDITION	MEDICATION
.....	.....
.....	.....
.....	.....

11. Are you currently taking any vitamins, minerals, or other dietary supplements? (Please specify).

.....

.....

SECTION 2: SYMPTOMS

Below is a list of symptoms that you may or may not have experienced in the last six months. Please read the symptom and indicate by a **circle** to what degree it bothers you. (If you have never suffered from a particular symptom please circle Never)

	Never	A little	Quite a bit	Extremely
1. Feeling nauseous	1	2	3	4
2. Headaches from reading	1	2	3	4
3. Diarrhoea or constipation	1	2	3	4
4. Toothaches	1	2	3	4
5. Heavy feeling in arms or legs	1	2	3	4
6. Spitting up blood	1	2	3	4
7. Bladder infection problems	1	2	3	4
8. Frequent tonsillitis	1	2	3	4
9. Dry mouth	1	2	3	4
10. Pain over sinuses	1	2	3	4
11. Painful nose	1	2	3	4
12. Blurred vision	1	2	3	4

	Never	A little	Quite a bit	Extremely
13. Heart beating quickly or strongly	1	2	3	4
14. Excessive gas	1	2	3	4
15. Loss of voice	1	2	3	4
16. Numbness in hands or feet	1	2	3	4
17. Twitching muscles	1	2	3	4
18. Swelling of body parts	1	2	3	4
19. Feeling drowsy	1	2	3	4
21. Broken hip	1	2	3	4
22. Sore throat	1	2	3	4
23. Earaches	1	2	3	4
24. Upset stomach	1	2	3	4
25. Ear drainage problems	1	2	3	4
26. Aches in back or neck and skull	1	2	3	4
27. Difficulty in chewing	1	2	3	4
28. Snoring	1	2	3	4
29. Lump in throat	1	2	3	4
30. Muscle pain, aches or stiff joints	1	2	3	4
31. Tinnitus (ringing in the ears)	1	2	3	4
32. Drainage down back of throat	1	2	3	4
33. Blood in urine	1	2	3	4
34. Chest infections	1	2	3	4
35. Tingling or pins and needles in hands and feet	1	2	3	4

	Never	A little	Quite a bit	Extremely
36. Inflammation of the eyes	1	2	3	4
37. Headaches or migraine headaches	1	2	3	4
38. Feeling dizzy or faint	1	2	3	4
39. Heartburn	1	2	3	4
40. Persistent cough	1	2	3	4
41. Sore tongue	1	2	3	4
42. High blood pressure	1	2	3	4
43. Difficulty breathing through nose	1	2	3	4
44. Poor appetite	1	2	3	4
45. Chest pains	1	2	3	4
46. Blind spots before eyes	1	2	3	4
47. Problems with urine control	1	2	3	4
48. Loss of hearing	1	2	3	4
49. Cold hands and feet	1	2	3	4
50. Shortness of breath	1	2	3	4
51. Pain on urination	1	2	3	4
52. Varicose veins	1	2	3	4

SECTION 3: HEALTH HABITS

1. Have you ever smoked cigarettes? Yes☐ ☐ No
   (if No, go to question 5)
2. Do you currently smoke cigarettes? Yes☐ No☐
3. If Yes, how many cigarettes do you smoke per day? .....
4. How many years have you or did you smoke for .....
5. Do you drink alcohol regularly (eg. once a week)? Yes☐ No☐

6. If Yes, how many standard drinks per week? .....
(1 standard drink = 1 glass of wine 150ml).

7. How many times per week do you exercise? .....

8. On average, how long do you exercise for each time .....

9. What type of exercise do you do? .....

SECTION 4: DISEASE HISTORY

Below are listed a number of diseases some of which you may have suffered from in your lifetime. Please indicate by ticking the YES or NO box whether you have suffered from any of these diseases. If you tick YES, and the disease is ongoing please tick the "ongoing" box. In either case, please indicate at what age you first suffered from it.

1.	High blood pressure	Yes☐ No☐	Ongoing☐	At what age? .....
2.	Arthritis	Yes☐ No☐	Ongoing☐	At what age? .....
3.	Heart disease	Yes☐ No☐	Ongoing☐	At what age? .....
4.	Kidney disease	Yes☐ No☐	Ongoing☐	At what age? .....
5.	Anaemia	Yes☐ No☐	Ongoing☐	At what age? .....
6.	Cancer	Yes☐ No☐	☐ Ongoing	At what age? .....
7.	Asthma	Yes☐ No☐	☐ Ongoing	At what age? .....

8. Diabetes, Type I	Yes☐ No☐	Ongoing☐	At what age? .....
9. Diabetes, Type II	Yes☐ No☐	Ongoing☐	At what age? .....
10. Liver disease	Yes☐ No☐	Ongoing☐	At what age? .....
11. Peptic ulcer	Yes☐ No☐	Ongoing☐	At what age? .....
12. Stroke	Yes☐ No☐	☐ Ongoing	At what age? .....
13. Thyroid disorder	Yes☐ No☐	Ongoing☐	At what age? .....
14. Epilepsy	Yes☐ No☐	Ongoing☐	At what age? .....
15. Multiple Sclerosis	Yes☐ No☐	Ongoing☐	At what age? .....
16. Parkinson's disease	Yes☐ No☐	Ongoing☐	At what age? .....
17. Pneumonia	Yes☐ No☐	Ongoing☐	At what age? .....

Appendix F: General Health Status

Participants were asked to select from five options (1 = poor, 2 = fair, 3 = average, 4 = good, 5 = excellent), the one that best described their health relative to other people their own age. Figure F1 illustrates that participants generally considered their health to be good and no one felt that their health was poor, compared to others of a similar age. A slightly less optimistic view was held when participants compared their health to the ideal of perfect health (Figure F2, p.305); but an average or good response was most common.

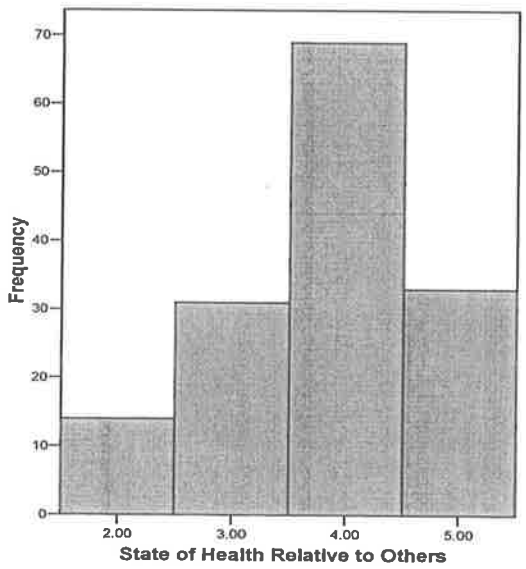


Figure F1 Participants' rating of their health compared to other people of a similar age

No one rated their health as poor (1) and the majority of participants rated their health as good (4).

Participants reported an average of one vision problem (mean = 1.12, SD = 1.03); the most a participant reported was four. The most common vision problem reported was cataracts (see Table F1, p.305). The category of "other" eye problems permitted participants to indicate other specific eye complaints. The most common response to this was macular degeneration (N = 5).

