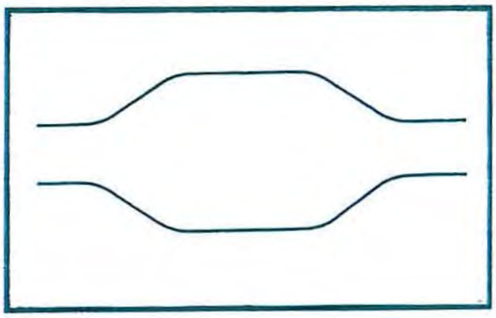
THE SHUTTLE EFFECT

THE DEVELOPMENT OF A MODEL FOR THE PREDICTION OF VARIABILITY IN COGNITIVE TEST PERFORMANCE ACROSS THE ADULT LIFE SPAN



Ann B Jordan

M Sc Clin Psych (UCT)

Thesis submitted in fulfilment of the requirements for the degree Doctor of Philosophy

Faculty of Science, Rhodes University
Grahamstown, South Africa

1997

This thesis is dedicated to my beloved parents
Betty and Stanley Shuttleworth

in proud appreciation of
their boundless energy for creative living
right into late old age

and with deep gratitude for
a lifetime of encouragement and support

ACKNOWLEDGEMENTS

I wish to convey my thanks to a number of people whose contributions over the past two years have finally come together in the form of this thesis. The order of my acknowledgement serves to tell the story of the project.

- Warrick Peacock, Neurosurgeon, friend and colleague, originally from Groote Schuur Hospital in Cape Town, whose initiative led to my taking up a year's research fellowship in traumatic brain injury at the University of California, Los Angeles (UCLA) during my 1995 sabbatical leave.

- Ann Peacock, who consistently provided generous assistance in practical matters such as finding me suitable accommodation close to the University before I even left home. This, in addition to the embracing manner in which I was welcomed as part of the Peacock family, set the scene for what turned out to be the year of a lifetime.

- Paul Satz, Chief Neuropsychologist at UCLA, who liaised with Warrick Peacock to create a space for me in the brain injury unit, and who generously offered to take on the supervision of this thesis. I am especially indebted to him for his warm reception of me in the UCLA Neuropsychology unit during my stay in the US, his enthusiastic encouragement, challenging mentorship, and for the commitment he has shown to me in this project long-distance in spite of onerous work schedules. I consider it an honour indeed to have worked together with someone of Paul's eminence in the field.

- Dan Kelly, co-ordinator of the UCLA head injury project on which I was working, who was quietly supportive and provided me with the time, office space and other facilities I needed to pursue my research in cognitive aging. Compared with what is available in South Africa, the UCLA library facilities, for example, had the quality of an 'Aladdin's lamp' dream.

- Sandy Bode and Sue Cornfield, my students and research assistants in the collection of South African normative aging data. Much of the raw data used in the Phase 2 analyses was collected by them for the purposes of central trends analyses. The availability of these data bases for further analyses in terms of inter-individual variability served an important role in the inception of the thesis.

- Craig Uchiyama, UCLA statistician, who was the first to tackle the array of statistical analyses involved in my thesis. He completed much of the Phase 2 raw data analyses at some sacrifice in terms of his busy schedule.

- Sarah Radloff, Rhodes University statistician, who took me under her supportive statistical wing on my return from the US. I am deeply indebted to her for the cheerful manner in which she undertook the mammoth task of all the data analyses, and for her patience, positive attitude and exceptional supervisory guidance with respect to the statistical aspects of the thesis.

- Susan Abraham and Bronwyn Tweedie, Rhodes University graphic artists, who completed the figures. It was a pleasure to work with people who were capable of such a high level of professionalism and fine attention to detail. The amount of graphical material required was no mean task, attempted and abandoned by several well-meaning others prior to my happy discovery of this specialist team.

- Ruth Jackson, *master* of secretaries, who played a central role in the typing, printing and organization of the manuscript. It is hard to acknowledge adequately the contribution of Ruth's unflagging energy and personal encouragement. In particular she never flinched in the typing of around 100 pages of intricate landscape tables with the utmost proficiency. She repeatedly helped me find energy resources (when I showed signs of wilting) to keep up the pace and persist.

- My children Nicky and Bobby, brother John, and other special family and friends, who showed their usual loving interest and support, as well as patience at times of inevitable dissertation 'absenteeism' during the past two years. I owe special thanks to my partner Dave whose incisive comments on my dissertation helped to clarify a number of important conceptual and practical issues. In particular his warm support in recent months gave me the strength to deal with my Mother's unexpected death. It is with a heavy heart that I come to terms with the fact that she will not be present to see me graduate.

ABSTRACT

The aim of this thesis was to investigate inter-individual variability on cognitive task performance in normal older adults. In a review of the cognitive aging literature, the implications of a differential perspective were drawn out in order to establish a theoretical and methodological basis for an investigation into variability. A number of regularly occurring patterns, identified on the basis of available reports in the literature, were used to develop a model of variability (*the shuttle model*). The empirically-based model was located broadly within a neuropsychological framework, and derived explanatory power from the tenets of brain reserve capacity (BRC) theory. It served to describe the bulge in inter-individual variability due to aging (*the shuttle bulge*), and the shifting occurrence of the bulge in relation to the age axis due to cohort and task-related influences (*the shuttle shift*). A two-phase research study was conducted in order to test hypotheses derived from the model. Phase 1 comprised between-groups analyses of normative data covering a broad range of neuropsychological tests in the domains of attention, memory, language, visual and hand motor skills, in order to examine the progression of variability effects across the adult age range. Phase 2 constituted between and within-groups analyses of normative data from a more limited number of neuropsychological tests. It included the examination of raw score distributions and the characteristics of outliers, and was undertaken to explore more closely the nature of the variability phenomena detected in the first phase of the analysis.

Taken together, the results of both phases of the investigation revealed statistically significant variability effects in support of the shuttle model. There was a consistent pattern of increased variability in association with older age regardless of functional modality; frequently, in association with later old age, there was also a subsequent decrease in variability (the shuttle bulge). The age of onset of the initial increase in variability occurred earlier or later (the shuttle shift) as a function of four factors: education, gender, task challenge and age-sensitivity of task. The finding of an earlier onset of variability effects for low education, male gender, high task challenge and high age-sensitivity of task was interpreted in terms of BRC threshold theory. The clinical and social implications of the outcome were discussed with special emphasis on the need for a differential perspective on aging, as a complement to the prevailing normative tradition. It was concluded that the shuttle model has considerable heuristic value. It presents an integrative framework for understanding existing variability data and provides clear indications for future research.

CONTENTS

VOLUME ONE

A differential perspective on cognitive aging and the development of a model

VOLUME TWO

An empirical investigation into variability and evaluation of the model

CONTENTS

VOLUME ONE

VOLUME TWO

VOLUME ONE

A DIFFERENTIAL PERSPECTIVE ON COGNITIVE AGING AND THE DEVELOPMENT OF A MODEL

1

CHAPTER 1: INTRODUCTION

'It is noteworthy that in the early days of experimental psychology all forms of random variability were regarded as "errors" insofar as they tended to restrict the applicability of general findings about human behavior.......to the psychometrician, however, these are not errors...... They are inalienable facts of behavior to be taken into account. Variability will not go away if you ignore it. If ignored, it remains to distort your results and probably lead to a wrong conclusion.'

(Anastasi, 1991, pp. 62-63, The Gap between Experimental and Psychometric Orientations).

'One of the most important questions in the psychology of aging concerns the origins of individual variations on the aging theme. Between successful agers, who maintain robust health and functional capacity to the end of a long life, and unsuccessful agers who experience premature senescence, early death, or incapacity, lies one of the most intriguing continuous distributions to challenge science'.

(Plomin & McLearn, 1990, p. 67, Human Behavioral Genetics of Aging).

1.1 BROAD AIMS OF THE THESIS

The first aim of this thesis was to conduct a review of the cognitive aging literature from the neglected differential perspective, in order to establish the basis for an investigation into the variability which occurs as part of the normal adult aging process. The second aim was to conduct a research study into variability by means of a statistical analysis of neuropsychological normative data available in the literature. Phase 1 of the empirical investigation set out to examine alterations in inter-individual variability as they occur *across* adult age groups for a broad spectrum of neuropsychological tests. This was primarily in order to describe group variability trends, and also to consider the implications of characteristic patterns in terms of normal aging processes. In Phase 2 of the investigation, on a more limited number of neuropsychological tests, it was decided to examine the raw score distribution of the inter-individual variability *within* adult age groups. It was considered that this raw score analysis would serve as a preliminary step towards describing and explaining group variability patterns across the adult age range identified in the first level of analysis. Taken together it was envisaged that the results of the two-phase research study would enhance conceptualizations of the aging process and the use of normative data in clinical work, from the neglected *differential* perspective as distinct from the more regularly employed central tendencies position taken in aging research and diagnostic practices.

1.2 THE NEGLECT OF INTER-INDIVIDUAL VARIABILITY

1.2.1 Literary perspective

In the areas of psychometrics and intellectual development in general, the importance of taking account of variability between individuals as a phenomenon to be studied rather than dismissed as error, has been strongly emphasized (Anastasi, 1958; 1991; Berg, 1992; Bidell & Fischer, 1992; Lezak, & Gray, 1984; Wohlwill, 1973). In particular two scholarly expositions exist which are dedicated to the study of individual differences in the form of a textbook by Anastasi (1958) covering basic psychometric, methodological and derivative aspects, and a chapter by Wohlwill (1973) which focuses on intra- and inter-individual differences in the area of child development. However the prevailing paradigm in psychological research has been strongly normative, and studies which use differences in inter-individual variability per se as the primary mechanism by which to demonstrate differences between groups rather than mean scores (for example, Rabbitt, 1993c; Shuttleworth-Jordan[1] & Saayman, 1989) appear as rare exceptions in the literature. Specifically with respect to the study of aging, a significant number of researchers, including leading experts in the field, make urgent reference to the centrality of the inter-individual variability issue, on the basis that there is considerable evidence to support a pattern of increasing heterogeneity with age in terms of social factors, psychological functioning and physical health status, (Aiken, 1995; Cunningham, 1990; Dannefer, 1988; Dolen, 1982; Gold & Arbuckle, 1990; Heron & Chown; Holland, 1990; Krauss, 1980; Labouvie-Vief, 1985; Lehr, 1980; Lovelace, 1990; Moscovitch & Winocur, 1992; Nettelbeck & Rabbitt, 1992; Owsley, 1991; Plomin & McClearn, 1990; Poon, 1985; Rabbitt, 1991; 1983; 1990; 1993a; 1993b; 1993c; Rabbitt, Donlan, Bent, McInnes & Abson, 1993; Rediess & Caine, 1996; Rowe & Kahn, 1987; Schaie, 1983; 1985; 1988a; 1988b; 1990; Schaie & Willis, 1986; Sheppard, 1978; Spirduso & MacRae, 1990; Sterns & Miklos, 1995; Thomae, Lehr & Schmitz-Scherzer, 1981; Weintraub, Powell, & Whitla, 1994; Willis, 1985).

The core problem which is raised by these authors, is that diversity within aging cohorts tends to be obscured by a prevailing research milieu which focuses on central tendencies.

---

[1] This is the name used by the present author for publication purposes. Thus all references to Shuttleworth-Jordan apply to the author's work.

As a consequence it is posed that a distorted picture of the aging process emerges which results in erroneous stereotypes about older people upon which public policy decisions are often based, and contributes to ageist discriminatory practices (see particularly Rowe & Kahn, 1987, Schaie, 1988a; Sheppard, 1978; Sterns & Miklos, 1995). When the heterogeneity of the older populations is focused on, a more optimistic view of aging is possible, in which the upper limits of potential in older adults can be acknowledged and gainfully employed. It is pointed out that the exploration of aging differences is a crucial "first step towards finding out what factors promote cognitive longevity" (Rabbitt, 1990, p.243). Consequently Rabbitt (1993b), goes so far as to suggest that when data on cognitive changes across age groups are examined, changes in the *variability* of performance are much more potentially informative, than changes in *average* scores. Thus several aging specialists appeal for a new paradigm in aging research calling for a *differential* rather than a normative gerontology (for example, Dannefer, 1988; Labouvie-Vief, 1985; Rowe & Kahn, 1987; Schaie, 1988; Thomae et al. 1981).

The reality is, however, as Nolan and Blass (1992, p.19) point out, that the neglect of the variability issue in aging has in the main "received only lip service". This is reflected in the fact that many seminal handbooks which deal with the psychological aspects of aging (for example Birren & Schaie, 1977; 1985; 1990[2]; Boller & Grafman, 1991; Craik & Salthouse, 1992; Huppert, Brayne & O'Connor, 1994), do not include any chapter which covers the differential aspects of such aging processes as a topic *in its own right*, nor is the term 'variability' indexed in the majority of such texts. Throughout *The Handbook of Aging and Cognition* (Craik & Salthouse, 1992), the milieu of normative trends predominates to such an extent that mean scores illustrating cognitive aging trends are even presented without accompanying standard deviations (Craik & Jennings, 1992, p.58; and Park, 1992, p. 465); similarly leading groups of aging researchers have presented comparative cognitive aging data in the journal literature without variance data (for example, Heaton et al., 1986; Hinkin, Cummings, Van Gorp, Satz, Mitrushina, & Freeman, 1990). Commensurate with these observations from the present literature overview, are the findings of Krauss (1980) who

[2] At the time of this review the most recent version of the Birren and Schaie *Handbook of the Psychology of the Aging (4th Ed.)* was still in press, and thus was not available as a source. However, from the reference in Finkel et al. (1996) (see next annotation), it appears that even in this most recent Birren and Schaie handbook the issue of heterogeneity in aging only forms part of a very broad-ranging chapter, and is not dealt with fully as a topic in its own right.

reviewed all the articles in the first four issues of Volume 39 of the *Journal of Gerontology* and noted that "few of the articles examined included variance information, even for purely methodological reasons" (p. 547), and those of Bornstein and Smircina (1982), who in a review of the journal literature published in 1979 and 1980, noted that only 41.8% of the articles reported the variance and only 7.2% even discussed the variance in any sections of the article. In the view of those who promote a differential psychogerontology (for example, Schaie, 1988, p. 181), central tendencies should never be reported in isolation without the standard deviation figures, due to the marked heterogeneity of aging populations.

Rare departures from the general rule of neglect of the variability issue in the aging literature, include a metatheoretical chapter by Dannefer (1988), and the psychogerontological reviews of Krauss (1980) and Bornstein and Smircina (1982), which are devoted entirely to general discussions on the issue of variability in itself. This is not to ignore the fact that there are several instances in which the subject has been strongly addressed *within* the chapters of some of the psychology of aging handbooks as an important concern alongside the theme of the particular report (for example Birren & Schroots, in press, cited in Finkel, Pedersen, McClearn, Plomin & Berg, 1996[3]; Hartley, 1988; Plomin & McClearn, 1990; Poon, 1985; Rabbitt, 1993c). Moreover, it is possible to detect a noticeable trend of increased attention to the matter over the years for instance across the Birren & Schaie aging handbook series (Birren & Schaie, 1977; 1985; 1990). Overall, however, (with the exceptions of book chapters by Dannefer and Krauss, and a single three-page journal article of Bornstein and Smircina), the marked absence of publications dedicated *exclusively* to the topic of variability in older populations over the space of the last seven years, does not reflect any predominant paradigmatic shift, as yet, towards a more differential aging psychology. Dannefer's (1988) paper is limited to a *sociohistorical* discussion which attempts to unravel the reason for the neglect of variability as a feature in aging research. Dannefer arrives at the conclusion that this has occurred, not as an isolated anomaly, but as the outcome of the systematic tendencies inherent in the language and paradigmatic assumptions of normativity which dominate research on age. Dannefer's point is forcefully endorsed by the two psychogerontological reviews of Krauss (1980) and

---

[3] The full reference for this citation which appears in Finkel et al. (1996) is as follows: *Birren, J.E., & Schroots, J.J.F. (in press). History, concepts, and theory in the psychology of aging. In J.E. Birren and K.W. Schaie (eds.), Handbook of the psychology of aging (4th Edition).*

Bornstein and Smircina (1982), which ironically, however, are extremely limited in themselves. They each focus in somewhat scattered manner only on very narrow subsections of the literature.