In terms of health care utilisation, most participants had visited their doctor between one and three times in the previous 6 months (see Table F2, p.305) and

participants had virtually no hospitalisations (mean = .28, SD = .56) within the same timeframe. Throughout their lifetime, participants reported an average of 5.02 general anaesthetics (SD = 4.81; range = 0 to 35).

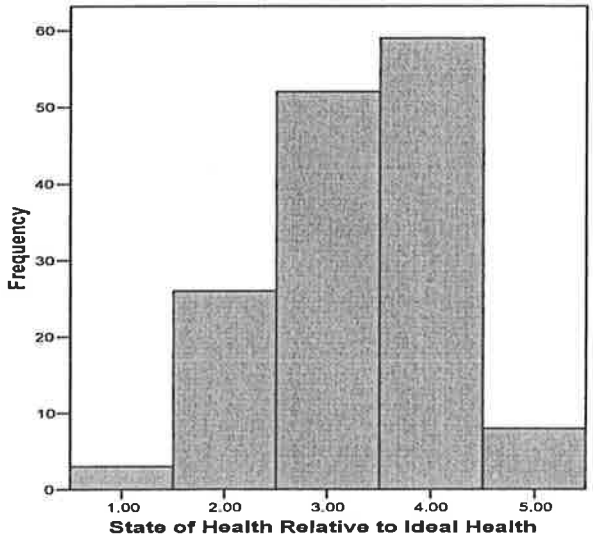


Figure F2 Participants' rating of their health compared to the ideal of perfect health

Few participants rated their health as either poor or excellent compared to ideal health.

Table F1

Prevalence of six types of vision problems in the current sample

Type of Vision Problem	N
Nearsightedness	31
Farsightedness	16
Glaucoma	10
Cataracts	57
Astigmatism	28
Other	22

Table F2

Frequency of visits to the doctor over the last 6 months

Number of visits	N
0	5
1 - 3	57
4 - 6	44
7 - 9	20
10+	18

The majority of participants (76%) indicated that they had not suffered from any acute illnesses (e.g. cold or flu) over the previous 6 months. As Table F3 (p.306) illustrates, chronic illnesses were the dominant health problem for the current elderly

population. Finally, on average, participants consumed three prescription medications (mean = 3.03, SD = 2.62).

Table F3
General health status: presence of acute and chronic diseases

Disease category	N
Neither	15
Acute disease only	2
Chronic disease only	104
Chronic and acute disease	22

In conclusion, the aforementioned statistics demonstrate that, on the whole, participants were very healthy; and health problems other than chronic diseases were minimal. In terms of the relationship between general health measures and biomarkers, few variables were significantly correlated. The exceptions were Reading Span and total number of vision problems and body mass index and number of hospitalisations (see Table F4, p.307). Correlations between general health measures and outcome variables were stronger and more numerous than for the biomarkers (see Table F5, p.307). However, the pattern of significant correlations was different for biomarkers and outcome measures. Consequently, general health measures did not have to be included as covariates in subsequent analyses.

Table F4
Correlations (Pearson's r) between general health measures and biomarkers

Biomarker variable	General health measure					
	N(a)	Total number of vision problems	Range of doctors visits	Number of Hospitalisations	Number of anaesthetics	Number of medications
Reading Span	119	.19*	-.15	-.11	-.12	-.14
Systlolic BP	109	.06	-.07	-.06	.18	.05
Diastolic BP	109	.04	-.15	.04	.13	-.14
Grip Strength	119	-.12	-.13	-.02	-.15	-.14
Body Mass Index	119	-.02	-.08	-.20*	.13	.13
Visual Acuity	119	-.08	-.03	-.06	.07	.10

*p < .05
N(a) = Listwise for general health measures and each biomarker
BP = Blood Pressure

Table F5
Correlations (Pearson's r) between general health measures and independent functioning outcome measures

Outcome measure	General health measure					
	N(a)	Total number of vision problems	Range of doctors visits	Number of Hospitalisations	Number of anaesthetics	Number of medications
Basic ADL	119	-.12	-.04	.05	-.23*	-.21*
Instrumental ADL	119	.05	-.26**	.02	-.11	-.03
Reasoning ability	119	.00	-.19*	.05	-.01	-.10
Life Satisfaction	117	-.12	-.35**	-.13	-.31**	-.30**

**p < .01; *p < .05
ADL = Activities of daily living

Appendix G (Part A): Body systems represented by symptoms

Symptoms were grouped based on biological theory rather than factor analysis.

Table G1
Summary of disease symptoms included in the health questionnaire in the current study

Name of general area of the body represented	Item number	Symptom name
1. Head/Ears, nose & throat		
Eyes	2.	Headaches from reading
	12.	Blurred vision
	35.	Inflammation of the eyes
	45.	Blind spots before eyes
Ears	22.	Earaches
	24.	Ear drainage problems
	30.	Tinnitus (ringing in the ears)
	47.	Loss of hearing
Nose	10.	Pain over sinuses
	11.	Painful nose
	27.	Snoring
	31.	Drainage down back of throat
	42.	Difficulty breathing through nose

Table G1 (continued)
Summary of disease symptoms included in the health questionnaire in the current study

Name of general area of the body represented	Item number	Symptom name
1. Head/Ears, nose & throat (continued)		
Mouth	4.	Toothaches
	9.	Dry mouth
	26.	Difficulty in chewing
	40.	Sore tongue
	43.	Poor appetite
Throat	8.	Frequent tonsillitis
	15.	Loss of voice
	21.	Sore throat
	28.	Lump in throat
2. Circulatory system	6.	Spitting up blood
	13.	Heart beating quickly or strongly
	18.	Swelling of body parts
	33.	Chest infections
	39.	Persistent cough
	41.	High blood pressure
	48.	Cold hands and feet
	51.	Varicose veins

Table G1 (continued)
Summary of disease symptoms included in the health questionnaire in the current study

Name of general area of the body represented	Item number	Symptom name
3. Gastro-Intestinal Tract and Genito-Urinary Tract		
Stomach and intestines	1.	Feeling nauseous
	3.	Diarrhoea or constipation
	14.	Excessive gas
	23.	Upset stomach
	38.	Heartburn
Genito-Urinary	7.	Bladder infection problems
	32.	Blood in urine
	46.	Problems with urine control
	50.	Pain on urination
4. Musculoskeletal	5.	Heavy feeling in arms or legs
	16.	Numbness in hands or feet
	17.	Twitching muscles
	20.	Broken hip
	25.	Aches in back or neck and skull
	29.	Muscle pain, aches or stiff joints
	34.	Tingling or pins & needles in hands & feet
5. Central Nervous System	19.	Feeling drowsy
	36.	Headaches or migraine headaches
	37.	Feeling dizzy or faint
	44.	Chest pains
	49.	Shortness of breath

Appendix G (Part B): Body systems represented by diseases

Diseases were grouped based on biological theory rather than factor analysis.

Table G2
Summary of disease conditions included in the health questionnaire in the current study

Disease Category	Item number	Name of disease
Cardiovascular	1.	High blood pressure (hypertension)
	3.	Heart disease
	12.	Stroke
Gastrointestinal	10.	Liver disease
	11.	Peptic ulcer
Haematological	5.	Anaemia
Metabolic	8.	Diabetes (Type I)
	9.	Diabetes (Type II)
	13.	Thyroid disorder
Musculoskeletal	2.	Arthritis
Pulmonary	7.	Asthma
	17.	Pneumonia

Table G2 (continued)
Summary of disease conditions included in the health questionnaire in the current study

Disease Category	Item number	Name of disease
Neurologic	14.	Epilepsy
	15.	Multiple sclerosis
	16.	Parkinson's Disease
Miscellaneous	4.	Kidney disease
	6.	Cancer

Appendix H: Diet Diary used to assess nutritional intake

Dear _____

We would like you to write down EVERYTHING you eat and drink into this Food Diary for the days specified at the top of each page.

Please record the AMOUNT that you eat and drink in the first column. For example, 1 glass of milk, 2 teaspoons of sugar, 50gm chips.

We would also like to know the BRAND NAMES AND TYPES of food eaten if possible. For example, 2 slices of Tiptop Multigrain Bread, 1 teaspoon of Meadow Lea margarine, 1 teaspoon of Cottee's Apricot Jam.

It is easy to forget little things like sugar in tea or coffee, a glass of water, and margarine on bread. The best way to avoid this is to FILL IN YOUR DIARY AS SOON AS POSSIBLE AFTER YOU EAT. Avoid leaving it to the end of the day because you may forget things.

We want to get a measure of what you usually eat so please DO NOT CHANGE YOUR EATING HABITS just because they are being recorded.

Finally, we would also like you to record ALL MEDICATIONS that you take over the same time period. Please tell us how many you take (eg. 2 tablets), the name of the medication (eg. Panadol), and the reason (eg. headache).

An example of a filled in FOOD DIARY is presented in the following 3 pages.

Day Wednesday		**Date** 18th March	
Quantity or weight	**Food and Drinks Consumed**	**Type and Brand**	**Preparation or Cooking Method**
	BREAKFAST		
4	Weetbix	Sanitarium	
20 gm	Dried fruit		
½ cup	Milk	Farmers Union	
1 teasp	Sugar		
1 cup	Multi-vitamin juice	Berri	
	MORNING TEA OR SNACK		
1 piece	Cake - chocolate	Home made	
1 cup	Coffee	Nescafé	
1 teasp	Sugar		
	Milk	Farmer's Union	
	LUNCH		
	Sandwich		
2 slices	White bread	Buttercup Wonderwhite	
1 teasp	Margarine	Flora	
1 slice	Cheese	Coon slices	
1 leaf	Lettuce		
3 slices	Tomato		
1 slice	Ham		
600 ml	Coke		

Day Wednesday		**Date** 18th March	
Quantity or weight	**Food and Drinks Consumed**	**Type and Brand**	**Preparation or Cooking Method**
	AFTERNOON TEA OR SNACK		
50 gm	Salt and Vinegar chips	Thins	
1	Apple (small)		
	DINNER/TEA		
	Lasagne		Baked in oven
3 sheets	Lasagne strips	San Remo	
1 cup	Tomato and beef mince sauce	Home made	Fried in oil
½ cup	Cheese sauce	Home made	
	Salad		
3 leaves	Lettuce		
4 slices	Tomato		
4 slices	Cucumber		
1 teasp	Salad dressing	Praise French	
1 glass	White wine	Wynns	
	EVENING SNACK (OR OTHER FOODS)		
200 gm	Strawberry yoghurt	Yoplait	

In the following table please indicate all medications that you took today and any vitamins, minerals or herbal supplements.

Day Wednesday	**Date** 18th March
Quantity	**Medication**
2 tablets	Vitamin C
1 tablet	Panadol (Headache)
2 tablets	Accupril (High blood pressure)

Day		**Date**	
Quantity or weight	**Food and Drinks Consumed**	**Type and Brand**	**Preparation or Cooking Method**
	BREAKFAST		
	MORNING TEA OR SNACK		
	LUNCH		

Day		**Date**	
Quantity or weight	**Food and Drinks Consumed**	**Type and Brand**	**Preparation or Cooking Method**
AFTERNOON TEA OR SNACK			
DINNER/TEA			
EVENING SNACK (OR OTHER FOODS)			

In the following table please indicate all medications that you took today and any vitamins, minerals or herbal supplements.

Day	**Date** .....
Quantity	**Medication**

Day		**Date**	
Quantity or weight	**Food and Drinks Consumed**	**Type and Brand**	**Preparation or Cooking Method**
BREAKFAST			
MORNING TEA OR SNACK			
LUNCH			

Day		**Date**	
Quantity or weight	**Food and Drinks Consumed**	**Type and Brand**	**Preparation or Cooking Method**
AFTERNOON TEA OR SNACK			
DINNER/ TEA			
EVENING SNACK (OR OTHER FOODS)			

In the following table please indicate all medications that you took today and any vitamins, minerals or herbal supplements.

Day	**Date** .....
Quantity	**Medication**

Day		**Date**	
Quantity or weight	**Food and Drinks Consumed**	**Type and Brand**	**Preparation or Cooking Method**
	BREAKFAST		
	MORNING TEA OR SNACK		
	LUNCH		

Day		Date	
Quantity or weight	**Food and Drinks Consumed**	**Type and Brand**	**Preparation or Cooking Method**
AFTERNOON TEA OR SNACK			
DINNER/TEA			
EVENING SNACK (OR OTHER FOODS)			

In the following table please indicate all medications that you took today and any vitamins, minerals or herbal supplements.

Day	**Date** .....
Quantity	**Medication**

Appendix I: Macronutrients and micronutrients

Table I
General categories of both macro- and micronutrients and their major constituents

Macronutrients		
Proteins	Fats	Carbohydrates
Nonessential amino acids	Triglycerides - saturated fatty acids	Sugars - simple carbohydrates
Essential amino acids	- unsaturated fatty acids	- complex carbohydrates
Conditionally essential amino acids	- mono-unsaturated fatty acids	Starches
Complete (animal)	- polyunsaturated fatty acids	Fibres
Complimentary (plant)	(e.g. omega 3 & omega 6)	
	Phospholipids	
	Steroids - cholesterol	

Micronutrients				
B vitamins#	Fat-soluble vitamins	Water-soluble vitamins	Major minerals	Trace Minerals+
Thiamin (B1)	A*^	C*	Sodium	Iron
Riboflavin (B2)	D		Potassium	Iodine
Niacin (B3)	E*		Calcium	Zinc
Biotin	K		Magnesium	Fluoride
Pantothenic acid			Phosphorus	Copper
B-6			Sulfur	Selenium
Folate			Chloride	
B-12				

#The B vitamins are also water-soluble
*Denotes anti-oxidants
^Beta-carotene is an anti-oxidant and is a precursor to vitamin A
+These nutrients are dependent on soil and water quality and the extent the food has been processed

Appendix J: Digit Ordering stimuli

In the following test, I will say several digits in random order. Please recall these digits in ascending order immediately after presentation.

For example, if I say '5 – 2 – 8 – 4', the correct recall for you to say would be, '2 – 4 – 5 – 8'

We'll do three examples before we start. So, after I say some numbers, please recall them in ascending/increasing order.