Thus in the author's knowledge, no text exists which has comprehensively reviewed and systematically integrated the available theoretical and research literature on cognitive aging with an exclusive focus on the issue of inter-individual variability. A central feature of the present thesis, therefore, was to conduct such an outstanding theoretical review. In this thesis, the current status of research findings and methodological issues in cognitive aging derived largely within a normative paradigm, were examined anew and redefined from the angle of a differential psychology. Such a wide-ranging and intricate task was deemed necessary to provide the foundation against which to design a major research investigation into the area, and to evaluate and interpret the findings of such an investigation with any meaning. It is suggested that the differential perspective as promoted in the thesis for the science of cognitive aging, is necessary to *complement* rather than to dominate or replace the normative perspective. The important metatheoretical message that emerges from the work is that a science of cognitive aging, in which the differential and normative perspectives operate *in parallel*, will produce a more complete rendering of the situation than will follow in the wake of a science which focuses on either one of these aspects in isolation.

1.2.2 Research Perspective

The neglect of the study of individual differences in aging *in practice* as distinct from more 'ivory tower' concern, is further revealed by the extremely limited number of research studies which have systematically investigated this issue. There is a marked lack of research that actually includes statistical analyses to test the hypothesis of increased variability in older populations, and an absence of rigour in the interpretation of variability effects in gerontological studies. This has caused Montgomery and Borgatta (1986) to suggest that a 'doctrine' of increasing variability among adults with age has been adopted in the aging literature on extremely flimsy empirical foundations, and could quite possibly be yet another one of the erroneous myths that are attached to aging! These authors' provocative comments, however, are based almost entirely on the very cursory review of Bornstein and Smircina (1982), and in the light of more recent research endeavours and gradually accumulating evidence, it appears highly *unlikely* that the presence of increased heterogeneity in the aging population is entirely mythical. Nevertheless, there certainly are glaring indications for a much less global and more differentiated understanding of its occurrence via further precise investigation.

At the least systematic level of analysis, studies can be found in the psychogerontological literature which explicitly comment on the nature of overlapping trends between older and younger age groups, or the between-group variability, on cognitive task performance reported in their aging studies (for example, Albert, Duffy, & Naeser, 1987; Beatty, 1993; Cornfield & Shuttleworth-Jordan, 1996; Mitrushina & Satz, 1991a; Mitrushina & Satz, 1995; Shuttleworth-Jordan & Bode, 1995a; Van Gorp, Satz, Kiersch, & Henry, 1986). However, typically in such studies, few have examined variability differences via statistical analyses, and/or variability has only been of extremely marginal concern in all of these reports which have had a separate main focus. Thus the issue has not been taken much further than brief commentary on, and limited explanation for, the marked heterogeneity noted in their aging groups.

A further series of researchers have reported on variability trends in older age groups at more length (Bornstein & Smircina, 1982; Christensen, Mackinnon, Jorm, Henderson, Scott, & Korten, 1994; Horn & Hofer, 1992; Maddox & Douglas, 1974; Morse, 1993; Poitrenaud, Barrere, Darcet & Driss, 1983; Rabbitt, 1990; 1993a; 1993b; 1993c; Rabbitt et al., 1993; Riegel, Riegel, & Meyer, 1967; Rush, Panek & Russell, 1990; Salthouse, 1991; Schaie, 1983; 1988a; 1988b; 1990; Thomae et al. 1981; Willis 1985; Weintraub et al., 1994). In some instances these reports have provided brief reviews of studies in the area, in some cases graphical representations of variability shifts across age groups for a variety of functional modalities, and in some instances fairly elaborated explanations for variability patterns in terms of social, psychological and physiological factors. For the most part in these reports, *speculative* discussions around the reasons for increased heterogeneity have been presented. In a limited number of cases possible connections between inter-individual variations and psycho-social and physiological mechanisms have been explored more formally via correlational analyses. From these studies it was found that cognitive performance varies on the basis of a complex pattern of demographic, personality and health variables (Bolla-Wilson, & Bleecker, 1986; Poitrenaud et al., 1983; Rush et al, 1990; Schaie, 1983; Thomae et al, 1981). Many of the above-cited reports have based their commentary on the variability findings noted in large, broad-ranging aging studies (for example Horn & Hofer, 1992; Riegel et al., 1967; Salthouse, 1991; Schaie, 1983), which were not designed with the prime purpose of investigating variability. In only a very few instances did the phenomenon of inter-individual variability across a number of older adult groups serve as the *main* focus of the investigations (Christensen et al., 1994; Maddox & Douglas, 1974; Morse, 1993; Rabbitt, 1993c; Thomae et al. 1981; Weintraub et al., 1994).

Taken together, all the available research investigations into variability noted above reveal the complexity of between-group variability patterns, which differ between cognitive modalities, which at times are different between studies for the same cognitive modality, and which do not always increase in linear fashion with increasing age. All the authors listed above (p.6, para.2), with the exception of Bornstein & Smircina (1982) and Salthouse (1991), present reviews that *support* the proposition that in general there is increasing heterogeneity with age. Bornstein and Smircina (1982) render a provocative, albeit brief, report which challenges the hypothesis of increased variability in aging, suggesting that this is a "doctrine" about aging which is based on marginal empirical evidence, and cite research which suggests that in many instances variability remains stable and at times even diminishes in older populations. However, whilst Bornstein and Smircina have been critical about the absence of statistical analyses of differences in variability data within gerontological studies, they go on to support their own argument with a review of variance within twenty-three studies via an informal "eyeballing" (sic) of the reported variability data. Salthouse's (1991) analysis of the situation is similarly brief and without statistical analysis. Acting, it appears, on the assumption that variability trends should be linear and identical across different studies and for all functional modalities, Salthouse does not endeavour to make sense out of the complicated graphical material he presents.

From a further set of more differentiated analyses than these, a number of consistent trends, as well as contradictory findings from research studies on variability trends within specific functional modalities have been reported. For example, whereas most studies indicate that variability increases with older age on 'fluid' tasks, and remains more stable with 'crystallized tasks', Horn and Hofer (1992) report on the large unpublished study of McArdle which identified a totally opposite trend. Generally, when reference has been made to increases in variability across functions in older groups, a *linear* increase has been noted or implied. Maddox and Douglas (1974), however, have presented the more differentiated observation that whilst variability in social and psychological social and psychological functioning may be accentuated at the beginning of old age, it is possible that de-differentiation may occur in the later years. Consistent with this observation, it has been noted with respect to variability data from the Seattle Longitudinal Study, that on some functions an initial increase in variability occurs with older age, followed by a sudden sharp decline in the older old age groups (Cunningham & Tomer, 1990; Schaie, 1983; Willis, 1985).

Standing out in the literature from the perspective of this thesis, are two very recent studies (Morse, 1993; Rabbitt, 1993c), in which variability has been systematically examined as the key focus of the investigations. Morse (1993) provides an isolated example of *an archival meta-analysis of variability trends via a statistical analysis of differences in variance between young and old age groups across a large spectrum of studies extracted from the cognitive aging literature*. Methodologically, as pointed out by Craik & Jennings, (1992), a meta-analysis of this type has heuristic power at a level above that of the individual study. This is due to its wide-ranging overview with respect to aging data available in the literature, thus facilitating the search for patterns of effects. Morse's wide-ranging *between-groups* study of data available in the literature broadly corresponds with the form of investigation conducted as the *first* phase of the present study. Rabbitt (1993c), also appearing as a solitary example in the literature, *plots the distribution of individual scores within older age groups, and makes comparisons across groups with respect to the characteristics of the outliers*. Methodologically this study has particularly interesting heuristic potential for the examination of precisely those individuals that are lost within any form of group statistics, and provides a prototype for investigation from a differential perspective. Thus Rabbitt's *within-groups* form of investigation, with a particular focus on the characteristics of *outliers*, is broadly commensurate with the type of investigation pursued in the *second* phase of the present study.

However, the studies of Morse (1993) and Rabbitt (1993c) both come from within the cognitive psychology paradigm of studies in aging, and their analyses are restricted to relatively narrow spectrums with regard to age. (In the milieu of clinical neuropsychology by contrast, normative adult aging data tend to cover a fairly wide and differentiated spectrum of age groups, from the 20's through to the 80's age decades, since the objective of these data is to provide reference points for clinical diagnosis in different age groups rather than to elucidate fine cognitive mechanisms). The age comparisons made by Morse (1993) are confined to the bimodal distribution of 'young' versus 'old' age groups, and those of Rabbitt (1993c) to the older age groups of 50-59, 60-69 and 70^{+}. This severely limits the observations these researchers can make about trends across the broad spectrum of the adult age group, and in particular would fail to distinguish changes within the older old-age groups, for example 70-79 versus 80-89. The examination of outliers by Rabbitt (1993c) was in any case narrowly focused with the purpose of contesting a hypothesis from Salthouse's (1985a; 1985b) single factor model of aging. Thus his objective was purely to identify outlying subjects with a dissociation between IQ and memory functioning, and he was not concerned with a more *general* examination of subject heterogeneity via his raw score scattergrams.

In summary, an overall critical appraisal of the existing research on variability reveals that it has tended to be very scattered, with a minimal number of studies which set out systematically to investigate the issue. Discussions about the overall nature of trends and conflicting findings have been limited and vague, and the absence of any major attempt to order available findings is very apparent. There is no broad-based archival study into patterns of variability for a range of cognitive modalities, across a spectrum of age groups which is more differentiated than merely the bimodal 'young' versus 'old' age dimension. There is no research which links an initial analysis of broad variability trends across adult age groups with an in depth analysis of individual raw score scatter, including the pattern and proportion of *outliers*, in order to elucidate the meaning of the identified trends, and to consider the influence of variables such as gender, educational and IQ level in these trends. Whilst the studies of Morse (1993) and Rabbitt (1993c), provide promising indications of future methodological directions to take in order to investigate heterogeneity in aging populations, it would appear that they have only made a bare beginning in terms of the rich heuristic potential available in such approaches.

1.3 AGING AND NEUROPSYCHOLOGY

As keynote speaker in a prestigious memorial lecture, Patrick Rabbitt recently made the point (with refererence to data from large-scale cross-sectional studies), that cognitive skills do not "all go together when they go" (1993b, p. 385). He suggests that there may be characteristic patterns, or *syndromes*, of cognitive aging. He confirms that many important questions remain unanswered, such as whether or not such 'syndromes' of cognitive aging exist, and the extent to which other variables such as educational level, intellectual level, illness, interact to produce inter-individual variability in cognitive patterns of advanced aging. He suggests that an approach to understanding characteristic age-associated patterns would be to relate differences in levels of performance across functions (such as "language functions", "visuo-spatial ability", "memory", and so on), to the neurophysiological indications that support them. He expresses his surprise that this challenge has not been pursued more actively in the field of neuropsychology. Craik and Jennings (1992), in a comment on future directions in aging research, similarly suggest that alongside traditional cognitive approaches, useful data could be generated from the neurophysiological syndrome-analysis angle of neuropsychology.

In recent years there has been a surge of normative data collected for a broad spectrum of commonly employed single neuropsychological tests, on the basis of cross-sectional studies across a wide range of adult age groups (for example, Albert et al., 1987; Beatty, 1993; Cornfield & Shuttleworth-Jordan, 1996; Shuttleworth-Jordan & Bode, 1995b; Van Gorp et al., 1986). In addition extensive collations of such disparate sets of normative data, often reported by decade across the entire 20 to 80-year-old age range, have been achieved and noted for diagnostic use in clinical neuropsychological practice (Lezak, 1983 and 1995; Spreen & Strauss, 1991)[4]. Since the goal of these normative data collections has been for pragmatic clinical purposes, as a body of data it has received marginal attention from the perspective of conceptual implications for aging processes. Further the focus in the presentation of these data has been on *average* scores across the age groups, and minimal attention has been paid to the crucial issue of differences in *inter-individual variability* across the age groups. The pattern of inter-individual variability has implications with respect to the *consistency* of presentation amongst individuals of characteristic 'syndromes' of the aging process as suggested by Rabbitt (see above). In addition it has indications for the clinical use of normative data since the reliability of the average score diminishes with increases in variability.

---

[4] The most recent collation of neuropsychological data is that of D'Elia, Boone & Mitrushina (in press) which is referenced in Lezak (1995). However this book had not appeared in print at the time of the present investigation, and thus was not available as a data source.

Thus, taking up the challenge of Rabbitt (1993b) and Craik and Jennings (1992) about the potential for neuropsychological investigations in future aging research, and making use of the large existing normative data base almost entirely neglected from the differential perspective, it was decided to design a study on inter-individual variability within a *neuropsychological framework.* In addition, *brain reserve capacity threshold theory* as formulated by Satz (1993) from a significant body of neuropsychological research, and which has major implications with respect to the differential presentation of functional aspects of aging, provided an explanatory theoretical context for the present study. Empirical indications from the literature review were used to develop an explicit *shuttle model of variability,* which was conceptually consistent with the postulates of the Satz brain reserve capacity threshold theory, and from which hypotheses for further verification were posed. In order to test these hypotheses an investigation was implemented to examine a selection of tests within the modalities of *attention, memory, language skills, visual skills, problem solving* and *motor activity*, from a broad range of existing normative data studies in the field of clinical neuropsychology, on age groups ranging from 20 through to 90 years.

The specific objectives of the research, which was divided into two phases, were as follows:

(i) a statistical comparison of differences between standard deviations was conducted to identify which tests show significant differences in *group* variability data between age groups, at which point alterations in variability occur across adult age groups, and to identify patterns of between-groups variability across functional domains so as to consider the implications of such patterns for the normal aging process;

(ii) the distributions of raw scores within age groups were plotted in order to examine the *individual* patterns of variability within age groups on a variety of tests, and the contribution of educational/IQ level, and gender to the inter-individual variability within-groups.

1.4 STRUCTURE OF THE THESIS

The thesis is bound in two volumes: *Volume 1* contains the theoretical framework of the thesis (Chapters 1 to 5); *Volume 2* comprises the empirical investigation (Chapters 6 to 9), Notes, References, and the Appendix. For ease of reference *all sections, tables and figures* *are prefixed with the number of the chapter to which they belong*. Thus, for example, Table 4.1 and Figure 4.2 are the first table and second figure, respectively, to appear in Chapter 4. Tables and figures always follow ahead of the point at which the are first mentioned. Mostly they appear at the next most convenient opportunity, or they are grouped together at the end of the subsection to which they pertain.

Chapters 2 to 5 (Volume 1) elaborate on the issue of inter-individual variability with specific reference to research methodology and conceptualizations of normal aging. The purpose of these chapters was to present a detailed review of the present position on the problem of inter-individual variability as it exists in the normal aging literature (a task which does not appear to have been conducted up to now), and further to provide indications for its enhanced portrayal within methodological and theoretical concerns in normal aging. In this sense these chapters can be seen to comprise a theoretical thesis on the subject of heterogeneity in cognitive aging. At the same time each subsection was structured so as to present essential aspects of the methodological and theoretical rationale on which the empirical aspect of the thesis was based. Specifically, Chapter 2 presents a delineation of the concepts of 'variability' and 'normal aging'; Chapter 3 covers existing theoretical approaches in cognitive aging, and the hypothetical and research implications that emerge from these perspectives with respect to the issue of inter-individual variability; Chapter 4 is devoted to a review of the research into inter-individual variability, and the reasons for heterogeneity in cognitive aging. *In particular, major variability studies including those of Salthouse, Schaie, Rabbitt, and Morse, which are repeatedly referred to in broad-ranging commentary in the Introduction, are critically examined in depth (Section 4.2.1)*; Chapter 5 poses the shuttle model for the prediction of inter-individual variability in cognitive aging, deduced from the literature review, and from which hypotheses for the current investigation were posed. Chapters 6 to 9 (Volume 2) constitute the Methodology, the Results and Discussion sections of the present research study itself, which served to test the predictions of the model.

CHAPTER 2: DEFINITION OF CONCEPTS

The purpose of Chapter 2 is to initiate a conceptual basis for the present research study, and to provide a methodological perspective on the issue of inter-individual variability in relation to the existing cognitive aging research milieu. Thus concepts which are central to the study are described and the issue of inter-individual variability is elaborated on within the psychometric and aging literature. The chapter is divided up into two main sections, each of which serves as an essential building block in the overall rationale for the study as follows: 'The concept of Inter-individual Variability' and 'The concept of Normal Aging'.

2.1 THE CONCEPT OF INTER-INDIVIDUAL VARIABILITY

> *"Aggregation of individual scores is inappropriate when there is sufficient individual within-group disagreement".*
> *(Verran, Mark and Lamb, 1992, p.240. Examination of psychometric instruments using aggregated data).*

> *"If age categories differ systematically in the amount of variability on a given characteristic, ...then it is a systematic distortion of the experience of age in a population to focus only on normative and central-tendency measures, as is usually done........ What is needed is a conceptual framework that recognizes......the differentiation within a cohort."*
> *(Dannefer, 1988, pp. 359; 368. What's in a name?: An account of the neglect of variability in the study of aging).*

In this section the concept of inter-individual variability is delineated in terms of its use in this study. The discussion includes psychometric implications of inter-individual variability in general, and then its relevance specifically with respect to cognitive aging studies. Further the difference between terms such as intra-individual, inter-individual, and between-groups variability is clarified, and the relationship between them discussed.