Example 1:	'7 – 6'	answer: 6 – 7	correct	☐
			incorrect	☐
Example 2:	'6 – 3 – 7'	answer: 3 – 6 – 7	correct	☐
			incorrect	☐
Example 3:	'9 - 8 – 4'	answer: 4 – 8 – 9	correct	☐
			incorrect	☐

Actual Test

Now that you have the hang of it, we will begin the test. After each set of three, the number of digits I present to you will increase by an extra digit. Do your best to recall the numbers in ascending/increasing order.

Span 2:

1.	'5 – 3'	answer:	3 – 5	correct	☐
				incorrect	☐
2.	'9 – 4'	answer:	4 – 9	correct	☐
				incorrect	
3.	'8 – 2'	answer:	2 – 8	correct	☐
				incorrect	☐

Span 3:

1.	'3 – 7 – 2'	answer:	2 – 3 – 7	correct	☐
				incorrect	☐
2.	'2 – 8 – 1'	answer:	1 – 2 - 8	correct	☐
				incorrect	☐
3.	'4 – 3 – 2'	answer:	2 – 3 – 4	correct	☐
				incorrect	☐

Span 4:

1.	'8 – 4 – 7 – 3'	answer:	3 – 4 – 7 – 8	correct	☐
				incorrect	☐
2.	'7 – 3 – 1 – 5'	answer:	1 – 3 – 5 – 7	correct	☐
				incorrect	☐
3.	'9 – 6 – 8 – 1'	answer:	1 – 6 – 8 – 9	correct	☐
				incorrect	☐

Span 5:

1.	'2 – 6 – 5 – 3 – 1'	answer:	1 – 2 – 3 – 5 – 6	correct	☐
				incorrect	☐
2.	'3 – 4 – 7 – 1 – 8'	answer:	1 – 3 – 4 – 7 - 8	correct	☐
				incorrect	☐
3.	'6 – 3 – 8 – 2 – 4'	answer:	2 – 3 – 4 – 6 – 8	correct	☐
				incorrect	☐

Span 6:

1. '4 – 8 – 3 – 9 – 2 – 7' answer: 2 – 3 – 4 – 7 – 8 – 9 correct ☐
   incorrect ☐

2. '7 – 5 – 3 – 6 – 4 – 8' answer: 3 – 4 – 5 – 6 – 7 - 8 correct ☐
   incorrect ☐

3. '5 - 6 – 8 – 4 – 3 – 1' answer: 1 – 3 – 4 – 5 – 6 – 8 correct ☐
   incorrect ☐

Span 7:

1. '7 – 4 – 3 – 1 – 8 - 2 – 6' answer: 1 – 2 – 3 – 4 – 6 – 7 – 8 correct ☐
   incorrect ☐

2. '2 – 9 – 1 – 4 – 7 – 5 – 8' answer: 1 – 2 – 4 – 5 – 7 – 8 – 9 correct ☐
   incorrect ☐

3. '1 – 3 – 7 – 6 – 2 - 9 – 5' answer: 1 – 2 – 3 – 5 – 6 – 7 – 9 correct ☐
   incorrect ☐

Span 8:

1. '5 - 9 - 6 - 4 - 7 - 1 - 3 – 8' answer: 1 – 3 – 4 - 5 – 6 – 7 – 8 – 9 correct ☐
   incorrect ☐

2. '3 – 6 – 5 – 2 – 8 – 7 – 4 – 9' answer: 2 – 3 – 4 - 5 – 6 – 7 – 8 – 9 correct ☐
   incorrect ☐

3. '4 – 7 – 6 – 2 – 9 – 8 – 3 – 1' answer: 1 – 2 – 3 – 4 – 6 – 7 – 8 – 9 correct ☐
   incorrect ☐

Appendix K: Reasons behind Dot Matrix missing data

There were procedural shortcomings that interfered with data collection. The extent of these was not fully recognised when designing the current study, and nor were the limitations apparent during pilot testing. Nonetheless, limitations that resulted in a smaller sample size for the Dot Matrix task (as highlighted in Chapter 4) stemmed from the characteristics of the sample, the environment of the test administration, and the nature of the test battery. The sample involved elderly participants, several of whom had health problems to varying degrees. For example, some participants had health problems that limited their ability to sit still for extended periods of time (e.g. arthritis, back problems); and others had eye problems (e.g. dry eyes) that made completing some tasks, especially the computer tasks, more difficult or, at least, lengthy. In addition to health problems, some participants showed displeasure when asked to complete the computer-administered tasks, and in some cases, refused to do so. The latter reaction was likely to be due to unfamiliarity with computer technology and associated feelings of frustration and self-perceived failure or lack of intelligence. Participants also tended to be particularly talkative and often digressed from the task at hand. Conducting data collection in participants' homes also increased the likelihood of digression from tasks; and the presence of their own memorabilia often prompted tangential conversations. Such occurrences appeared to be less common with individuals who completed test sessions at the University.

The types of limitations described above led to an increase in the amount of time some participants required to complete the test battery. Another demand on time was assisting participants to complete the questionnaires mailed out prior to the test session (as explained at the end of Chapter 3). Where possible (i.e. without impeding other test sessions), additional time was allocated to these situations. However, Dot Matrix was a

difficult, computerised task that was scheduled towards the end of the test battery (the rationale for this was also provided in Chapter 3). Altogether, these factors resulted in missing data for Dot Matrix. As will be demonstrated below, there appeared to be no systematic differences between individuals with and without Dot Matrix data.

Independent samples t-tests were carried out to compare performance on measures that may have influenced the likelihood of completing versus not completing the Dot Matrix task. These measures included: demographic information (i.e. age, education, sex), working memory ability (i.e. Reading Span), general intelligence (i.e. Raven's Standard Progressive Matrices), functional ability (Activities of Daily Living) and health (i.e. total number of symptoms and diseases). The mean and standard deviation for each measure are provided below (Table K1).

Table K1
Various performance comparisons between participants who did, and did not, complete the Dot Matrix task

	Completed Dot Matrix (N = 63)		Did not complete Dot Matrix (N = 87)	
Comparison measures	Mean	SD	Mean	SD
Age	77.60	4.50	77.58	4.33
Years of education	11.68	4.49	11.68	3.86
Reading Span	13.78	8.64	14.01	8.99
Raven's	16.76	5.16	16.51	4.41
Basic ADL	6.76	.67	6.86	.57
Instrumental ADL	25.78	2.04	25.82	1.75
Total symptoms	68.94	12.20	69.97	12.52
Total diseases	2.43	1.66	2.63	1.53

ADL = Activities of daily living; Raven's = Raven's Standard Progressive Matrices

None of the above measures showed significant mean differences between participants who completed the Dot Matrix task and those who did not. Similarly, 61 females and 26 males did not complete Dot Matrix; this figure was generally representative of the 2:1 gender ratio of the overall sample. Therefore, it appears that there were no systematic differences between individuals with and without Dot Matrix data.