2.1.1 Definition of variability

In statistical terms inter-individual variability refers to the extent of irregularity which occurs amongst the individual raw scores of a group of individuals, in relation to the mean of the data set. Frequently what occurs in the assessment of an aspect of cognitive functioning is that *measurement* takes place at the *individual* level, but *analysis* of the data takes place at the *group* level via the aggregation of individual responses and the derivation of a central tendencies (mean) score. Consequently, in order to compare an element of cognitive performance across different age groups, this occurs via a comparison of *mean scores* which are the scores considered to be most representative of the comparative groups. The extent to which the individual raw scores vary about the mean score constitutes the *inter-individual variability*. This must be clearly distinguished from *intra-individual variability* which is the extent to which scores on a task may vary for a particular individual at different periods of time. Intra-individual variability is intrinsically linked to inter-individual variability in that at any point in time an individual's score will be representative of one score from within a whole range of his/her own possible scores (Nesselroade, 1991a; 1991b; Nesselroade, & Featherman, 1991; Wohlwill, 1973). This is a subject that will be taken up more fully in Chapter 4 in a section in which artifactual reasons for increased inter-individual variability are discussed (see section 4.2.1.3). However, the concept is raised here in the context of defining terminology, and in order to note that *for the purposes of this thesis the focus is on <u>inter-individual variability</u>, and unless otherwise specified, when the term 'variability' is used in isolation, it will be always on the understanding that it refers to 'inter-individual variability'.*

The inter-individual variability of a group (that is, variability *within* a group), is for most purposes most precisely depicted via the calculation of a *standard deviation* (Anastasi, 1958). The standard deviation (SD) represents the extent to which the score of each individual deviates from the mean, and is computed by calculating the difference between each person's score and the group mean. These differences are squared and the squares are averaged; the standard deviation is the square root of this mean squared difference. Such calculations of inter-individual *within-groups* variability in the form of standard deviations, can then be compared across groups (that is, *between* groups), and this is referred to as the *between-groups variability*. Anastasi (1958) notes that meaningful comparative statements can be made about the extent of variability across groups within a specified frame of reference, such as for different sets of scores achieved on a particular psychometric test. This can be achieved informally by perusing trends (for example, Bornstein & Smircina, 1982; Wohlwill, 1973), or by means of formal statistical analysis (for example Anastasi, 1985; Bode & Shuttleworth-Jordan, 1993; Cornfield & Shuttleworth-Jordan, 1996; Morse, 1993; Shuttleworth-Jordan & Saayman, 1989). The crucial point that arises about between-groups comparisons, is that the utilization of individual level responses at a group level of analysis via mean scores, may result in erroneous conclusions under conditions where such an aggregation of scores is not justified due to high variability of scores around the mean (Anastasi, 1991; Verran, Mark, & Lamb, 1992). For the aggregated score to be reliable at the group level, within-group variability should be significantly less than the between-group variability (Verran et al., 1992.). Thus the more within-group variability increases, the less reliable (that is, representative) is the mean score for that group, and the less reliable are between-group comparisons. Verran et al. stress that it is important to make statistical checks on within-group variability as it relates to between-group variability in order to make valid statements about central tendencies across groups.

Lezak and Gray (1984) have highlighted the problem of large standard deviations reflecting the scattered distributions of scores commonly present amongst brain-damaged populations. With a number of case studies, they illustrate how parametric statistical comparisons between group mean scores accompanied by such inflated variability can submerge differences between means. They show how comparisons of such means using nonparametric evaluation, where the number of individual *events* are taken into account rather than just scores, prove to be significant. Thus via an emphasis on scores rather than individual occurrences, real differences can be obscured. Moreover, as Anastasi (1991, p. 61) points out, inter-individual variability "is an essential feature of human behavior" in itself, which should not simply be ignored as error. Clearly to the extent that variability is enhanced within a group, the more crucial it becomes as an element in its own right to be examined rather than overlooked. For example, "we may want to know what range of individual variation would be appropriate to consider for a particular purpose......Such questions deal with real, objectively assessed individual differences in performance. They are as close as you can get to empirically observed facts" (Anastasi, 1991, p.63).

The statistical concerns and suggestions reflected above (Anastasi, 1991; Lezak & Gray, 1984), are well illustrated by the study of Shuttleworth-Jordan and Saayman (1989) in which a comparison was made of two different modes of therapeutic intervention. Pre- and post-treatment change scores in anxiety showed no difference on a mean score comparison of intervention effects. However on statistical analysis a highly significant difference in the variability of scores between the two groups was noted. On further examination of the raw score distribution it was noted that one method yielded a moderate reduction in anxiety levels *for all subjects*; the other method revealed a *bimodal distribution*, with some subjects showing raised levels of anxiety, (one subject in particular had a disproportionately raised level of anxiety), and the other subjects showing reduced levels of anxiety. In the search for psychotherapeutic treatment modalities which are relatively safe, and will not have implosively deleterious effects, the examination of the data in this way was highly illuminating. The variation in *numbers* of potential casualties between methods, rather the mere aggregation of positive effects, is a crucial element of information.

The important point that arises out of the above research example, is that highly artifactual inferences can occur when individual differences within treatment groups are treated as error variance. Hoyer (1974) has emphasized this issue and drawn attention to a hypothetical experiment designed by Eckstrand (1962), with the purpose of illustrating the confounding of between subjects treatment comparisons due to individualized treatment effects. Eckstrand proposes a study to assess the effects of three treatments on a measure of learning efficiency. For certain individuals one treatment was best, and for others an alternative treatment was more effective. He illustrates how with four possible types of individual which are randomly assigned to each of the three treatment conditions, no significant treatment effects would be obtained in spite of large individual differences in treatment effects. On this basis to conclude that the treatment conditions were equally effective across all individuals is totally misleading. Hoyer highlights the fact that such fallacious conclusions occur as a consequence of operation from within a narrow scientific paradigm which overlooks human complexity and diversity. He is highly critical of the fact that in all psychological research with the exception of operant behaviourism, individual variability is "usually treated, statistically and conceptually, as error" (Hoyer, 1974, p. 821). He calls for studies in which the unique characteristics of the individual are taken into account.

In that older adult populations, as with brain-damaged populations, are recognized for their heterogeneity over and above the variability in middle-aged and younger human beings (for example Schaie, 1998b; Sheppard, 1978), the above points have the utmost pertinence for investigations into aging (Hoyer, 1974). In particular this issue has been taken up by Rabbitt (for example, 1983; 1993c), and elaborated on in some detail in the cognitive aging research context. He notes on the basis of a study of over 6000 individuals aged from 50-96 on two different batteries of cognitive tests, that:"The most interesting insight into age changes that we obtain from screens of such large populations is not the rate at which average levels of performance change across age decades but the enormous increase in variability in performance between individuals on the same tasks.......with the ages of the cohorts that we study" (Rabbitt, 1993c, p.11). Rabbitt points out that there is a progressive skewing of distribution with age, in the direction of a bimodal apportionment of scores, in which a diminishing number of individuals can still achieve maximum scores whilst an increasing number fall towards the lower end of the distribution. He illustrates (Rabbitt, 1983) how such a bimodal distribution of a significant number of extreme high and disproportionately low scorers may artificially be reflected in the group mean score as an overall, moderate decline in functioning which obscures the true nature of events.

In the interpretation of findings of inflated variability, or differences in variability across groups, Rabbitt raises two important points to be taken into account. The first is that consideration must be given to the possible artifacts of *floor* or *ceiling* effects on a cognitive task, a point that Salthouse (1985, p.407) has also raised. Salthouse cautions strongly about 'fruitless speculations' being developed around reasons for changes in variability if this potential artifact is not first excluded. A ceiling effect on a task, evident in a younger more proficient population, might be absent in an older population resulting in increased variability in the older population; a floor effect might produce the opposite effect of diminished variability in older populations; either ceiling or floor effects might obscure differences in in variability between the two populations. Ideally as Rabbitt (1993c) points out, an efficient measure of cognitive differences in an aging population must have both a high ceiling, and a very low floor in order to allow both very able and very frail individuals to be assessed along the same continuum of performance, which is a difficult constraint to meet. However, the possibility for such effects must be addressed in the consideration of variability effects. In the light of this problem, Rabbitt makes the point that for their sample genuine increases in variability are confirmed on a number of tests for which the possibility of floor or ceiling effects have been ruled out.

The second point to be aware of in the interpretation of variability trends, is that of subject bias. Rabbitt (1993c) cautions about the fact that selection bias towards the atypically mobile, motivated, well-educated and intelligent older people becomes more severe with increasing sample age, and this will have consequences on the variability measure, and its generalizability to more representative populations. However, as in the case of the relatively elitist, high functioning cohort he is describing, Rabbitt notes that this particular bias is likely to diminish rather increase variability. Thus he suggests that the marked age-related increases in variation found with respect to his cohort probably underestimates an even higher true variation in the general population.

Finally, Rabbitt (1993c) to great effect, plotted the scores of individuals in three older age groups, and used a comparison of the *number* of individuals who varied more than two standard deviations below the mean between the three groups in terms of a particular constellation of cognitive functions, in order to inform theoretical postulations on cognitive aging. Thereby he was basically operating within the non-parametric psychometric framework recommended by Anastasi (as outlined above), for when comparisons are made between groups with high variability. In this mode the *number* of individual events is taken into account in preference to the central tendency of ordinal scores, and may have more heuristic value when comparisons are made on groups with skewed distributions such as occur in progressively aging populations.

2.1.2 The concept of inter-individual variability: Section summary

In concluding this section, the central points that have emerged are as follows:

(i) Variability between individuals implies behavioural facts that need to be scientifically investigated and not ignored as error, and where variability is particularly enhanced as in aging populations this point becomes even more relevant;

(ii) Between-groups variability can be examined by a comparison of the standard deviations across groups;

(iii) Within-groups variability can be examined by plotting the distribution of individual scores and by paying attention to the number of, and characteristics of, the extreme scorers;

(iv) In the interpretation of high inter-individual variability within a group, or differences in inter-individual variability between groups, account must be taken of the possibility for floor or ceiling effects on cognitive tasks which may artificially distort estimates of variability, and the influence of restricted sampling which would be likely to diminish the amount of variability true for the general population.

2.2 THE CONCEPT OF NORMAL AGING

In this section the concept of normal aging is discussed and delineated in terms of its use in the present study. Broad definitions of aging are followed by a discussion of three related methodological issues in normal aging research design: (i) the use of chronological age as a variable, (ii) the concepts of normal age 'differences' as against 'changes' within cross-sectional and longitudinal research designs, and (iii) the notion of what constitutes a generally representative sample in normal aging. In each case the issue is discussed in terms of its relationship with the problem of inter-individual variability, and implications for the design of the present study are elaborated.

2.2.1 Definition of normal aging

The literature reflects considerable difficulty with arriving at a precise definition of aging (Birren & Birren, 1990; Birren & Cunningham, 1985; Ingram, 1991; Kirkwood, 1994; Plomin & McClearn, 1990). Birren and Birren (1990) note that there has been a reluctance amongst aging researchers to define aging and suggest that this may be a consequence of an emphasis on operational rather than conceptual rigour. From a discussion by Birren and Cunningham (1985) it would appear that part of the problem in arriving at clear definition, is that such delineation depends very much on the perspective taken on the aging concept, and definitions vary considerably depending upon whether it is viewed as a phenomenon from a social, biological or psychological angle.

However, it can be argued that the concept of aging is so broad that definitions are only likely to derive meaning from being narrowed down to a particular perspective, and operationalized to an extent. Thus based on the discussion presented by Birren and Cunningham (1985), *biological aging* for instance can be conceptualized as a process of deterioration due to a number of essentially irreversible physiological changes with a focus on the ultimate end point of death as the dependent variable, whereas *psychological aging* needs a broader definition in which environmental factors are given more emphasis, and in which recognition is given to the fact that there can be incremental as well as decremental changes across the adult years, and it is these changes which serve as the dependent variables.

Implicit in any definition of 'normal' psychological aging however, is the assumption of a *typical, characteristic* pattern of change (Birren & Birren, 1990), in other words the assumption of *normativity*. Birren and Cunningham (1985) suggest that the following definition of Birren and Renner (1977) is a suitable one from the slant of the behavioural (psychological) sciences, in that it encompasses the possibility for both incremental and decremental occurrences as part of the regular aging process:

> "Aging refers to the regular changes that occur in mature genetically representative organisms living under representative environmental conditions". (Birren & Renner, 1977, p.4).

In a paper devoted to concepts and models of the psychology of aging, a similar definition is opted for by Birren and Birren (1990), but which has been extended so as to emphasize the notion of changes across time:

> "Aging .....refers to an orderly or regular transformation with time of representative organisms living under representative environments." (Birren, 1988, p. 160, cited in Birren and Birren, 1990).

These two Birren definitions are supposedly attempting to define 'normal aging' as in 'normal' populations and environments, whereas in fact the way in which they are phrased does not strictly do this. Within the context of the above definitions, 'representative' is still open to definition, in that for instance the definitions can encompass aging as it occurs regularly in an institutionalized, brain-damage population, or aging as it occurs normally in a population that has been exposed to war conditions. Thus in order to restrict the definition to 'normal aging' as in normal populations and environments, it would add clarity to extend the adjectival description of 'representative' to '*generally* representative'. Furthermore, the problem with both the above definitions from the viewpoint being promoted in this thesis, is that the emphasis is laid heavily on *normative* and *regular* processes. Thus, in terms of the sociohistorical perspective promoted by Dannefer (1988) on the neglect of variability in aging research, this definition exemplifies the kind of language which is inspired from within a normativity research paradigm. As Dannefer has proposed, such discourse has a powerful influence on scientific attitudes, and consequently serves to maintain the prevailing prototype of scientific endeavour with an overriding focus on central tendencies. Clearly in the light of the acknowledged heterogeneity of aging changes, any adequate definition of the phenomenon should reflect not only the intrinsic normativity but also the *variability* in the aging process. These definitions, therefore, could usefully be modified to provide a much more dynamic emphasis which is informed from a differential perspective on aging as follows: (Emphasis is added to the aspects of language which convey the implication that there will be variation as part of the regular, normally occurring aging process).

> *Aging refers to the differential changes that occur regularly with time across mature, generally representative organisms in interaction with generally representative environmental conditions.*

In the adoption of the above definition as the underlying standard for this thesis, there are three key notions that require further elaboration in a research context: (i) the assumption of chronological age as the independent variable, (ii) the idea of what constitutes normal aging 'changes', and (iii) what is understood by the word 'generally representative' in relation to environmental conditions and the human organism within the aging process.

2.2.1.1 Chronological Age as the independent variable

Inherent in the concept of normal aging is the notion of time, in that the concern is with differences which occur, or patterns which are typical in people of different ages, at different periods in time. (The complex metatheoretical aspect of time in relation to aging has been addressed in some depth in the aging literature - see Birren & Cunningham, 1985, Schroots & Birren, 1990 - and will not be dealt with here, since it is tangential to the focus of this thesis). From on operational perspective, the most immediately obvious measure of time in human aging processes is that of chronological age. Consequently most studies of aging use chronological age as the central variable of interest, and most frequently as the independent variable from which point the presence or influence of other variables are considered. However, considerable dissatisfaction has been expressed with this practice (see reviews in Birren and Cunningham 1985; Schroots and Birren, 1990). Complex methodological pitfalls reside in the relativity of the phenomenon of chronological age with respect to sociohistorical factors, and also in that conceptually this variable is devoid of fundamental explanatory power. A serious related problem is the distorted stereotyping about aging which arises when chronological age is used as an undifferentiated yardstick and given explanatory power which it does not rightly have (Schaie, 1988a; Rowe & Kahn, 1987; Sheppard, 1978).

Thus it has been posed (see reviews on the topic in Birren & Cunningham, 1985; Schroots & Birren, 1990), that the concept of chronological aging in isolation should be replaced by a more differentiated notion of aging including 'biological aging' which is the individuals' position in relation to their potential life span in terms of vital and life-limiting organ systems; 'social aging' which refers to the roles and habits of individuals in relation to other members of the society; and 'psychological aging' which encompasses the behavioural capacities of individuals in relation to the environment. The important point is that these three may be relatively independent of one another, and independent of chronological aging, and conceptualized in this way there is protection against time-based ageist stereotyping. It has been suggested that such constructs should supplement, or even replace, chronological age in aging research practices.

However, whilst more conceptually satisfying, the ramifications involved in establishing indices of aging along any one of the biological, social and psychological dimensions, are fraught with ambiguity in terms of definition and ambiguity of scale properties. As Schroots and Birren (1990) note, all three of these constructs are ultimately scaled on a physical time scale, and hence in the final analysis, do not escape from measurement in chronological time. Furthermore, the biological, social and psychological aging constructs, whilst not identical with chronological age, are all highly predictive of chronological age; similarly chronological age accounts for a high proportion of the variance in each of the three constructs. If for argument's sake any of either psychological, social, or biological aging were used as the independent variable in an attempt to investigate possible reasons for successful versus unsuccessful aging, the key confounding variable is likely to be chronological age which would have to be partialled out before any meaningful conclusions could be reached. In effect this would mean the need to match subjects on the basis of chronological age, and the inescapable return to its use as a baseline variable. Thus, whilst chronological age needs to be acknowledged as a limited construct in explanatory terms, which stands for an undifferentiated combination of possible effects and influences, it is the most pragmatically available index to use as the base point in the study of aging since we live in a time-related age-graded society (Schaie & Hertzog, 1985). Consequently, as pointed out by Birren & Cunningham (1990, p. 19) it is likely 'to remain as the "meat and potatoes" variable for the bulk of aging research'.

Implications for the present study. The present study was exclusively restricted to examination of studies which use the index of chronological age as the independent variable. However, the rationale for this was based on substantially more than the pragmatic issue raised above which has caused it to be retained in most aging research. It is evident from the literature, that the objection to the use of chronological age as the prime measure in aging research resides significantly, not only in the recognized heterogenous nature of the aging process itself in psycho-social and biological terms, but also in the *variable* presentation that occurs across these dimensions *between individuals*. In other words the whole difficulty relates to the matter of inter-individual variability, although the point does not seem to have been encapsulated as succinctly as that. At the root of the 'neglect of variability' issue, which is becoming increasingly highlighted in the aging literature, is the neglect of variability *precisely around the measure of chronological age.*

As noted above, one route out of the dilemma of this variability problem has been the inclination to throw out chronological age as the chief yardstick; however, in reality this has not emerged as an entirely pragmatic option. The alternative which is posed from the perspective of this thesis, is to develop a more systematically *differentiated* approach to a study of aging in which chronological age is *retained* as the base measure, but in which there is *systematic investigation of the variability in addition to the central trends* that occur in relation to that measure. Turned on its head in this way, chronological age becomes a robust and easily accessible blank screen against which the richness of the heterogeneity of the aging process can be thrown into relief for enhanced scrutiny, rather than a barren phenomenon with limited heuristic potential which serves only to restrict and distort understanding, and thus needs to be thrown out.

Against this conceptual background, the use of chronological age as the fundamental unit of measurement in the present research was fully justified as the scale of choice.

2.2.1.2 Age 'changes' versus age 'differences'

> *"Cross-sectional comparisons ....are not merely weak, expedient substitutes for proper experiments but essential to all gerontological investigations".*
> *(Rabbitt, 1983, p.12. How can we tell whether human performance is related to chronological age?)*

Inherent in the concept of normal aging, as defined above (see section 2.2.1), is the developmental concern of *change* over the maturing years. From an operational point of view in the aging research context, it is essential to differentiate between the concepts of 'age change' and 'age differences', two notions which from a methodological perspective are inextricably bound up with the disparity between longitudinal and cross-sectional research design. Consequently the advantages and disadvantages of these two broad research modalities are of central concern in the aging literature and have been discussed repeatedly and at length (Cunningham & Tomer, 1990; Birren & Cunningham, 1985; Finkel, Pedersen, McClearn, Plomin & Berg, 1996; Heaton, Grant & Mathews, 1986; Holland, 1990; Labouvie-Vief, 1985; Nesselroade & Labouvie, 1985; Mitrushina & Satz, 1991b; Nolan & Blass, 1992; Rabbitt, 1983; 1990; Rabbitt, Watson, Donlan, Bent, & McInnis, 1994; Reitan & Wolfson, 1995; Riegel, 1967; Rogosa, 1988; Salthouse, 1992a; Schaie, 1988a; 1988b; 1990; Schaie & Hertzog, 1985; Schroots & Birren, 1990; Willis, 1985).

The important points that emerge from the literature can be synthesized as follows. *Age changes* are the intra-individual shifts in the organization of behaviour which occur with age, and thus theoretically can only be directly assessed via longitudinal research studies in which the same individuals are repeatedly assessed over long periods of time. *Age differences* on the other hand can be identified intra-individually in repeated regular investigations over time as in longitudinal research, and can also be identified on the basis of cross-group comparisons of individuals of different ages, who are all investigated at the same period of time, as in cross-sectional research. However, the key point is that age *differences* established across groups via cross-sectional research, should not mistakenly be seen to be identical with intra-individual age *changes*, the nature of which can only cautiously be inferred from cross-sectional research since there may be significant variations in older people depending on the characteristics of their birth cohort, in addition to a host of additional unknown nonequivalent characteristics between the groups which are controlled for in a within-subjects design. For example, cross-sectional data have implied much earlier decrement in aspects of intellectual functioning such as inductive reasoning and verbal meaning, than has been indicated from longitudinal data, and it has been suggested that this can be attributed to cohort effects (Schaie, 1988b; Labouvie-Vief, 1985). The indication is that cross-sectional studies may tend to overestimate the extent of cognitive decline with age. In addition Finkel et al. (1996) specifically draw attention to the possibility of over-estimating the extent of increased variance in association with advancing age in cross-sectional studies due to the confounding factor of generational differences.

However, interpretations of longitudinal studies are also subject to *severe* methodological problems which reduce the extent to which aging changes can in fact be accurately measured in this research modality, and these have been extensively reported (see references in this subsection, paragraph 1 above). Whereas cross-sectional studies appear to *exaggerate* age-related change due to confounding generational cohort differences, longitudinal studies appear to *underrate* the extent of decremental change, and Cunningham and Tomer (1990), go so far as to suggest that with careful analysis which takes account of cohort effects, cross-sectional research may even constitute the lesser of two evils in terms of bias.

The well-recognized difficulties which are likely to contribute to *inflated* scores on cognitive tests across the years in longitudinal research are as follows: (i) there is the problem of selective attrition of the non-elite, which is that subjects that do not return for testing either because of illness, death, or poor cooperation, tend to be the 'less-elite', lower functioning individuals; in particular dropouts from longitudinal studies have been identified as more rigid, dogmatic, to have had a more negative attitude to life, fewer interests, lower FSIQ and Verbal IQ (Riegel, 1967), and lower scores on memory and psychomotor tests (Mitrushina & Satz, 1991b); (ii) repeated testing results in increasing familiarity and relaxation with initially unfamiliar test contexts and cumulative test practice effects; (iii) there is the phenomenon of statistical regression to the mean across two occasions of measurement (see Nesselroade & Labouvie, 1985); (iv) over long periods of time unwitting shifts in protocol occur and there is the problem of differential experience and motivation of staff resulting in instrumentation and intervention effects; (v) importantly, it has been emphasized (for example, Rabbitt, 1990; Salthouse, 1992a), that longitudinal studies are not immune to generational effects in that in many countries standards of nutrition, education, hygiene, and health care have steadily improved, and changes in the external environment may be associated with improvements which will significantly reduce the apparent rate of change with age; and finally (vi) it has been pointed out (Rabbitt, 1990) that the measures used to assess cognitive change with age, particularly IQ tests, acquire reduced validity over time in that the norms against which they were originally standardized become inappropriate due to increasing test sophistication in the cultural and educational environments, and this will produce artificially reduced indications of decremental change due to age.

From the perspective of this thesis, the important question to be addressed is the extent to which the methodological problems noted here with respect to longitudinal studies will not only flatten the curve of genuine cognitive decline with aging, *but may also result in a significant underestimation of the increase in inter-individual variability of cognitive performance with advancing age.* That this is highly likely to occur due to the effect of selective attrition of the less able has been emphasized by a number of researchers (Finkel et al., 1996; Rabbitt et al., 1994; Riegel, 1967). However, the manner in which the cumulative effects of longitudinal research might effect variability is not entirely clear-cut and there are indications for differential effects.

On the one hand it appears almost certain, for example, that the consequence of test familiarity and practice effects, statistical regression to the mean, enhanced sequential environmental effects, in addition to the effects of selective attrition of the non-elite, will compound and artificially create a much *reduced* impression of the variability of performance within older versus younger age groups than is truly the case. On the other hand it is not possible to rule out the possibility of distortions in the direction of artificially *increased* variability, in that over time subtle changes in intervention and instrumentation may serve to increase variability. Furthermore laboratory research on the cumulative effect of test practice (Anastasi, 1958)[5], reveals a U-shaped *non-linear* effect on variability with a decrease for the initial four to five test occasions, but a subsequent increase in variability on the later test occasions, implying that the end result of practice in longitudinal research with *many* trials may be one of increased variability. Finally disproportionately reduced numbers of subjects in the later longitudinal samples relative to the earlier samples may result in statistical distortions across age groups in the direction of increased variability. Generally, however, the proposed possible sources of inflated estimates of variability in longitudinal

---

[5] In a continuous practice experiment, Anastasi (1958, p.198) examined the alterations in variance across successive trials on a cross-section of cognitive tasks including cancellation, symbol-digit, vocabulary and hidden word tasks. She found a statistically significant *increase* in variability due the practice effect between the first and the twentieth trials on each task, and on this basis she poses that individuals tend to become more *dissimilar* following a period of practice. However, careful examination of her data indicates a U-shaped curvilinear effect, with a clear trend of *diminished* standard deviations in the initial five to six practice trials, with increases occurring only in the subsequent trials. Hypothetically, this research from Anastasi provides some empirical support for the probability that in longitudinal research for the first five or so consecutive test occasions dispersed over wide time periods, if there is any effect it is likely to be one of *reduced* inter-individual variability as a consequence of test familiarity such as is evident from the initial trials of Anastasi's laboratory data. Clearly, the higher the number of trials, the less valid it becomes to extract implications from Anastasi's data for longitudinal research. This is because her laboratory data are restricted in time period, and the same number and calibre of subjects are retained across the consecutive trials. By contrast in longitudinal research the numbers of subjects are reduced over time due to selective attrition of the non-elite.

research are of a very subtle nature, and possibly may not even apply (see footnote number 5). For instance it must be borne in mind that the laboratory sample of Anastasi would not have been subject to the problem of selective attrition of low functioning subjects which contributes to reduced variability. Overall, therefore, *it appears much more probable that longitudinal research would result in underestimates of increased heterogeneity in association with older age*, in particular due to the well-recognised, strongly confounding effect of selective attrition which would be likely to override other more trivial effects.

Possible routes through the dilemma of the methodological problems inherent in both cross-sectional and longitudinal designs are (i) to implement the cohort-sequential design of Schaie (described and evaluated in Nesselroade and Labouvie, 1985), which is essentially a longitudinal study carried out at the same age levels on multiple groups at different times of measurement, and is very time-consuming and expensive, or (ii) to use both cross-sectional and longitudinal methodologies in the same study as exemplified by the studies of Fromm and Holland (1989), and, Mitrushina and Satz (1991b), which is also time-consuming and expensive, or (iii) for different researchers to conduct cheaper and more simply yet carefully designed studies in one or other of the cross-sectional or longitudinal modalities, which are very carefully interpreted with due respect to the acknowledged potential confounding factors pertinent to the type of design, and for cumulative data across different research studies to be taken into account in the attempt to unravel the nature of normal aging.

As Nesselroade and Labouvie (1985), point out, the highly complex cross-sequential cohort-sequential design aims to obtain a more adequate description of behavioral change by identifying the sources of those changes in relation to maturational and environmental histories. However, unless sampling procedures for the sequential cohorts are very carefully selected, and contain sufficiently large numbers so as to eliminate cross-cohort bias, observed cohort differences will not indicate social change with anymore validity than cross-sectional age differences are valid indicators of age change. Thus Nesselroade and Labouvie state that because of poor sampling procedures they are "unconvinced that the application of sequential strategies has always yielded empirical descriptions that have greater internal and external validity than those of simple designs" (p.52). On the same point Rabbitt (1983) notes that longitudinal studies "have both high potential yield and high potential risk" (p. 12).

Heaton et al. (1986) provide a good example of the third research proposition listed above for research into normal aging - that is the use of carefully designed and interpreted studies in one or other of the cross-sectional or longitudinal modalities. These researchers have conducted a cross-sectional study in isolation, but with careful explication, cumulatively using evidence from other studies to enhance the understanding of their results and to provide compelling indications about the aging process. They concede that it is not legitimate for them to establish age changes on the basis of their data since they have not conducted a longitudinal study. However, they argue that all three age subgroups at each of their education levels had similar scores on the Wechsler Adult Intelligence Scale (WAIS)[6] Vocabulary and Information subtests, thus there did not appear to be differences between age-cohorts on crystallized intelligence. They argue that such generational differences would be more likely to occur for crystallized rather than fluid aspects of intelligence, and it was the fluid functions which showed the greatest age change in their study. In addition they argue that their results are commensurate with several other studies conducted 20 years previously, and that the norms of their test battery for younger people have been appropriate for decades. Hence they come to the conclusion that the age-effects found in their cross-sectional study do reflect the true extent of change on their test modalities with advancing age.

Thus to some extent it is probably true to say that longitudinal studies may have become endowed with a 'halo' effect which is not entirely warranted in relation to more simple cross-sectional studies which are frequently viewed as second best. This may be due to the fact that longitudinal studies are much more difficult to implement than cross-sectional studies, or because it has taken time to fully acknowledge the methodological pitfalls in longitudinal studies. Moreover, longitudinal developmental change is not the only research issue in question, and there are heuristic benefits which pertain *particularly* to cross-sectional studies which need to be emphasized (Cunningham & Tomer, 1990; Mitrushina & Satz, 1991b; Rabbitt, 1983; Reitan & Wolfson, 1995; Schaie, 1988b). Importantly, there are questions about aging which relate to *current* within- and between-groups differences, and patterns of

---

[6] The overall format of the WAIS (Wechsler Adult Intelligence Scale) as it was first published in 1955 is basically unchanged for other versions of the IQ test such as the WAIS-R (1981) updated edition, and the South African WAIS (1969) version. Thus for the purposes of general discussion of the Wechsler Adult Intelligence Scale subtests within this thesis, the term WAIS is used generically regardless of the specific version of the WAIS that was involved in any particular study.

cognitive functioning across age groups *at a specific point in time*, which are more suitably addressed in cross-sectional than longitudinal studies, and which will not be subject to the burdensome problems of selective attrition of the non-elite and test practice effects.

Firstly, cross-sectional analyses are very much needed to give us information on the characteristics of populations across different age groups at a particular stage in time, in order to *inform policy issues*, or *conduct clinical practices* (Reitan & Wolfson, 1995; Schaie, 1988b). Thus in terms of policy matters around employment or retirement for example, knowledge about the *present* level of functioning for particular older age groups is vital. For practice in psychological assessment, normative data on cognitive tests need to be constantly updated via cross-sectional studies so as to establish the differences in performance across age groups for *current* use. Secondly, there is the need for cross-sectional studies in that they can provide valuable elucidation on the *patterns* of performance on different tasks which are the same or different within or between age groups (Rabbitt, 1983). And, as has been aptly pointed out (Mitrushina & Satz, 1991b), for clinical diagnostic purposes it is important to establish this pattern on the *first* test performance across age groups which is only possible via cross-sectional research design.

Implications for the present study. The present study was restricted exclusively to cross-sectional data in both phases of the analysis. Based on the discussion outlined above, it was understood that in the first phase of the analysis, broad trends of variability differences across age groups would need to be carefully interpreted and not taken as support for genuine age changes due to the possible influence of cohort effects. As such this phase of the analysis was viewed as *an essential preliminary step* in the understanding of variability *differences* across adult age groups on the neuropsychological tests investigated, which would need to be followed up with further cumulative and balancing indications concerning age *changes* in variability from longitudinal research. In view of the fact that it is considered likely that longitudinal studies will be *particularly* prone to produce artifactual variability effects in older groups (see the argument presented earlier, pp. 26-7), it can be argued that cross-sectional analysis might well constitute the lesser of the two evils in terms of bias in a variability study. Thus in this initial phase of investigation into variability trends, it was considered that cross-sectional analysis was the method of choice.

To the extent that the purpose of the present study was to examine the *patterns* of inter-individual variability across functional modalities within age groups (as in the second phase of the study), and to the extent that the degree and nature of inter-individual variability on particular tests across age groups (to be identified in both phases of the study), would serve to inform *current clinical practice*, it was considered that the use of normative cross-sectional data as was done in this study, was the appropriate choice of methodology. This was based on the fact that the cross-sectional data would be representative of *current trends* across age groups, and the supposition that they would be *less likely* to distort the true nature of variability patterns as much as might occur with longitudinal data.