A related issue for missing Dot Matrix data was the location of test sessions. Participants were given the option of completing test sessions in their home or at the University. The majority (80.6%) of participants chose to complete the test session in their home. Providing the option of at-home testing expanded the number and type of elderly people who were able to participate in the current study. This is in part reflected by the fact that participants who completed the test session at home had significantly lower scores (i.e. were more dependent) in basic and instrumental activities of Daily Living (see Table K2). Independent samples t-tests involving other salient variables also showed that participants who were tested at home were significantly older than those tested at University.

Table K2
Various performance comparisons between participants who completed the initial test session at their home and at the University

	At-home (N = 121)		University (N = 29)	
Comparison measures	Mean	SD	Mean	SD
Age*	78.04	4.26	75.71	4.49
Years of education	11.52	4.13	12.35	4.12
Reading Span	13.88	8.76	14.03	9.19
Dot Matrix	8.28	4.62	11.06	6.93
Raven's	16.45	4.80	17.31	4.43
Basic ADL**	6.78	.68	7.00	.00
Instrumental ADL*	25.67	1.96	26.34	1.34
Total symptoms	69.54	12.78	69.51	10.59
Total diseases	2.60	1.63	2.34	1.40

**p < .01; *p < .05
ADL = Activities of daily living; Raven's = Raven's Standard Progressive Matrices
Age: t = 2.61*, df = 148; Basic ADL: t=3.63**, df = 120; Instrumental ADL: t = 2.20*, df = 60.

Chi-square analysis also showed that males and females were not disproportionately represented across the above groups (Yates' Correction for Continuity = 2.54, df = 1, p > .05). The group of participants who completed the initial test session at University (N = 29) included 15 females (52% of this total) and 14 males. The remaining participants (N = 121; 84 females and 37 males) completed the test session at home. These at-home figures represent 84.8% of all females tested and 72.6% of all males tested. Further comparisons were made regarding group membership

between test location (i.e. at-home versus University) and completion of the Dot Matrix task (i.e. completed versus not completed). Chi-square statistics showed that 75 participants tested at home (62% of the at-home total) did not complete the Dot Matrix task compared to 12 participants tested at University (41.4% of the University total). These figures lend support to the rationale provided earlier for missing data on the Dot Matrix task. However, this difference in proportions did not reach significance (Yates' Correction of Continuity = 3.28, df = 1, p = .07).

Appendix L: Australian recommended dietary intake levels for various nutrients

Table L1

Recommended dietary intakes for adults (expressed as mean daily intake)

	Males (64 years^)	Females (54 years and older)
B-vitamins		
Thiamin (B1)	0.9	0.7
Riboflavin (B2)	1.3	1.0
Niacin (B3)	16.0	11.0
B-6	1.0 - 1.5	0.8 - 1.1
Folate#	200.0	200.0
B-12#	2.0	2.0
Anti-oxidants		
Vitamin C	40	30
Vitamin E	10	7
Vitamin A#	<7500	<7500

#measured in micrograms (μg); all others are measured in milligrams (mg)

^This is the oldest age range available for males as produced by the National Health and Medical Research Council (1991)

Appendix M: Relationships between depression and biomarker and independent functioning variables (initial assessment occasion)

Tables M1 and M2 contain bivariate correlations (Pearson's r) for the maximum sample size. Instrumental activities of daily living (ADL) and reasoning ability appear to have significant, negative relationships with the number of depressive symptoms. None of the biomarker variables was significantly associated with depression.

Table M1
Correlations between depression and outcome variables

Outcome variables	N	Depression
Activities of daily living		
Basic	150	-.07
Instrumental	150	-.21*
Reasoning ability	150	-.22**
Life Satisfaction#		

**p < .01; * p < .05
#Explained in thesis text (see Chapter 5, p.144)

Table M2
Correlations between depression and biomarkers

Biomarkers	N	Depression
Age	150	-.09
Systolic BP	137	.02
Diastolic BP	137	.03
Grip Strength	150	-.12
BMI	150	.07
Visual Acuity	150	-.09
Reading Span	149	-.13

A boxplot of the depression variable showed the presence of several outliers (i.e. scores greater than 10). Table M3 shows that the correlations between independent functioning outcome measures and depression were no longer significant when these outliers were excluded.

Table M3
Correlations (Pearson's r) between depression and outcome variables

Outcome variables	N	Depression
Activities of daily living		
Basic	137	-.12
Instrumental	137	-.04
Reasoning ability	137	-.17
Life Satisfaction#		

#Explained in thesis text (see Chapter 5, p.144)

Appendix N: Relationships between health and biomarker and independent functioning variables (initial assessment occasion)

Tables N1 and N2 contain bivariate correlations (Pearson's r) for the maximum sample size. Reasoning ability, grip strength and Reading Span appear to have significant, negative relationships with either symptom severity or total number of diseases.

Table N1
Correlations between health and outcome variables

Outcome variables	N	Symptom severity	No. of diseases
Activities of daily living			
Basic	150	-.08	-.11
Instrumental	150	-.16	-.14
Reasoning ability	150	-.14	-.18*
Life Satisfaction#			

*p < .05
#Explained in thesis text (see Chapter 5, p.145)

Table N2
Correlations between health and biomarkers

Biomarkers	N	Symptom severity	No. of diseases
Age	150	-.14	.07
Systolic BP	137	-.01	.06
Diastolic BP	137	-.08	.02
Grip Strength	150	-.21*	-.18*
BMI	150	.04	.13
Visual Acuity	150	.05	.07
Reading Span	149	-.06	-.20*

BP=Blood pressure; BMI=Body mass index
*p < .05

Inspection of boxplots revealed that reasoning ability and the total number of diseases did not have any outliers. However, when sex, another covariate was taken into consideration, the significant relationship between number of diseases and reasoning ability disappeared for females ($r = -.06$, $p > .05$, $N = 99$). In contrast, this significant relationship remained for males ($r = -.38$, $p < .01$, $N = 51$). Given that males and females could not be analysed separately in the regression analyses (current sample was too small for regression requirements), including total number of diseases as a covariate was impractical. However, this would be a variable of interest if a similar study was

conducted with males only. Perhaps males react more negatively (at least in terms of thought processes) to poor health than females.

Analysis by sex for biomarker variables demonstrated similar results as described above. Grip strength was no longer reliably related to symptom severity or total number of diseases in females ($r = -.17$, $p > .05$, $N = 98$ and $r = -.15$, $p > .05$, $N = 99$, respectively). This pattern also applied to males ($r = -.20$, $p > .05$, $N = 51$ and $r = -.12$, $p > .05$, $N = 51$, respectively). Consistent with this, Reading Span was also no longer significantly related to the total number of diseases for females ($r = -.15$, $p > .05$, $N = 98$) and males ($r = -.24$, $p > .05$, $N = 51$).

Appendix O: Relationships between lifestyle variables and biomarker and independent functioning variables (initial assessment occasion)

Tables O1 and O2 contain bivariate correlations (Pearson's r) for the maximum sample size between smoking status and outcome variables and biomarkers. None of the outcome variables was significantly different across smoking categories (see Table O1). Age was the only biomarker to vary significantly across the smoking categories (see Table O2, p.339). However, given that none of the outcome variables was significantly related to smoking category, including smoking category as a covariate would be redundant (biomarker variables were investigated in terms of their ability to predict outcome measures, not to predict covariates).