2.2.1.3 <u>The representative normal aging sample</u>

Explicit in the definition of normal aging arrived at above (see section 2.2.1), is that it is a process which occurs in 'generally representative individuals' within the context of 'generally representative environments'. From the perspective of variability, the issue of what constitutes 'generally representative individuals' is conceptually highly intricate and has significant implications for research in this area; hence it will be discussed at length. This will be followed by a brief delineation of the less complicated issue of 'representative environments'.

<u>*Representative individuals*</u>. Within the context of aging research the phrase 'generally representative individuals' has come to imply normal, *non-disease related* individuals - the population that Birren and Cunningham (1985) have identified as forming the focal thrust of current normal aging research. However, the extent to which this healthy older sector of the population can be taken as generally representative is debatable, and furthermore the establishment of what constitutes a normal non-disease related aging population raises a number of complex conceptual and practical issues.

Initially it is possible to question in itself, the usual practice of separating out *non-diseased elderly* in order to study normal aging, since the prevalence of dementing disorders, other cognitive impairments and physical disease increase so dramatically with age (Birren & Cunningham, 1985; Johansson, 1990). Accordingly, Johansson reports that the prevalence of dementia and cognitive impairment was in the range of 18% to 33% for 84-to-90-year-olds in a Swedish epidemiological study, and Birren and Cunningham state that the assumption of non-disease related normal aging "cheerfully ignores the high prevalence of chronic disease even in the community-dwelling elderly" (p. 21). Thus in order to present normal aging data on a fully representative elderly population it is arguable that such disease factors should not be considered extraneous. However, since there are a significant proportion of older people who are relatively disease free, and who would not be subject to a diagnosis of a dementing disorder, it does seem legitimate to study what Birren and Cunningham (1985) refer to as *primary* cognitive aging in such a population in order to examine the most representative non-disease related normal aging cognitive processes, as distinct from *secondary*, disease-related processes, on the understanding that neither of these processes on their own are generally representative of older populations in an epidemiological sense. Johansson makes the point that it is precisely *because* of the increasing variability in aging populations due to intervening pathology, that the division of the population in this manner is deemed necessary in aging research. In such a heterogenous context the presentation of a *general* aging pattern across the full spectrum of older people becomes extremely dubious. Further, in research on healthy cognitive aging processes that involves the comparison of younger with older adult age groups, the exclusion of older-age disease related factors that would be confounding is particularly pertinent.

Having established the legitimacy of studying a primary, healthy aging population to describe a normal aging process, the next consideration relates to the delineation of what constitutes the non disease-related aged population itself. In research practice, in order to establish such a healthy nonclinical sample, most normal aging studies include a combination of screening clinical interviews and screening neuropsychiatric instruments to exclude the presence of major physical illness, neurological disease (including in particular Alzheimer's dementia), and psychiatric disorder, and in *operational* terms this is relatively straightforward. However, *conceptually*, the clear definition of what comprises the normal healthy aging population is highly complex, and requires further discussion here. The difficulties are inherent in (i) the continuity/discontinuity debate with respect to Alzheimer's dementia and its relationship to normal aging, an issue which has been developed as the central theme of Huppert, Brayne and O'Connor's (1994) handbook on dementia and normal aging; (ii) 'pre-clinical' stages of dementia, and the phenomena of Benign Senescent Forgetfulness (BSF), Age-Associated Memory Impairment (AAMI), Age-Related Cognitive Decline (ARCD), and Mild Cognitive Impairment (MCI); and (iii) the tertiary stage of cognitive aging designated as 'terminal decline'.

(i) Continuity/Discontinuity debate in Alzheimer's Disease. The central problem raised here (and amply addressed in Huppert et al., 1994), is as to whether the brain changes involved in Alzheimer's dementia are *qualitatively* the same as those of normal aging but merely occur in greater *quantity* than in normal aging (the *continuity* model), or whether the brain changes in Alzheimer's dementia constitute qualitatively different markers representative of pathology - histological, biochemical, viral, behavioural - which separate it from normal aging (the *discontinuity* model). The issue is currently unresolved with some theorists leaning towards the continuity hypothesis (for example, Brayne & Calloway, 1988; Huppert & Brayne, 1994; Lishman 1994), and others, on the basis of suggested indications of discontinuities in physiological and cognitive processes, favouring the discontinuity theory (for example, Hodges, 1994; Mitrushina, Satz, Drebing, Van Gorp, Mathews, Harker & Chervinsky, 1994; Roth, 1994).

The fundamental lack of resolution in continuity versus non-continuity conceptualizations of Alzheimer's dementia, however, poses particularly interesting implications for the definition of a normal aging population in terms of the *variability* of that population (see discussions in Brayne, 1994; Roth, 1994). Were a continuous Alzheimer's/normal aging hypothesis to be confirmed, then it is possible to view Alzheimer's patients as merely the extreme lower end of a highly heterogenous normal aging range; on the other hand, if a firm biological marker of Alzheimer's was to be discovered, then a lot more individuals in the very early stages of the disease might be distinguished than is possible on neuropsychiatric diagnostic criteria at present, and this would result in a marked upward shift of the normal aging distribution, due to the exclusion of individuals from the lower end of the distribution. The consequence would be a significant reduction in the variability of such a normal aged population.

In current aging research, in order to define a normal non disease-related aging population for research investigations, the conceptual issues above are circumvented operationally via the use of well-defined clinical criteria (e.g., DSM-IV), or screening instruments such as the Mini Mental State Examination (Folstein, Folstein & McHugh, 1975), on the basis of which individuals can be identified as having dementia, and excluded from normal aging samples. However, in terms of conceptualizations of what constitutes the variability in normal older populations as elucidated above, it is important to be aware that implicit in this diagnostic exclusion process is fundamental adherence to the idea of Alzheimer's dementia as a *disease*, which implies in large part also, acceptance of the notion of a *discontinuity* between that process and normal aging.

(ii) 'Pre-clinical' dementia, BSF, AAMI, ARCD, and MCI. Of relevance to the selection of healthy normal aging samples, is the perennial problem of uncertainty about the extent to which a sample contains individuals in pre-clinical stages of a progressive dementia who are showing subtle signs of cognitive deficit, but who would not be excluded on the basis of the usual screening procedures. Obfuscating this differential further, is the fact that subtle signs of memory impairment are so common in healthy older age individuals who do *not* have a malignantly progressive dementia, that Kral (1962; 1966; 1970) has proposed the syndrome of Benign Senescent Forgetfulness (BSF) to distinguish older people who report more impaired memory than their peers in the absence of dementia. The National Institute of Mental Health (NIMH) has taken this notion further and developed the diagnostic category of Age-Associated Memory Impairment (AAMI) to characterize age-related memory decline (Crook, Bartus, Ferris, Whitehouse, Cohen, et al., 1986). The criteria define AAMI as comprising those individuals over 50 who have subjectively noticed declining memory function compared to their young adult years, perform at least one standard deviation below the mean level of memory test performance of young adults, and who are not suffering from dementia, depression or any other existing medical or psychiatric condition that could account for memory decline. Although AAMI, and particularly BSF, are problematic and variable in terms of their definition (see discussions in Henderson, 1994; Hindmarch, 1993; Howard, 1993; Huppert, 1991; Larrabee & Crook, 1994; O'Connor, 1994), the important point that arises here is that both of these proposed syndromes of memory impairment (AAMI and BSF) are fundamentally conceptualized to exist *as part of the normal aging process*.

Due largely to criticisms of AAMI criteria (see discussions in Ferris & Kluger, 1996; Rediess & Caine, 1996), it was not included in the DSM-IV as a diagnostic category. Instead in more generalized and simplified form it has appeared as a new DSM-IV V-code with the designation Age-Related Cognitive Decline (ARCD). The ARCD concept is broader than the notion of AAMI in that it provides the scope to cover areas of cognitive functioning that decline with age such as complex problem solving ability, in addition to memory. However, as with the notions of BSF and AAMI, the cognitive alterations specified for ARCD are explicitly not identified as part of a disease entity, but rather as being within the normal limits of an aging population. Hindmarch (1993) notes that AAMI affects 40% of those in their fifties, and 75% of those above 75 years old. An archival data analysis from Larrabee and Crook (1994) suggest that only 15% of subjects 80 years and older do not have AAMI. Kral's (1962) research on residents of a nursing home, revealed that mortality in BSF was only 38% compared to that of 62% in those with memory problems from discrete pathological processes, on the basis of which he concludes that BSF is part of the normal aging process. Thus, although it is acknowledged that individuals included in the categories of AAMI and BSF might actually be in the early stages of dementia in a small proportion of cases, subjects classed in either of these two categories are considered to show the memory decline associated with normal aging, and thus are not conventionally excluded from what is considered to be a healthy, non-disease-related, aged population. The corollary of this is that the *variability* of cognitive performance due to either AAMI or BSF, in addition to the more recently developed outgrowth of these concepts in the form of ARCD, is considered *continuous* with normal aging and thus must be described as essentially part and parcel of the normal aging process. This is in contrast to the variability due to a diagnosed dementia which is, as discussed above, conventionally considered *discontinuous* with normal aging, and thus is not deemed intrinsic to the normal aging distribution for the purposes of research.

Very recently Smith, Petersen, Parisi, Ivnik, Kokmen, Tangalos, and Waring (1996) have produced yet another category which they term Mild Cognitive Impairment (MCI), which is applied to a borderland area which falls between dementia and ARCD. This category of individuals, who do not show the functional impairment sufficient to qualify for a positive rating on dementia, are identified by having objective cognitive impairment as evidenced by scores > 1.5 SD's below the age appropriate mean. Whilst the isolation of this yet further category of compromised individuals has been strongly motivated by the search for an 'at-risk' group for dementia, Smith et al. (1996) report that only 55% of MCI cases converted to a full dementia syndrome on clinical follow-up. The MCI category of older adults clearly presents the most problematically grey area with respect to defining a normal aging population for research purposes. However it is still grouped together with the non-disease entities of AAMI and ARCD as a label that refers 'to the same putative group of individuals who present with mild cognitive changes, but have no significant decline in functional status' (Rediess & Caine, 1996, P. 109). Since the category has not yet been established as a clear precursor of dementia in nearly half of the follow-up cases, it would seem highly problematic at this stage to view MCI individuals as continuous with a dementing population, and to exclude this category of people from normal aging samples.

A clear conceptualization and acknowledgement of the above subtle continuity versus discontinuity issues and their relationship to the normal distribution and variability in the aging population is necessary, but often lacking, in the design and interpretation of normal aging research studies. Mitrushina, Patel, Satz, D'Elia and McConnell (1989), and Mitrushina, Satz, and Van Gorp (1989), have conducted studies comparing the patterns of cognitive impairment and changes in semantic memory processing between 'healthy' and 'nonpatient at-risk for Alzheimer's Disease' groups of 58- to 85-year-olds. In both studies, the healthy subjects were extracted from a well-functioning cohort of elderly subjects being used in a larger study of normal aging; the at-risk group were also selected from the large well-functioning cohort of elderly subjects, but on the basis of being the *outliers* who scored more than two standard deviations below the well-functioning group on two or more of the twelve core neuropsychological tests in the battery. In both studies the results indicated significant differences in the performance of the healthy and at-risk groups. A further study on the same basic normal elderly cohort (Mitrushina, Uchiyama, & Satz, 1995), explored the factor clusters of the subjects' performance on a battery of neuropsychological tests. On the basis of six clusters which were extracted these authors highlight *the extreme variability* in the pattern of responses of the subjects, and with a focus on the disproportionately low scorers, they suggest that three of these factors have distinct patterns of cognitive deficit expected with Alzheimer's Dementia, and are likely to represent preclinical stages of the dementing process.

Although not explicit, all three of the studies of the Mitrushina et al. studies are concerned with issues of variability, and consequently throw some important issues in this regard into relief. In the case of all three of the above analyses, the authors have taken an originally defined *normal* older adult cohort, and then on the basis of *extreme inter-individual variability* within this cohort, redefined the subjects as 'at-risk' subjects for Alzheimer's Dementia. This amounts to a reduction in the heterogenous structure of the original normal adult group. However, since the 'at-risk' subjects in these studies have not been followed up longitudinally for confirmation of the eventual emergence of Alzheimer's Disease, their 'at-risk' status, is purely hypothetical, and conceptually this presents an interesting issue:

When there is the specific intention to describe the nature of inter-individual variability *as it is observed to occur within a normal aging population*, do you -

(i) proceed as these researchers have done, and identify the outlying individuals or unusual cognitive patterns as aspects which need to be parcelled out of the originally defined normal aging population, and which are then redefined as something else such as early Alzheimer's Disease?

or rather, do you (in the absence of confirmatory evidence that outlying scores or distinctive patterns of cognitive performance are representative of the performance of subjects in the early stages of Alzheimer's Disease),

(ii) consider these findings to describe the normal aging picture in itself?

A study by Arsanow, Steffy, MacCrimmon and Cleghorn (1977) sheds further light on this issue. These authors have illustrated the heuristic potential of the identification of a subset of outliers on patterns of cognitive test performance within a group of 'at-risk' schizophrenic children, by means of a cluster analysis such as Mitrushina et al. (1995) have performed on their normal older adult group. However, Arsanow et al. use the identification of this outlying pattern to *enhance the characterization of the group within which the cluster was found* - that of the 'at-risk' schizophrenic children. They do not suggest that this subgroup may in some way erroneously represent the original group. Moreover they make the important observation that this sub-set (which was to be expected in a population in which only 10-15% subsequently become schizophrenic), accounted for the overall low group means. Accordingly, in a study such as that of Mitrushina et al. (1995), there is *significant* heuristic potential in the identification of an outlying group of individuals with an outlying decremental cognitive pattern in the way they have done. However, this should be seen as a first step in the analysis, and until subsequent investigation confirms that the noted inter-individual variation is due to something other than the normal aging process of the pre-defined non disease-related population, it would appear conceptually imperative, and more empirically rigorous, to retain and examine the outliers and heterogenous patterns as *intrinsic* to the originally-defined healthy, normal population. In terms of emergent indications of the extreme variability of the normal aging process (Schaie, 1983; 1988b; Schaie and Willis, 1986; Rabbitt, 1993a; 1993b), and the complexity of possible reasons for normal inter-individual variability in aging populations noted by these researchers (to be discussed in detail in Chapter 4), such differential factors become the *thrust* of analysis in such a population. Research rigour is particularly important in this regard so as to avoid the stereotyped, over-pessimistic and often erroneous myths that tend to occur in relation to aging (Montgomery & Bergatten, 1987; Schaie & Willis, 1986).

In the case of the three Mitrushina et al. studies under discussion, for example, the outlying low scorers and heterogenous cognitive patterns on further follow-up examination might conceivably emerge as due to host of other isolated or interactive reasons to account for cognitive decrement in addition to the possibility of early Alzheimer's dementia. For example, it is possible that there is a clustering of those individuals who have unidentified malignant physical illness and are in terminal decline, or more optimistic possibilities such as the presence of undiagnosed remediable physical illness such as diabetes, or reversible psychiatric illness such as depression (see Montgomery & Bergatten, 1987; O'Connor, 1994), or it may be that there is a grouping of under-stimulated normal individuals whose cognitive functioning could be significantly enhanced via learning enrichment programmes (see Schaie & Willis, 1986). There is also the proposition that Mitrushina et al. (1995) may have identified in their aging cohort, the *benign* pattern of neuropsychological impairment designated as ARCD in the DSM-IV, which is more general than just memory impairment, but which as in BSF and AAMI may tend *not* to evolve into a malignant dementia in the majority of cases, and thus can be clearly conceptualized as part of the normal pattern of cognitive variation in a healthy aging process. Finally it is possible that these researchers have identified the individuals who fall into the more impaired category isolated recently by Smith et al. (1996) of Mild Cognitive Impairment (MCI). However as discussed above, even in this MCI group only 55% have developed a dementia on clinical follow-up and it is not a group which is formally conceptualized as a 'disease' entity.

Thus in conclusion, from the perspective of a differential psychogerontology, once a normal non disease-related aging population has been properly delineated in terms of the criteria outlined above, it seems *empirically obligatory to retain a strong hypothesis amongst other postulations, that the variability may indeed be representative of the normal aging population within which it was identified*. This is particularly so in that on further analysis the reason for outlying patterns such as identified on the Mitrushina et al. (1995) study may be located, not in pathological patterns, but in aging processes which are amenable to reversibility.

Moreover, the finding of extreme heterogeneity in a normal aging population may account for overall low group means due to disproportionately low scorers; however this central tendency may obscure the fact that a significant number of individuals may have relatively preserved functioning, and these more positive patterns also need to be taken into account.

(iii) Terminal decline. Finally, in considering the constitution of a valid healthy normal aging population, it is necessary to address the issue of terminal decline. The concept of *primary* aging is delineated by Birren and Cunningham as slow, gradual changes which are believed to be a reflection of either a genetic aging program, or a cellular wear-and-tear deterioration program, or a combination of both. As discussed above, this process is distinguished from *secondary*, specifically disease-related aging. Both of these, however, are further distinguished from a third phenomenon which is that of *tertiary* aging (terminal decline). The notion of tertiary aging is derived from the empirical observation of cognitive decline in association with the terminal years which has been identified as an important source of variability in cross-sectional aging studies. It refers to a decremental phase on cognitive test performance that occurs a few years prior to death, and is hypothesized to be due in some way to pre-mortality physiological deterioration (Botwinick, 1977; Rabbitt, 1986). It is not possible to eliminate such subjects in advance from studies on cognitive aging and hence hypothetically they are seen to contribute significantly to the cognitive heterogeneity identified across older age groups (Botwinick, 1977; Rabbitt, 1986), although this has not invariably been confirmed on longitudinal follow-up (see Maddox & Douglass, 1974).

Importantly, from the perspective of this thesis, Botwinick (1977, p. 585) suggests that these undiagnosable adults constitute part of the reality of what contributes to intellectual variation in the normal older population, and to exclude them from aging studies would be "to make for highly abstract, artificial samples for study." However, this is an intricate issue, and it would appear that whilst it is important not to attribute a pattern of low scores to terminal decline *prematurely* on the basis of cross-sectional analysis (as per the arguments developed above), if post-hoc analysis or longitudinal follow-up reveals such an association, then it could be considered important to make the differentiation between terminal aging patterns from non-terminal cognitive aging patterns. This is because the former may serve artificially to lower the overall indications of cognitive potential in normal older adults (see discussion in Rabbitt, 1986).

In summary, this review on the complexities of what constitutes a generally representative population for research into normal cognitive aging, particularly in terms of the heterogeneity of the normal aging process, reveals an intricate set of conceptual issues which must be given due attention in the design and interpretation of research studies. For empirical purposes it is argued that such a generally representative population should comprise healthy, non-diseased persons, who do not have a diagnosis of Alzheimer's Dementia although they may have age-associated, 'normal' cognitive problems such that they would qualify for the categories of AAMI, BSF or MCI. Further it is posed that patterns of poor performance in normal healthy older populations, should not prematurely be considered representative of phenomena such as pre-clinical dementia, or terminal decline, but rather as possibly illustrative of the 'normal' heterogeneity in the aging process until post-hoc or longitudinal analysis provides contrary indications.

Representative environments. Birren & Cunningham (1985), have noted that the incorporation of the term 'representative' with respect to the environment in the normal aging definition of Birren & Renner (1977), is to imply that the notion of normal aging would exclude age changes that occur in mature individuals under unusually appearing environmental conditions such as wartime, epidemic illness, or natural disaster type situations. However, beyond such extreme situations, environmental conditions for the elderly are particularly complicated relative to other age groups in terms of what is to be considered 'generally representative'. Aging amongst older individuals can be seen to occur normally under at least the three distinctly different conditions of (i) residence in the open community, (ii) residence in retirement community villages, or (iii) institutionalization in a home for the elderly. Thus a further factor which contributes strongly to the heterogenous nature of the aged population, is the inherent variability in normal older age environmental situations.

Several researchers have emphasized the differential effect of the aged institutionalized and non-institutionalized environmental conditions on standard cognitive test performance and recall of early memories (Holland & Rabbitt, 1991a; Morris & Kopelman, 1986). Memory in particular, is considered to reflect the needs of older individuals in terms of their particular surroundings (Hindmarch, 1993). In spite of the differential effects of these contrasting living situations in older populations, it is clearly not possible to exclude any of the above mentioned three environmental conditions, including that of institutionalization in a senior citizens' home, from what typifies the normal aging situation. Consequently clear delineation of environmental conditions of samples within aging studies is particularly crucial in order to account for such situational effects.

Birren & Cunningham (1985) note that there is a tendency for empirical endeavours in healthy normal aging to be focused on *community-dwelling* older adults (which includes individuals from retirement villages in which people live independently), rather than on older adults in senior citizen homes, in that the former would be the least likely to be confounded by unrecognized disease-related aging factors which may have contributed to diminished functioning and a decision for admission to a home. Furthermore when comparisons of cognitive processes with younger adult age groups are at issue, it is essential to keep the environmental conditions as similar as possible. Some older adult normative data studies have used hospitalized cardiac or hospitalized convalescent subjects (Goul & Brown, 1970; Stanton, Jenkins, Savageau & Zyzanski, 1984), and Hindmarch (1993) actually lists *in-patient and out-patient elderly populations* amongst typical sample resources for aging studies in addition to *volunteer elderly, senior citizens clubs, retirement communities*, and *social service homes*. However, it appears problematic to consider hospital populations, even if they are convalescent populations, as suitably representative of a primary, healthy normal aging population. It is highly likely that research results would be confounded with disease- and/or medication-related effects, which in turn would result in significantly increased variability in outcome that would not be strictly attributable to non disease-related aging processes.

In sum, from this discussion, a generally representative environment for the purposes of studying normal cognitive aging, is considered to be the largely community-based environment which is that situation which would be least likely to confound research with clinically unidentifiable, subtle disease-related effects.

<u>Generally representative individuals and environments: Implications for the present study</u>. From the above discussion it is clear that the extreme heterogeneity of the aging population is intricately linked to increased variability in both *normally representative older individuals* as well as in the *normally representative environmental conditions* in which older individuals reside. Thus clear delineation of both in operational terms is crucial for normal aging studies. It is acknowledged that in broad epidemiological terms a normal older population

contains a large proportion of individuals with Alzheimer's dementia and individuals with physical disease. However, the consequent extreme heterogeneity in the epidemiologically-defined older population makes the documentation of *general* aging cognitive processes of dubious empirical value. Thus it is considered legitimate to investigate the normal aging process in a non disease-related (primary) older population, which is separated out from a disease-related (secondary) population. Accordingly, it is appropriate to restrict the environmental situation of the research sample to community based elderly individuals as far as possible, a focus which is less likely to result in the inclusion of unidentified cases of physical disease or dementia and confound the picture of a primary aging process with that of a secondary aging process.

Thus it was decided that the data to be analyzed in the present study would be based on non disease-related, largely community based samples of individuals. On empirical grounds, it was decided that distinctive patterns of performance identified on the basis of such pre-defined healthy normal older populations, *both positive and negative*, would be considered to be *intrinsic aspects of the delineated non disease-related populations under analysis, and not parcelled out as erroneous*. Thus the identification of heterogenous patterns would serve as part and parcel of the characterization of normal cognitive aging processes.

2.2.2 The concept of normal aging: Section summary

In concluding this section on the concept of normal aging, the central points that have emerged are as follows:

(i) Standard definitions of normal aging have been expressed in language which adheres to the normativity tradition in aging research serving to perpetuate the paradigmatic neglect of the importance of variability. Thus a definition has been proposed which is couched in discourse with a more differential focus, and specifically the definition reflects variability as *regular* occurrence in association with aging;

(ii) The normal aging process is intrinsically bound up with the concept of changes across time, and hence normal aging research has tended to use chronological age as the central scale of measurement - a scale which in the context of a normative aging research paradigm has been heavily criticized because it obscures the variability of individuals on biological, psychological and social dimensions. However, from the perspective of a differential gerontology it is proposed that this problem disappears, and chronological age serves as the ideal pragmatically accessible and robust independent variable against which to study precisely those biopsychosocial patterns of variability which it throws into relief;

(iii) Research descriptions of the normal aging process occur via the two broad methodological strategies of cross-sectional or longitudinal research in which the former tends to exaggerate age differences due to generational effects, and the latter tends to obscure age changes due to selective attrition of the non-elite, practice effects, statistical regression to the mean, and environmental enhancement effects. It is proposed that these classical confounding factors in longitudinal research will also serve to distort indications of variability inherent in normal age change. Thus, whilst both cross-sectional and longitudinal studies are necessary to understand the nature of changes in variability with aging, cross-sectional studies have particular heuristic value: if the samples are balanced numerically and well-matched on key influential variables such as educational and IQ level, it seems that they may be less likely to obscure subtle patterns of inter-individual variability; further they provide the necessary *current* perspective on age-related cognitive patterns upon which policy and clinical decisions should be based;

(iv) In order to document differences across age groups in normal aging research, the optimal target population is considered to be that of healthy non disease-related individuals, living in non-disaster community-based situations, although it is acknowledged that this population is not generally representative of the elderly population at large; the latter is so extreme in its variability as to make general statements about cognitive patterns in relation to it of dubious value;

(v) Typically the conceptualization of a discontinuity between diseased versus normal aging processes is adhered to in the operationalization and interpretation of normal aging research studies. However, from the focus of a differential psychogerontology the issue around what degree of variability is continuous or discontinuous with normal aging is highlighted and acts as a reminder not to discard variability prematurely as discontinuity (that is as erroneous) when describing the normal aging distribution.

2.3 CHAPTER SUMMARY

In this chapter the notions of '*variability*' and '*normal aging*' have been delineated and the methodological implications for a study of the inter-individual variability in cognitive performance across the adult years have been considered. From the *differential* perspective of this chapter, it has been emphasized that *inter-individual variability constitutes the presence of important behavioural facts that should be systematically investigated and not dismissed as error, nor should extreme variability be viewed as necessarily discontinuous with the genuine distribution of the normal aging population.* Failure to focus on the variability particularly in studies of heterogenous populations such as older adults due to operation within the conventional 'blinkers on' central tendencies paradigm, may result in artifactual statistical conclusions and distorted impressions about the aging process.

Methodologically it has been noted that variability *across age groups* can be examined by the statistical comparison of standard deviations across groups, and variability *within age groups* can be described via the plotting of individual distributions and taking account of the numbers of individuals that contribute to distinctive heterogenous patterns around the central tendency of the distributions. It has been further argued that for a study of inter-individual variability in cognitive test performance, the index of *chronological age* serves as an easily accessible robust screen against which to highlight variability data for scrutiny; *cross-sectional data* are less likely to obscure subtle patterns of inter-individual variability than longitudinal data, and furthermore are essential to depict the current status of cognitive patterns for clinical and socio-political purposes; and whilst less heterogenous than the generally representative older adult populations in an epidemiological sense, *non-disease related, largely community-based older populations* are more comparable to younger adult populations, and thus more suitable for making valid comparisons of healthy normal aging processes across age groups.

These arguments support the methodological framework of the present study which set out to examine the standard deviations and the individual distributions of neuropsychological test data across and within-groups defined on the basis of chronological age; the data selected for analysis in the study were derived on the basis of cross-sectional research with healthy, largely community dwelling, adult populations.

3

CHAPTER 3: THEORIES AND MODELS IN NORMAL AGING

The aim of Chapter 3 is to locate the issue of inter-individual variability within the current state of theories, concepts and research paradigms of normal cognitive aging, and in this way to develop a theoretical framework for the present study. The chapter begins with a general overview of theorizing and research orientations in the psychology of aging, and against this background the theoretical status of the present research is clarified. Next, the specific approaches which formed the basis of the present study are presented as follows: (i) the *neural* perspective, and (ii) the *psychometric/neuropsychological* perspective. The manner in which these frameworks served as the springboard for the present investigation are elucidated.

3.1 OVERVIEW OF SCIENCE IN COGNITIVE AGING

In this section a metatheoretical analysis on the state of theorizing and research in the psychological science of aging is presented in terms of its implications for the present study, and is divided up as follows: (i) the metatheoretical status of cognitive aging science, (ii) the two broad research paradigms of 'normativity' and 'differentiality', and (iii) an analysis of research orientations as they lie along a descriptive-exploratory/causal-analytic/hypothetico-deductive continuum.

3.1.1 The state of theory in cognitive aging science

In the literature, attention has repeatedly been drawn to the fact that there is no general unifying theory for the psychology of aging, and the discipline consists rather of a collection of concepts which explain limited aspects of aging behaviour, and a number of theoretical stances from which contributions are made towards the description and understanding of the phenomenon (Birren & Renner, 1977; Birren & Cunningham, 1985; Salthouse, 1991). For example, ideas and research about cognitive aging are derived from the six perspectives of psychometric, Piagetian, neo-Piagetian, information-processing, learning, and contextual perspectives which have been delineated by Berg (1992); or the broadly comparable psychometric, cognitive-analytical, reduced processing resources, competence and environmental approaches isolated by Salthouse (1991). In addition further insights arise from neural approaches (Bondareff, 1985), neuropsychological approaches (for example, Lezak, 1995), and brain reserve capacity theory (Satz, 1993), and a variety of single and multifactor information processing models from within cognitive psychology (for example Rabbitt, 1993c; Salthouse, 1985b). Aging theorists such as Birren and Cunningham (1985) and Shock (1977), however, point out that in such a complex subject as the psychology of human aging, which involves diverse phenomena ranging from molecular and cellular events through to physiological, psychological and social events, it seems appropriate that it is approached, in this way, from many points of view.

The multiconceptual framework as evidenced in psychogerontology is in anycase entirely commensurate with the multiparadigmatic status of psychological science as a whole (Kuhn, 1970), and there is no reason or likelihood that it should become an exception. Nor is there a need to be apologetic about this situation in broad terms. From the relativistic philosophical standpoint of the *'principle of complementarity'* it has been argued that a multitheoretical approach is in fact the only pragmatic solution to scientific exploration in the case of such a complicated phenomenon as human behaviour (Shuttleworth-Jordan, 1982; 1984). The principle of complementarity has its origins within the context of the modern physical sciences, (in which it serves to explain the dual, apparently opposing theories of light), and from the perspective of this philosophical framework the following tenets apply:

(i) multiple theoretical viewpoints are conceptualized as both *necessary* and *complementary* in the description of a complicated event under investigation;

(ii) these multiple viewpoints should not be seen as mutually exclusive and competing but rather as *supplementary* to each other;

(iii) description from a *single* theoretical viewpoint is not viable since such this is likely to produce a very much reduced and distorted view of a complex phenomenon.

An important *corollary* of the principle of complementarity is that since more than one perspective is necessary when studying a complex phenomenon, the *particular* theoretical position adopted in any scientific investigation cannot automatically be assumed as in a single paradigmatic science, and hence needs to be *especially* well defined. Shuttleworth-Jordan (1982; 1984) has argued that it is useful to apply the above tenets from within the framework of 'complementarity' for use in psychological science as follows. First it is accepted that multiple frameworks are a *necessary* solution to the complexity of the phenomenon being studied (that is aspects of human behaviour). Second it is assumed within this context, that it is also necessary for a theoretician or researcher (who is conducting an investigation) to adopt a *clearly defined and conceptually coherent position* within one of the many possible psychological perspectives. In this way it will be possible to avoid the lack of conceptual integration inherent in an eclectic or a-theoretical data-gathering approach.

Accordingly, with specificity to this thesis, the principle of complementarity is a metatheoretical framework which can be seen to legitimize the current multi-perspective status of the science of cognitive aging; at the same time however, it does not condone an a-theoretical, shot-gun research stance, which would only serve to add to the disparate plethora of data and confusion which easily occurs in a multiparadigmatic science. Rather as an antidote to this potential pitfall, the principle of complementarity strongly emphasizes the need for a high level of theoretical integration from whichever vantage point is taken for the purposes of scientific investigation. *Thus an important objective for the present research was to adopt and develop a very clearly defined theoretical orientation from the plethora of possibilities within the science of cognitive aging, in order to contribute to a conceptually integrated description of the phenomenon.*

3.1.2 Broad research paradigms: Normative/Differential

The principle of complementarity can be effectively advanced as a metatheoretical framework to inform broad modes of scientific investigation in the psychology of aging, and in particular this has implications for the dual research paradigms of 'normativity' and 'differentiality'. Currently, as emphasized in the previous two chapters of this thesis, the overriding research paradigm in psychology *in general*, which occurs regardless of which theoretical approach is adopted *within* the discipline, and which is prevalent also in the study of psychological aging, is that of 'normativity'. The proposal of this thesis is that an *additional* balancing research paradigm should be promoted, which is that of 'differentiality'. Thus it is not suggested that a differential approach should replace a normative approach, but that *both* should be used in *complementary* fashion in order to provide a more accurate description of the phenomenon of aging. In order to illustrate the subject/object (observer/observed) interdependency of scientific observation, which is the basis of the principle of complementarity, Pirsig (1984) has used the analogy of how best to explicate the phenomenon of a 'motorcycle'. He highlights how a very skewed and erroneous description of what constitutes a motorcycle will emerge if it is only observed from a single vantage point, and that the true nature of the motorcycle will only be approximated by the summation of many descriptions from multiple observational points. And accordingly, it is this distorted view of the aging process derived only from the single, exclusive view of central tendencies research which has led to the ardent objections of those who are arguing for a differential psychogerontology (for example, Rabbitt, 1993b; Schaie, 1988a; Thomae et al., 1981).