Table O1
Influence of smoking status on outcome variables

	Smoking Status		
Outcome Variables	Never smoked (N = 64)	Former smokers (N = 78)	Current smokers (N = 7)
Activities of daily living (ADL)			
Basic	6.73	6.87	7.00
Instrumental	25.55	26.05	25.71
Reasoning ability^			
Females	14.82	16.54	15.00
Males	20.38	18.44	17.00
Life Satisfaction	148.44	145.46	143.71

^ As a result of pre-established differences between males and females on this measure, one-way ANOVAs were conducted for each sex

As can be seen by Table O3 and O4 (p.339), no variable was significantly related to the overall amount of exercise conducted each week.

Table O3
Correlations between exercise and outcome variables

Outcome variables	N	Overall exercise
Activities of daily living		
Basic	150	.10
Instrumental	150	.16
Reasoning ability	150	.08
Life Satisfaction	147	.11

Initial results indicated that performance on basic activities of daily living (ADL) was significantly different among alcohol consumption groups (Table O5,

p.340). However, this result was a consequence of outliers. When outliers were excluded, all groups had a mean of 7.00 for basic ADL. Ceiling effects on basic ADL led to scores below 7.00 (the maximum achievable) being outliers. In terms of the biomarkers, grip strength was significantly different between abstainers and light drinkers for females and between light and moderate drinkers for males (see Table 06, p.340). In both cases, light drinkers had significantly stronger grip strengths. However, given that none of the outcome variables was significantly related to alcohol consumption group, including this variable as a covariate would be redundant. As aforementioned, biomarker variables were investigated in terms of their ability to predict outcome measures, not to predict covariates.

Table O2
Influence of smoking status on biomarkers

	Smoking Status		
Biomarkers	Never smoked (N = 64)	Former smokers (N = 78)	Current smokers (N = 7)
Age*	77.40	78.02	73.55
Systolic BP	148.32	146.42	156.39
Diastolic BP	79.88	78.88	75.67
Grip Strength^			
Females	15.09	14.14	14.33
Males	30.89	29.45	27.50
BMI	26.51	26.61	24.28
Visual Acuity	93.48	93.55	92.43
Reading Span^			
Females	12.51	13.51	12.20
Males	17.00	15.85	16.00

BP = blood pressure; BMI = body mass index
^ As a result of pre-established differences between males and females on this measure, one-way ANOVAs were conducted for each sex
*Age: $F(2, 148) = 3.53$, $p < .05$, $eta^2 = .05$.

Table O4
Correlations between exercise and biomarkers

Biomarkers	N	Overall exercise
Age	150	-.08
Systolic BP	137	-.02
Diastolic BP	137	-.09
Grip Strength	150	.06
BMI	150	-.06
Visual Acuity	150	.02
Reading Span	149	.11

Table O5
Influence of alcohol consumption group on outcome variables

Outcome Variables	Alcohol consumption group: Abstainers (N = 63) Mean	Light drinkers (N = 63) Mean	Moderate drinkers (N = 22) Mean
Activities of daily living (ADL)			
Basic*	6.70	6.95	6.95
Instrumental	25.49	26.02	26.14
Reasoning ability^			
Females	15.34	15.58	16.57
Males	19.08	18.57	18.53
Life Satisfaction	144.13	149.00	147.68

^As a result of pre-established differences between males and females on this measure, one-way ANOVAs were conducted for each sex
*Basic ADL: $F(2, 145) = 3.61$, $p < .05$, $eta^2 = .05$.

Table O6
Influence of alcohol consumption group on biomarkers

Biomarkers	Alcohol consumption group: Abstainers (N = 63)	Light drinkers (N = 63)	Moderate drinkers (N = 22)
Age	77.56	77.43	78.06
Systolic BP	146.76	148.89	146.38
Diastolic BP	78.94	80.35	76.05
Grip Strength^			
Females**	13.38 (N = 47)	16.77 (N = 39)	13.52 (N = 7)
Males*	29.05 (N = 13)	31.97 (N = 23)	26.44 (N=15)
BMI	26.77	25.94	26.65
Visual Acuity	93.27	93.44	94.05
Reading Span^			
Females	11.94	13.97	12.43
Males	14.23	16.57	16.80

BP = blood pressure; BMI = body mass index
^As a result of pre-established differences between males and females on this measure, one-way ANOVAs were conducted for each sex
**Grip strength (females): $F(2, 94) = 5.55$, $p < .01$, $eta^2 = .11$ (Post hoc analysis: abstainers & light drinkers, $p < .01$); *Grip strength (males): $F(2, 48) = 3.74$, $p < .05$, $eta^2 = .16$ (Post hoc analysis: light drinkers & moderate drinkers, $p < .05$).

The majority of outcome variables and overall level of nutrient intakes were not significantly correlated (see Table O7, p.341). However, Vitamin C and Vitamin A were significantly related to life satisfaction, and Vitamin C to instrumental ADL. When analysed by sex, the significant correlation between Vitamin C and instrumental ADL applied to females ($r = -.25$, $p < .05$, $N = 89$) but not males ($r = -.03$, $p > .05$, $N = 46$). Boxplot inspection also showed that Vitamin C had many outliers (i.e. participants

taking nutrient supplements). Once these were excluded, the relationship between instrumental ADL and Vitamin C for females was no longer significant ($r = .06$, $p > .05$, $N = 80$). This same process, applied to the correlation between life satisfaction and Vitamin C, yielded non-significant results for females ($r = -.21$, $p > .05$, $N = 75$) and males ($r = .04$, $p > .04$, $N = 46$). Analysis by sex for Vitamin A showed non-significant relationships; females, $r = .20$, $p > .05$, $N = 87$ and males, $r = -.08$, $p > .05$, $N = 45$).

Table O7
Correlations between overall nutrient intake of B-Vitamins and anti-oxidants and outcome variables (N = 135)

	Basic ADL	Instrumental ADL	Reasoning ability	LS
B-Vitamins				
Thiamin	.00	-.02	-.01	.02
Riboflavin	.01	-.02	-.03	.01
Niacin	.05	.02	-.02	-.03
B-6	.02	.01	.07	.05
Folate#	.08	.10	.06	.03
B-12#	.02	-.03	-.12	.02
Antix-oxidants				
Vitamin C	.03	-.20*	.06	-.22*
Vitamin E	.01	-.11	-.09	-.07
Vitamin A#	-.03	.08	-.07	.17*

ADL = Activities of daily living; LS = Life Satisfaction
#measured in micrograms (μg); all others are measured in milligrams (mg)
*$p < .05$

Of the biomarkers, body mass index (BMI) showed the greatest number of significant correlations with different nutrients; and folate was significantly associated with grip strength and Reading Span (see Table O8, p.342). However, like the outcome variables, most of these disappeared when sex and outliers were taken into consideration. When folate was analysed by sex, it was significantly related to grip strength in males ($r = .36$, $p < .05$, $N = 46$) but not females ($r = .04$, $p > .05$). This significant relationship in males persisted with the exclusion of outliers ($r = .39$, $p < .05$, $N = 40$). Therefore, reported higher intakes of folate appear to be associated with increased grip strength in males. However, folate was not employed as a covariate here because it was not reliably associated with any of the outcome variables (see above for