The philosophical position of complementarity as proposed here, serves as a reminder that a *differential approach needs to be incorporated <u>in parallel</u> with a normative approach* in order to enrich the understanding of aging. As Horn and Hofer (1992) have also noted, these two broad-based research paradigms (the normative and the differential) are not mutually exclusive, but rather are complementary and interdependent. Hence, whilst this thesis has set out to provide a differential antidote to normative perspectives on cognitive aging, it is presented in the form of a *complement* and not a *replacement* within the discipline, in full acknowledgement of the fact that both viewpoints are necessary for a more accurate description of the aging process.

3.1.3 Levels of investigation: descriptive/analytic/theorizing

Whilst it can be argued from the perspective of 'complementarity', that intrinsically a multiconceptual approach is indicated for the psychology of aging, the available models that actually exist within the discipline do not reflect in the main an advanced stage of theoretical sophistication. The picture is one of a diversity of narrow conceptualizations about aging processes, rather than more fully articulated and integrated scientific theories from which deductions can be made and for hypothesis-testing research. Consequently, research tends to be largely problem and data oriented (Birren & Renner, 1977; Birren & Bengtson, 1988), and researchers from the domain of the cognitive information processing perspectives, such as Salthouse (Salthouse, 1985a; 1985b) and Rabbitt (Rabbitt, 1990; 1993a; Maylor & Rabbitt, 1994), who propose fairly elaborate theoretical models and engage in hypothesis-testing of these models, are notable exceptions to the general rule of non-theoretically articulated investigation that continues to prevail in the domain of aging research.

Commensurate with the a-theoretical tone of research investigation in cognitive aging in general, research which has specifically set out to investigate the phenomenon of *inter-individual variability* has tended to fall mostly into the category of the descriptive, data-gathering type; and even at this level the number of variability studies are very few, narrow in scope, scattered and yield unexplained contradictory findings (see discussion, section 1.2.2). Further, only a handful of studies (notably Rush, 1990; Schaie, 1983; Thomae, 1981) have taken any steps to analyze the reasons for the inter-individual variability described in their studies by means of empirical investigation, and even fewer attempts have been made to gather data at the level of hypothesis-testing research in order to inform theory (Rabbitt, 1993c; Maylor & Rabbitt, 1994). It has been emphasized, nevertheless, that at this undeveloped stage of theory in psychogerontological science, there is in fact a mandate for a predominance of *descriptive* and *analytical* research rather than deductive investigation (Birren & Cunningham, 1985; Montgomery & Borgatta, 1986; Nesselroade & Labouvie, 1985). A considerable amount of description and exploration of phenomena are the essential precursors of the formulation of theoretical frameworks, in that conceptual refinement occurs alongside the progression of such research, and inefficiency emerges if there is a premature move towards structured hypothesis-testing research (Birren & Renner, 1977; Birren & Cunningham, 1985). Thus whilst the pursuit of causal relationships grounded in theoretical exposition and demonstrated in hypothesis-testing research can be seen as the ultimately more valuable scientific goal, the initial importance of descriptive research is considered "critical" (Birren & Cunningham, 1985, p. 37).

However, the heuristic value of descriptive research lies centrally in the notion of it being a critical *initial* step in the research and theorizing process, and not as a means in itself. Poon (1985, p. 452), who has a particular interest in inter-individual variations in memory function, highlights the value of using descriptive data as the basis for systematic "post hoc" investigations for exploring such effects as health, gender and education on inter-individual variability in cognitive functioning. Hoyer (1974), who also has focused on the investigation of inter-individual variation in aging, states that it is not sufficient merely to describe the "wrinkles" - there is a need go further and discover the causes of the "wrinkles" via the systematic investigation of individualized behavioural characteristics. The exceptional theorizing work of Salthouse exemplifies how descriptive, data gathering research can form the crucial building blocks that in turn lead to the illumination and formulation of theory when considered as "a means, rather than as an end, in the attempt to understand fundamental aging processes" (Salthouse, 1985a, p. 401).

Drawing this analysis to a conclusion, therefore, along a *descriptive-analytic-theorizing* continuum, the state of the art of research into variability in cognitive aging, even more so than with research into cognitive aging in general, can be considered to be at a stage of extreme infancy in which the need for descriptive/exploratory research appears mandatory. Thus first and foremost the present research was designed with the purpose of taking up the mandate for extensive *descriptive and exploratory* investigation. Phase 1 set out to *describe* the group variability patterns across a broader range of cognitive tests, and across a wider range of adult age groups than had previously been accomplished. These results were considered in relation to previous research descriptions of such trends, an integrative endeavour that had also not been attempted before in any comprehensive depth. Phase 2 set out to *describe* the individual raw score patterns which gave rise to the group variability. This was a sequential, more detailed descriptive phase which had not been previously conducted in cognitive aging variability research.

As noted, however, whilst the gathering of descriptive data is considered vital as a much needed base-line in the developmental phases of scientific investigation, research ideally should not remain at that level, but serve rather as the means for further integrative analysis and theory building. Thus, it was decided to complement the *descriptive* phase of the present research, with a *causal-analysis* phase, by examining the individual scatter in terms of the influence of education, IQ and gender effects where such data were available. Further, on the *theoretical* dimension, whilst not at the stage of highly specific hypothesis-testing, causal-experimental research, the present research set out to go a significant way beyond a mere "shotgun" level, and was conceived within a *three-tiered theoretical hierarchy* of increasing specificity with respect to the development of hypotheses and the explication of the phenomenon of 'variability'. Most broadly the study was located within an umbrella *neuropsychological framework*, and the spectrum of tests chosen for examination were selected on the basis of representing a range of key functional modalities which inform brain-behaviour relationships as commonly delineated within this modality, and which provided the organizing structure for the demonstration of variability patterns in cognitive aging. Next the neurally-based and neuropsychologically-supported *brain reserve capacity theory* as formulated by Satz (1993), was used to explicate broad causal influences in inter-individual variability in aging. Finally, arising out of the rubric of both of these more broad ranging conceptual modalities, a specific *shuttle model of inter-individual variability* in aging was developed and was used to delineate the particular course and pattern of variability to be expected in association with cognitive aging.

3.1.4 Overview of science in cognitive aging: Section summary

In summary from this metatheoretical overview of theoretical and research paradigms in the psychology of aging, the following points have emerged:

(i) Cognitive aging science falls into the category of a multiparadigmatic science in that there are a plethora of approaches with no single unifying theory;

(ii) This multiperspective status is considered legitimate in terms of the relativistic principle of complementarity which advocates the need for many perspectives in the study of complex phenomena; however, in this situation it is necessary to take up a clearly defined theoretical stance within the multiple possible approaches, from which to conduct a research investigation;

(ii) In terms of the principle of complementarity, it was posed that the broad normative research paradigm should be complemented by the addition of a differential perspective in order to redress the existing limited and skewed view on the phenomenon of cognitive aging. This provided the overriding motivation for the *supplementary (*as distinct from replacement) contribution of the present research from a differential perspective;

(iii) Finally, on a continuum of data-gathering/analytic/and theorizing research, the current state of the art in cognitive variability investigation is in its extreme infancy. Thus the present study was designed with a view to producing a higher level of research sophistication in the area, by producing both (a) the critical initial requirement of a broad-based set of descriptive data, as well as (b) hypothetico-deductive theoretical possibilities for the integrative analysis of these data.

3.2 SPECIFIC PERSPECTIVES ON AGING

The aim of this section is to identify the specific aging perspectives which will be called upon for use in the present research, and to review them with reference to their implications for inter-individual variability in cognitive aging.

3.2.1 A choice of perspective

Psychogerontological texts which deal with conceptualizations of aging in general have tended to centre their discussions within the three broad groupings of *biological, psychological* and *social* perspectives on aging (Birren & Renner, 1977; Birren & Cunningham, 1985). With a more specific focus on *cognitive* functioning in aging, Berg (1992) has offered the six more theoretically differentiated frameworks of *psychometric, Piagetian, neo-Piagetian, information-processing, learning*, and *contextual*, which she proposes represent a broad view of the guiding approaches to research in the field of intellectual development; and Salthouse (1991) for his review of theoretical perspectives on cognitive aging has delineated the broadly comparable *psychometric, cognitive-analytical, reduced processing resources, competence* and *environmental* approaches. However, the categories isolated by Berg and Salthouse are entirely located within the cognitive-psychological and social dimensions of cognition, whilst omitting the biological dimension altogether.

A bridge between the psycho-social and biological dimensions is provided by the *neuropsychological* perspective (for example as described in Lezak, 1995; Walsh, 1991). This is an approach which, as noted in the introduction, is receiving recognition by leading psychogerontologists (Craik & Jennings, 1992; Rabbitt, 1993b) as an important future direction to take for unravelling the nature of aging processes, through an exploration of the behavioural correlates of brain pathology. A further neuropsychological integrative bridge between broad biological/and psychometric dimensions is the perspective of *'brain reserve capacity theory'* as delineated by Satz (1993). Although not formally catalogued in the major aging texts as an 'aging theory', this is a notion which has relevance for aging theory, and has particular significance with respect to the issue of inter-individual variability.

Thus for the purposes of this specifically *neuropsychological* thesis, which is concerned with the issue of variability in normal aging, the following two broad perspectives are offered for discussion: the *'neural'* perspective which is intrinsic to the neuropsychological perspective, and a broadly defined *'psychometric'* perspective, which includes the *'neuropsychological'* perspective. An overview of these viewpoints will be presented since they have been extensively discussed elsewhere, and the main focus will be to consider and expand on the central tenets of these frameworks derived from existing reviews, in terms of the issue of inter-individual variability. The Piagetian, information-processing, and learning perspectives are not included for discussion here as perspectives in their own right, in that they concern delineation of the mental mechanisms of cognitive processing and learning largely in isolation from the neural perspective, and are not central to the neuropsychological perspective of the present thesis. However, in that the single-factor versus multiple-factor information-processing debate of Salthouse versus Rabbitt (for example, Maylor & Rabbitt, 1994), has implications for the variability of cognitive aging processes as conceptualized within a neuropsychological framework, a discussion of the debate will be incorporated within the psychometric/neuropsychological subsection which follows below. In broad terms it is considered that the 'contextual' or 'environmental' approaches are contained within the neuropsychological perspective of this thesis, in that socio-cultural and environmental factors are central to psychometric test construction and the interpretation of neuropsychological test patterns.

3.2.2 The neural perspective

Neural aspects of aging fall under the broader rubric of biological and physiological aspects of aging which include genetic theories and organ system mechanisms of aging such as cardiovascular or metabolic systems, in addition to the mechanisms of the nervous system. (For a detailed review of these broader biological and physiological theories and assumptions see Riegel, 1977; Shock, 1977). In the broadest terms the biological view of aging conceives of the life span of the animal species as determined genetically, and thus the aging program that resides in the genes is conceptualized to underlie all biological aspects of aging, including age changes within the nervous system. Over and above the genetic factor, biological aging is considered to consist of the complex integration of multiple cellular and organic agents which result in changes in the human biological system and ultimately in death.

Specifically with regard to neural agents within the broader biological system, the nature of the changes that are noted to occur in the brain during the course of normal aging have been extensively described (for example, Berlot & Volk, 1980; Bondareff, 1985; Hinkin et al., 1990; Ivy, MacLeod, Petit, & Markus, 1992; Seitelberger, 1980). In summary from these reviews it appears that the well-documented brain changes that occur consist of a number of distinct processes as follows: (a) neuronal atrophy in the form of cell disintegration and neuronal death, (b) the accumulation of lipofuscin in the nerve cells (lipopigment dystrophy), (c) increased density of glial cells contributing with neuronal loss to an increased incidence of glia to neuron ratio with age with possible consequences on the maintenance of brain homeostasis, (d) neuroaxonal dystrophy, (e) the accumulation of extracellular amyloid fibril formation (senile plaques) and neurofibrillary tangles in the soma and dendrites of the nerve cells, contributing to manifestations of cellular degeneration, (f) marked morphological changes in microvasculature from being relatively straight and evenly dispersed to more coiled, tortuous, and differentially dispersed, resulting in decreased cerebral blood flow. Accompanying, and apparently consequent on these processes, age-related reductions in brain mass are regularly reported (Berlot & Volk, 1980; Peress, Kane, & Aronson, 1973; Tomlinson, Blessed, & Roth, 1968). It is observed, further that all these neural changes which accompany normal aging, are characteristic also of Alzheimer's dementia, except that in Alzheimer's disease the neural attrition is much enhanced. As yet no clear neuropathological marker of Alzheimer's disease exists to differentiate it from normal aging apart from severity of the degenerative brain changes, and thus a grey area exists in the differentiation between what constitutes pre-clinical Alzheimer's dementia and normal age changes (De Kisky & Bass, 1980; Esiri, 1994). (See prior discussion, section 2.2.1.3).

Importantly, as Seitelberger (1980) has emphasized, within the above-listed collection of relatively independent partial neurophysiological aging changes that occur, there is *extensive variability* in the individual forms that they take. Paradoxical effects have also been documented in that whilst dendrites generally decrease in size during aging, in some instances what may be compensatory age-related dendritic proliferations are reported (Ivy et al., 1992). In addition environmental enrichment programmes have produced increased neuronal density although with increasing difficulty with advancing age (see discussions in Diamond & Connor, 1982; Huppert, 1990; Ivy et al, 1992; Labouvie & Vief, 1985). Further variations are apparent in that not all nerve cell populations are affected by atrophy due to nerve cell disintegration and death, and the accumulation of plaques and neurofibrillary tangles and vascular changes localize to certain sites (for example, Ivy et al. 1992; Seitelberger, 1980). Direct brain investigations via autopsy or neuroradiological mechanisms, consistently reveal that the majority of tissue loss and vascular changes appear in the cerebral cortex, and the atrophy is most marked over the frontal lobes. However, significant losses have also been documented in the parietal and temporal lobes, and moderate amounts of tissue loss demonstrated in hippocampal regions; and there are variable reports of cell loss in subcortical brain regions (see reviews in Hinkin et al., 1990; Ivy et al. 1992).

Further multiplying all these divergent possibilities around the nature of neuronal changes with normal aging itself, are variations in the time course of neuronal loss between individuals, as well as the effect of all of these on how they differ in their expression in various aspects of behaviour, including cognitive behaviour (Bondareff, 1985). Thus with respect to variability, the decremental neural changes in normal aging can be conceptualized as a form of slowly progressive brain damage which is likely to result in highly variable inter-individual patterns of behavioural expression in older age groups. Further, since not all parts of the brain are equally affected, there are likely to be differential effects on cognitive task performance, depending on the particular stage and site of the aging attrition. However, the description of such differential effects on cognition to be expected in association with the normal, progressive, relatively independent and partial age-related processes of neural aging that have been described in this subsection, falls outside the field of *neural* investigation. The source of research insights in this regard lie in the psychometric dimension.

3.2.3 The psychometric/neuropsychological perspective

The view of aging from the psychometric perspective is heavily dependent on the results of performance on psychometric tests on the basis of which differences in cognition between individuals and age stages are delineated (Berg, 1992). For the purposes of this thesis the psychometric framework is used as the organizing concept for an overview of the patterns of cognitive functioning that have been established by means of psychometric tests across adult age groups, regardless of whether the data have been derived from the cognitive or neuropsychological research perspectives. The neuropsychological perspective, which has as its focus the documentation of the behavioral expression of brain dysfunction (Lezak, 1995), has evolved out of a combination of the neurological and psychometric traditions (Russell, 1986), and thus can appropriately be used in this section as the integrating higher-order concept by which to connect the psychometric aspects to be presented, with the neural aging features discussed in the previous subsection.