Table O8

Correlations between overall nutrient intake of B-Vitamins and anti-oxidants and biomarkers (N = 135)

	Age	Systolic BP	Diastolic BP	Grip Strength	BMI	Visual Acuity	Reading Span
B-Vitamins							
Thiamin	.15	-.10	-.06	.00	-.23**	-.02	-.01
Riboflavin	.14	-.12	-.09	.06	-.22*	-.00	.02
Niacin	-.04	-.00	-.03	.03	-.16	.16	-.00
B-6	.06	-.08	-.04	.01	-.20*	-.01	.02
Folate#	-.10	-.11	.01	.17*	-.13	.09	.20*
B-12#	.06	-.05	-.03	-.03	-.20*	.02	-.08
Antix-oxidants							
Vitamin C	-.10	.03	.08	-.04	-.04	.08	.04
Vitamin E	.05	.13	.04	-.02	-.08	-.03	-.05
Vitamin A#	-.06	.01	.00	-.14	.16	.08	-.02

measured in micrograms (μg); the remainder is measured in milligrams (mg)

**p < .01;*p < .05

rationale). In contrast, analysis by sex and the exclusion of outliers resulted in a non-significant relationship between folate and Reading Span for females ($r = .09$, $p > .05$, $N = 89$), and males ($r = .00$, $p > .05$, $N = 40$).

The same steps were applied to body mass index (BMI) and thiamin, riboflavin, Vitamin B-6 and Vitamin B-12. There were no significant correlations between Vitamin B-6 or B-12 and BMI for females ($r = -.19$, $p > .05$, $N = 89$ and $r = -.19$, $p > .05$, $N= 89$, respectively); riboflavin was not significantly related to BMI for males ($r = -.28$, $p > .05$, $N = 46$). In addition, the exclusion of outliers made the relationships between BMI and thiamin ($r = -.13$, $p > .05$, $N = 71$) and riboflavin ($r = .02$, $p > .05$, $N = 67$) non-significant in females. Similarly, Vitamin B-6 ($r = .06$, $p > .05$, $N = 43$) and B-12 ($r = -.06$, $p > .05$, $N = 43$) were no longer significantly correlated with BMI in males. Alternatively, for males, thiamin and BMI remained significantly correlated ($r = -.40$, $p < .01$, $N = 41$). Perhaps thiamin prevents weight gain in older males. However, thiamin was not included as a covariate (reasons aforementioned).

Independent samples t-tests were conducted to compare performance on different variables across recommended dietary intake (RDI) groups. Initially, of the outcome variables, mean basic ADL performance was significantly different across riboflavin RDI groups. Similarly, Life Satisfaction was significantly different across Vitamin A groups (see Table O9, pp.344-345). Examination of these results by sex showed that the difference in basic ADL performance for riboflavin was only significant for females (above RDI, mean = 7.00; below RDI = 6.73; $t = 3.37$, $df = 81$, $p < .01$). However, exclusion of outliers for basic ADL made this difference non-significant. The difference in life satisfaction for Vitamin A group was also significant for females only ($t = 2.57$, $df = 85$, $p < .05$); no males were in the above RDI group. This difference was still significant after the exclusion of outliers. However, with a

Table O9

A comparison of means for outcome variables across different RDI groups for various B-Vitamins and anti-oxidants (initial test occasion)

	N	Basic ADL	Instrumental ADL	Reasoning ability	Life Satisfaction
B-Vitamins					
Thiamin					
Below RDI	11	6.91	24.91	16.91	143.36
Above RDI	124	6.81	25.86	16.78	147.12
t (df)		.52 (133)	1.61 (133)	.08 (133)	.85 (130)
Riboflavin					
Below RDI	11	7.00	26.00	17.55	146.36
Above RDI	124	6.80	25.77	16.73	146.85
t (df)		3.45 (123)**	.39 (133)	.54 (133)	.11 (130)
Niacin					
Below RDI	0				
Above RDI	135				
B-6					
Below RDI	12	7.00	25.58	18.67	144.67
Within RDI	23	6.83	25.83	16.74	146.43
Above RDI	100	6.79	25.80	16.58	147.15
F (df)		.61 (2,132)	.08 (2,132)	1.02 (2, 132)	.18 (2, 129)
Folate#					
Below RDI	21	6.86	25.52	15.52	143.57
Above RDI	114	6.81	25.83	17.03	147.42
t (df)		.34 (133)	.69 (133)	1.32 (133)	1.16 (130)
B-12#					
Below RDI	27	6.74	25.82	15.85	147.80
Above RDI	108	6.83	25.78	17.03	146.58
t (df)		.69 (133)	.09 (133)	1.14 (133)	.39 (130)

measured in micrograms (μg); all others are measured in milligrams (mg)

**$p < .01$

Table O9 (continued)
A comparison of means for outcome variables across different RDI groups for various B-Vitamins and anti-oxidants (initial test occasion)

	N	Basic ADL	Instrumental ADL	Reasoning ability	Life Satisfaction
Antix-oxidants					
Vitamin C					
Below RDI	11	6.91	25.27	16.36	145.45
Above RDI	124	6.81	25.83	16.83	146.93
t (df)		.52 (133)	.93 (133)	.31 (133)	.34 (130)
Vitamin E					
Below RDI	79	6.86	25.80	17.18	147.49
Above RDI	56	6.75	25.77	16.25	145.85
t (df)		.95 (87)	.09 (133)	1.11 (133)	.66 (130)
Vitamin A#					
Within RDI	133	6.81	25.77	16.80	146.39
Above RDI	2	7.00	27.00	16.50	174.00
t (df)		.42 (133)	.91 (133)	.03 (133)	2.84 (130)**

#measured in micrograms (µg); all others are measured in milligrams (mg)
**p < .01

sample of N = 2, this differences was considered unreliable.

Table O10 (pp.347-348) shows the same initial comparisons for biomarkers across RDI groups. Examination of these results by sex showed that age was not significantly different across thiamin and riboflavin groups for males ($t = 1.42$, $df = 44$, $p > .05$ and $t = .58$, $df = 44$, $p > .05$, respectively). For females, age was significantly different across riboflavin groups ($t = 2.21$, $df = 87$, $p < .05$) but not thiamin groups ($t = 1.55$, $df = 87$, $p > .05$). Riboflavin RDI group was not covaried because it only effected age. Similarly, independent samples t-test showed that grip strength was not significantly different across Vitamin B-6 groups for females [$F(2,86) = 1.37$, $p > .05$] or males [$F(2,43) = .01$, $p > .05$]. The same outcome occurred when sex was accounted for across folate group and visual acuity (females: $t = 1.29$, $df = 6$, $p > .05$; males: $t = 1.29$, $df = 6$, $p > .05$). The significant difference in visual acuity across Vitamin E groups was not significant for females ($t = 1.54$, $df = 66$. $p > .05$). Initially it was for males ($t = 2.96$, $df = 35$, $p < .01$) but significance disappeared when outliers were excluded. Lastly, BMI was significantly different across Vitamin A RDI groups for females only ($t = 2.75$, $df = 87$, $p < .01$; there were no males in this group). Exclusion of outliers did not alter the significance but, as aforementioned, with a sample size of N = 2, this differences was considered unreliable.

Table O10

A comparison of means for biomarkers across different RDI groups for various B-Vitamins and anti-oxidants (initial test occasion; *p < .05)

	N	Age	Systolic BP	Diastolic BP	Grip Strength	BMI	Visual Acuity	Reading Span
B-Vitamins								
Thiamin								
Below RDI	11	74.97	142.99	76.85	23.33	26.01	93.45	16.18
Above RDI	124	77.67	148.02	79.41	19.60	26.52	93.53	14.21
t (df)		2.05 (133)*	.70 (121)	.72 (121)	1.30 (133)	.39 (133)	.09 (133)	.72 (132)
Riboflavin								
Below RDI	11	74.79	150.20	79.50	21.57	26.60	92.64	16.73
Above RDI	124	77.68	147.34	79.15	19.76	26.47	93.60	14.16
t (df)		2.20 (133)*	.38 (121)	.09 (121)	.63 (133)	.10 (133)	1.07 (133)	.93 (132)
Niacin								
Below RDI	0							
Above RDI	135							
B-6								
Below RDI	12	75.75	154.17	81.17	25.61	25.60	92.83	16.25
Within RDI	23	77.82	149.78	79.72	20.85	26.36	92.65	14.87
Above RDI	100	77.56	146.22	78.80	19.00	26.61	93.81	14.03
F (df)		1.09 (2,132)	.76 (2,120)	.26 (2,120)	3.05 (2,132)*	.33 (2,132)	1.93 (2, 132)	.39 (2, 131)
Folate#								
Below RDI	21	76.98	154.24	82.05	17.68	26.92	92.33	13.10
Above RDI	114	77.53	146.43	78.69	20.31	26.40	93.75	14.61
t (df)		.55 (133)	1.35 (121)	1.18 (121)	1.22 (133)	.53 (133)	2.11 (133)*	.73 (132)
B-12#								
Below RDI	27	77.69	145.61	75.76	17.41	26.64	92.59	15.92
Above RDI	108	77.38	148.07	80.05	20.53	26.44	93.76	14.00
t (df)		.33 (133)	.48 (121)	1.72 (121)	1.60 (133)	.23 (133)	1.91 (133)	1.01 (132)

Table O10 (continued)
A comparison of means for biomarkers across different RDI groups for various B-Vitamins and anti-oxidants (initial test occasion)

	N	Age	Systolic BP	Diastolic BP	Grip Strength	BMI	Visual Acuity	Reading Span
Antix-oxidants								
Vitamin C								
Below RDI	11	75.74	148.60	79.03	22.45	25.39	93.73	11.27
Above RDI	124	77.60	147.48	79.19	19.68	26.57	93.51	14.65
t (df)		1.40 (133)	.15 (121)	.04 (121)	.96 (133)	.91 (133)	.24 (133)	1.23 (132)
Vitamin E								
Below RDI	79	77.46	145.27	78.91	20.96	26.94	93.13	14.09
Above RDI	56	77.42	150.81	79.56	18.41	25.82	94.09	14.78
t (df)		.05 (133)	1.33 (121)	.31 (121)	1.61 (133)	1.57 (133)	2.14 (133)*	.45 (132)
Vitamin A#								
Within RDI	133	77.47	147.55	79.19	19.89	26.34	93.50	14.37
Above RDI	2	75.72	149.60	78.00	20.15	35.47	95.00	14.50
t (df)		.58 (133)	N/A (1)	N/A (1)	.04 (133)	3.20 (133)**	.73 (133)	.02 (132)

measured in micrograms (μg); all others are measured in milligrams (mg)
N/A= Not applicable
**p < .01;*p < .05

Appendix P: Bivariate correlations (Pearson's r) between dependent and independent variables for cross-sectional regression results (Chapter 5)

	Dependent variables			
	Basic ADL	Instrumental ADL	Raven's	Life Satisfaction
Covariates				
Sex	-.17*	-.03	.32**	-.03
Education	-.04	-.07	.29**	.14
Biomarkers				
Age	.25**	.31**	-.24**	.20*
Systolic BP	.04	.01	.13	-.04
Diastolic BP	-.13	-.07	.26**	-.04
Grip strength	-.29**	-.14	.39**	.07
BMI	.15	-.08	.19*	-.05
Visual acuity	-.03	-.16	.26**	-.02
Reading Span	-.15	-.14	.43**	.09

*p < .05; **p < .01

ADL = Activities of Daily Living; Raven's = Raven's Standard Progressive Matrices; BP = Blood Pressure; BMI = Body Mass Index;

Appendix Q: Bivariate correlations (Pearon's r) between dependent and independent variables for longitudinal regression results (Chapter 6)

Covariates		Dependent variables (final measurement scores) Basic ADL	Instrumental ADL	Raven's	Memory-based ADL
Sex		-.25**	-.10	.21*	-.05
Education		-.08	-.02	.35**	.13
Biomarkers					
Age	Initital score	.22*	.29**	-.17	-.09
	6 month change	.13	.05	.02	-.04
	18 month change	.06	.03	.06	-.03
Systolic BP	Initital score	.03	.05	-.02	.07
	6 month change	.01	.09	-.04	-.23*
	18 month change	-.06	.10	-.11	-.11
Diastolic BP	Initital score	-.11	-.01	.09	.05
	6 month change	-.03	-.04	-.12	-.18
	18 month change	-.01	.01	-.02	.03
Grip strength	Initital score	-.27**	-.22*	.26**	.04
	6 month change	.08	.20	-.13	-.12
	18 month change	.02	.13	.19*	-.03
BMI	Initital score	.20*	.03	.12	.10
	6 month change	.09	.01	-.03	.01
	18 month change	-.13	-.19*	.06	.08
Visual acuity#	Initital score	.04	-.02	.06	.02
Reading Span	Initital score	-.21*	-.25**	.35**	.27**
	6 month change	-.07	-.07	.23*	.08
	18 month change	.01	-.11	.22*	.17

*p < .05; **p < .01

ADL = Activities of Daily Living; Raven's = Raven's Standard Progressive Matrices; BP = Blood Pressure; BMI = Body Mass Index;
#Standardised regression-based change scores were not calculated; see p.171 for more details.

	Dependent variables (18 month change scores)			
Covariates	Basic ADL	Instrumental ADL	Raven's	Crystallised ability
Sex	.35**	.51**	.05	.16
Education	.27**	-.07	.05	.06
Biomarkers#				
Age	.01	.09	-.03	-.03
Systolic BP	-.12	-.03	-.17	-.11
Diastolic BP	-.21*	.04	-.07	-.07
Grip strength	.27**	.35**	.05	.18
BMI	.12	.06	-.01	.02
Visual acuity	.12	.08	-.10	-.07
Reading Span	.10	-.04	.14	.23*

*p < .05; **p < .01

#Initial scores only were employed to predict change in the dependent variables; see p.178 for the rationale behind this

ADL = Activities of Daily Living; Raven's = Raven's Standard Progressive Matrices; BP = Blood Pressure; BMI = Body Mass Index