The review of psychometric findings to be presented will begin with a discussion of *general intellectual functioning*, followed by the specific functional domains of *attention, memory, language skills, visual skills, problem solving, and motor activity* based on the divisions adopted by Lezak (1983; 1995) and Spreen and Strauss (1991) in their seminal texts on neuropsychological tests and assessment, and further narrowed down to include those functions which have received most attention with regard to age-related changes as are incorporated in the overviews of Lovelace (1990) and Nolan and Blass (1992). The presentation will take the form of a relatively brief synthesis of the general normative findings of aging patterns within each of these functional domains, since these have been extensively reported and discussed in the major aging and normative data texts (for example, Birren & Schaie, 1977; 1985; 1990; Boller & Grafman, 1990; 1991; Craik & Salthouse, 1992; Heaton et al., 1986; Huppert et al., 1994; Lovelace, 1990; Lezak, 1983; 1995; Spreen & Strauss, 1991). The section will be concluded with a synopsis of the theoretical debates surrounding the meaning of the typical aging decrements which have been identified. The objective of this overview is, in the mode of *complementarity,* to provide a synopsis of the common findings from the prevailing aging literature which up to now has revealed an almost exclusive focus on central tendencies in such functional reviews, as the necessary background against which a corresponding synthesis of the available indications on variability within functional domains will be reviewed in the following chapter (Chapter 4).

3.2.3.1 General intellectual functioning: Average trends

Across extensive reviews of the age-related changes in intellectual functioning with advancing age, as noted on the basis of tests covering a broad range of cognitive functions, and including age-related effects on IQ tests, there is general agreement about a number of reliably established average trends. It is well-established, for example, that decline in intelligence is clearly part of the aging picture and that it is a differential decline in terms of abilities (Albert, Duffy, & Naeser, 1987; Birren & Cunningham, 1985; Charness & Bosman, 1992; Cunningham & Tomer, 1990; Botwinick, 1977; Heaton et al., 1986; Hinkin et al., 1990; Labouvie-Vief, 1985; Lovelace, 1990; Nolan & Blass, 1992; Schaie, 1988b; 1990; 1994; Zarit, Eiler & Hassinger, 1985). The exact stage of onset of decremental changes is still a matter of some debate. Albert et al. (1987) summarize the picture of generalized changes in intelligence as follows: It appears that there is relatively little decline in intelligence test scores until around the age of 50, and after this age the indications differ depending on whether cross-sectional or longitudinal methods are used. Cross-sectional studies reveal sharper and earlier declines of over one standard deviation from around 60, dropping sharply after 70; the longitudinal method shows later declines beginning only in the late 60's. Both show substantial declines from the mid 70's.

Whether cross-sectional data overestimates and longitudinal data underestimates the average age changes remains unresolved (Schaie, 1988b), since both methodologies are subject to significant confounding effects (see earlier discussion, section 2.2.1.2). Some researchers tend to emphasize the evidence from longitudinal studies which suggests later onset of intellectual decline (for example Labouvie-Vief, 1985; Schaie, 1983; 1994), whereas others tend to highlight signs of earlier decline (for example Holland & Rabbitt, 1991b). Thus, whilst the above overview from Albert et al. (1987) indicates relatively late onset of intellectual decline in both cross-sectional and longitudinal studies (that is from around 60 onwards), Holland and Rabbitt (1991b) report that some careful studies have detected age-related changes in tests of intelligence and creativity, and the ability to master a fast, complex interactive video game, as early as in the late third and early fourth decades.[7]

There is overall consensus on the fact that age changes are clearly domain-specific with some abilities showing substantial decline in the middle adult years, others only in the 60's and 70's, and some remaining stable beyond those years (for example, Schaie, 1994). Hence the generalized statements with respect to intellectual decline which have been described above, are of dubious value. The unstable fluid versus stable crystallized view on aging (Horn & Cattell, 1967), encapsulates the consistent very broad observation that aging decrementally effects novel problem solving ability and new learning (fluid abilities), whilst minimally effecting the ability to make use of old-acquired and frequently-employed information (crystallized abilities). This is a differential feature that is marked on cross-sectional

---

[7] Aging researchers such as Labouvie-Vief (1985) have raised the point that age-related declines which are so robustly demonstrated on cognitive test performance in the elderly may not always represent *decrement*, in that the change may be in a direction which has positive indications for adaptive functioning in later age stages. Further it is suggested that tests designed for younger populations may lack validity for uncovering the aspects of functioning which are pivotal in the societal contributions of older populations. Since the focus of discussion in this thesis is on the pattern of test differences between age groups on psychometric test data *in themselves*, subtle ramifications of the meaning of differences, such as raised by Labouvie-Vief, are not incorporated into the main body of the present discussion. However, it is important to be aware of such possible contributory influences on psychometric test patterns aside from genuine decremental aging effects.

studies, and is revealed also on longitudinal studies in spite of the minimization of age changes due to selective attrition (Schaie, 1994; Zarit et al., 1985). Nolan and Blass (1992) describe how this effect is reflected on cross-sectional data with the WAIS: unadjusted Performance IQ (which contains a greater predominance of fluid functions), peaks in the 20's and then declines sharply, with the mean falling below adolescent level in the 40's and reaching the borderline retarded range in the 70's; unadjusted Verbal IQ (which has a greater loading of crystallized tasks), peaks between the ages of 25 and 34, but subsequent decline does not reduce the mean score below the adolescent level until the early 70's.

With respect to specific tests, this trend of fluid (non-hold) versus crystallized (hold) is clearly revealed in a summary of the percentage of test variance accounted for by education versus age reported in Heaton et al., (1986). On the basis of Heaton et al.'s schema, tests from the WAIS and Halstead-Reitan battery *most to least effected by age* (with the inverse for education), are ranked as follows: *Halstead-Reitan (H-R) Grooved pegboard; H-R Trails B; Picture Arrangement; Digit Symbol, Object Assembly; H-R Finger Tapping; Wechsler Block Design; Picture Completion; Digit Span; Arithmetic; Similarities; Comprehension; Information; Vocabulary.* Falling at the end of the spectrum *least* affected by age and most dependent on acquired knowledge, in the Heaton et al. research, are Vocabulary and Information, followed by Comprehension, and Similarities. At the end of the spectrum *most* affected by age are the novel tasks of visuo-spatial sequencing (Trails B and Digit Symbol), visuo-spatial problem solving (Object Assembly and Block Design). Of note is the fact that all of the latter tasks involve perceptual speed, whereas none of the former involve timed scoring procedures. Digit Span (which taps sustained attention and concentration), Picture Completion and Arithmetic (both of which involve reasoning ability), lie around the middle. Clearly falling into the age-affected sphere are the speeded hand motor tasks (grooved pegboard and finger tapping). The grooved pegboard, which involves *perceptual* speed to a greater extent than does finger tapping, is the more strongly associated with aging.

Consistent with the data that emerge from Heaton's continuum, is the differential pattern of aging deficit across functional modalities which was identified by Schaie (1983), and has been encapsulated in the 'cascade model' described in Birren and Cunningham (1985). In this model it is proposed that in descending order, the three cognitive constructs of 'verbal comprehension', 'reasoning', and 'perceptual speed' are *increasingly* differentially sensitive to aging.

3.2.3.2 Specific functional modalities: Average trends

Specific reporting and discussion of the expected patterns of deficit within functional domains have been collated across chapters distinguished by functional area in the seminal cognitive aging literature (Boller & Grafman, 1990; 1991; Birren & Schaie, 1985; 1990; Craik & Salthouse, 1992), and in summary form by Nolan and Blass (1992), and Lovelace (1990). Most commonly the areas focused on in these sources can be distinguished as *attention* (Hartley, 1992; McDowd & Birren, 1990; Lovelace, 1990; Nolan & Blass, 1992), *memory* (Craik & Jennings, 1992; Hultsch & Dixon, 1990; Huppert, 1990; Light, 1990; 1992; Lovelace, 1990; Nolan & Blass, 1992; Poon, 1985; Sagar, 1990), *language skills* (Huff, 1990; Kemper, 1992; Kempler, 1994; Light, 1990; 1992; Lovelace, 1990; Nolan & Blass, 1992; Older & Albert, 1985), *visual skills* (Filoteo, 1994; Lovelace, 1990; Ogden, 1990; Nolan & Blass, 1992), *problem-solving* (Cronin-Golomb, 1990; Lovelace, 1990; Nolan & Blass, 1992; Reese & Rodeheaver, 1985), and *motor activity* (Lovelace, 1990; Welford, 1977). A further major source of age-related patterns on cognitive test performance is available from normative data studies within the neuropsychological literature. Age-related normative data for central trends on single neuropsychological tests which fall into all of the above functional domains have been collated and are reported in neuropsychological assessment compendiums (Lezak, 1983; 1995[8]; Spreen and Strauss, 1991). In addition there are a substantial number of normative studies which report mean score data on older populations across a spectrum of neuropsychological tests falling within these functional domains, which do not appear in these assessment compendiums (for example, Albert, et al., 1987; Cullum, Butters, Troster & Salmon, 1990; Heron and Chown, 1967; Mitrushina, Satz, Van Gorp and associates; and Shuttleworth-Jordan and associates). For the purposes of this discussion, age-related central trends will be reviewed from the above-mentioned two major sources: the cognitive aging literature, and the neuropsychological normative data literature.

Although there is evidence to suggest that for different aspects of cognition, changes with age occur at variable rates across the adult life span (see discussions on fluid versus crystallized

---

[8] The second edition of Lezak (1995) differs from the first edition of Lezak (1983) in that the 1983 edition served more extensively as a compendium of normative data, whereas such data are largely only referred to rather than tabulated in her most recent text. The reason for this is that a companion text (D'Elia, Boone & Mitrushina, in press) will contain the tabulated normative data. Since the latter text was not available at the time of completion of this thesis, the present research study and literature review was based on the spectrum of norms containing older adult mean score data (and standard deviations), as collated in Lezak (1983) and Spreen and Strauss (1991), and supplemented by older adult studies available from the literature.

age changes above), the summary reviews in the above-cited cognitive aging literature on specific functions have not systematically attempted to track the exact age gradient of change across functions. Rather the overall views are presented in relatively undifferentiated terms in terms of onset of age change, and hence will be similarly reported here. Further, a comprehensive synthesis of age effects across neuropsychological normative test data was not the purpose of the clinically-oriented compendiums of tests and norms of Lezak (1983; 1995), and Spreen and Strauss (1991), nor has the formal identification of age-effects been the focus of many norming studies available in the neuropsychological assessment literature. Thus frequently the statistical significance of mean score age differences across age groups has not been established, nor have such results (if available) been systematically reported in the collations of normative data. Thus age-related mean score *trends* from a cross-section of tests which cumulatively support deficit patterns within functional domains are incorporated into the present discussion, in conjunction with statistical indications.

The neuropsychological studies focused on (from the profusion that exist in the literature) are those for which older adult normative *mean scores* are available, in addition to the variability data in the form of *standard deviations,* both of which are necessary for the statistical analyses to follow in the later development of this thesis. Thus data which are documented in the form of percentile ranges in Lezak (1983; 1995) and Spreen and Strauss (1991), and the large older adult normative data bases that exist from Mayo's Older Americans Normative Studies (Ivnik, Malec, Smith, Tangalos, Petersen, Kokmen & Kurland, 1992; Ivnik, Malec, Smith, Tangalos, & Petersen, 1996), from the Canadian Study of Health and Aging (Tuokko & Woodward, 1996), and on the extended Halstead-Reitan battery (Heaton et al., 1986), which are reported in the form of age- and education-adjusted scaled scores and percentile ranges (Ivnik et al., 1992; 1996), or T-score equivalents of scaled scores (Heaton et al., 1986; Tuokko & Woodward, 1996), are specifically not included. However, whilst this discussion exclusively focuses on *mean score trends* across age groups on specific neuropsychological tests within functional domains, none of the overall aging indications from the large percentile, and scaled score normative data bases stand in contradiction to the broad trends evident from the studies which are incorporated into the present review.

For each of the functional domains delineated above the discussion proceeds as follows: (i) operational definitions derived largely from Lezak (1995), are followed by (ii) a brief summary of broad age-related central tendency findings based on a synthesis of the above-cited cognitive aging literature, which in turn is followed by (iii) the mean score indications from normative data studies on the selected neuropsychological tests.

Attention. Attentional deficits appear as distractibility or the inability for focused behaviour despite an individual's intention, and intact attention is necessary for concentration and tracking activities, as well as for speed of reactions. There is no pure test of attention, and it can only be assessed as a component of other cognitive tasks. Attention is commonly delineated to consist of 'sustained attention' and 'divided attention', the former showing minimal decline with aging, whereas there is a robust finding of age-related decrements in the latter. A third dimension of 'selective attention' (the ability to discriminate between relevant and irrelevant stimuli), has unclear findings with respect to age, but does appear to be maintained when perceptual factors are controlled. Nolan and Blass (1992) cite the longitudinal study of Crossley, Hiscock and Bessie (1991) which produces compelling evidence of a decline in global attention capacity with age.

Commensurate with the above, are the findings on cross-sectional normative data studies for specific neuropsychological tests. For the Paced Auditory Serial-Addition Test (PASAT) of Gronwall and Wrightson (1974), a test of *divided attention*, a regular trend of age decrements is noted (Roman, Edwall, Buchanan & Patton, 1991; Spreen and Strauss, 1991; Stuss, Stethem, & Poirier, 1987). On the other hand an absence of *marked* age decline is noted on the auditory and visual continuous attention tasks of Mirsky (Albert et al., 1987), as well on the digit span forwards task (Heron & Chown, 1967; Mitrushina & Satz, 1989; Shuttleworth-Jordan, 1992), all of which test *sustained attention*. It is important to be aware that performance on the relatively non age-sensitive digit span forwards test, which is a test of immediate attention, may be obscured when reported as part of the total WAIS Digit Span test which incorporates digit span backwards, a test of mental tracking which is more sensitive to aging than digits forward (see discussion in Lezak, 1995). Thus ideally this test should not be scored in composite form. However, it appears that the composite Digit Span measure also does not reveal a marked age effect (Craik and Jennings, 1992; Cullum et al., 1990; Heaton et al., 1986; Poon, 1985).

Memory. Memory functioning is not a unitary phenomenon and is differentiated along multiple dimensions which are assessed by means of separate tasks. Memory tasks can be intentional (that is there is the intention to memorize at the test time), or incidental (when recall is required of the testee without prior warning). Memory performance is also differentially represented in terms of the verbal versus visual nature of the task. Further, memory is variably expressed across a range of forms of memory task which tap the functions of immediate retention, short-term retention, recent and delayed retrieval (in terms of minutes), recognition, new learning ability, and remote memory (in terms of months and years). Finally there are tasks which tap working memory, a construct used to describe the dual process of retention of verbal or visual material in short-term memory storage, and the simultaneous cognitive manipulation of that material (Baddeley, 1987).

Deteriorated performance due to age is clearly demonstrated across age-related normative data on both intentional and incidental visual recall tasks. (There is sufficient similarity between these two aspects of recall with respect to brain damage effects, that in clinical neuropsychological assessment not a great deal of emphasis is laid on differences between these two modalities - for discussion see Shuttleworth-Jordan & Bode, 1995a). Further it appears that both the verbal and visual modalities of memory show aging effects, with reviews providing no compelling evidence for greater aging effects on visuospatial over verbal tasks (Lovelace, 1990). However, further to these broad dimensions, that different components of memory functioning show vastly differential effects due to aging has been robustly demonstrated and is explicated very clearly in graphical form in Poon (1985, p.431): Whereas Digit Span remains almost stable with age, an opposite substantial decline is seen with the retention of paragraphs, and an intermediary fall-off is noted on the retention of paired associates and memory for designs. Thus it can be seen that, as noted under the section on 'Attention', relatively small age changes are noted in immediate memory (as tested with digit span). In contrast, the memory aging literature cited consistently reflects marked age-related changes in recent and long-term retrieval and new learning ability.

There are subtle, and at times seemingly paradoxical nuances which surround the underlying mechanics of memory in association with normal aging (see Perlmutter, 1980). Perlmutter (1980) describes a number of research instances (aside from those noted with respect to immediate attention), in which age does not appear to show an effect on memory tasks. She suggests that the apparent paradoxical effects can be explained by the fact that older individuals have a much richer data base than younger individuals, and that when tasks are relatively more dependent on acquired data then memory processing, there will be an absence of age-related effects. Commensurate with this, the cognitive aging literature reflects the observation that age-related memory deficits appear to be due to deficiencies in encoding or retrieval rather than diminished storage capacity, and this is supported by the fact that consistently recognition tests of memory do not show significant age effects.

With regard to the modality of remote memory, in spite of the fact that older people often indicate that their youthful memories are much more distinct than memories of current events, the empirical evidence suggests otherwise. Whilst there is some sparing of salient, personalized aspects of autobiographical memory, generally there is also a notable decline in this aspect of memory. Commensurate with this is evidence that suggests that in older populations memory for recent, public events may be better than memory for past, public events (Cohen, Conway & Maylor, 1994; Maylor, 1991; see also reviews in Craik & Jennings, 1992; Hindmarch, 1993; Sagar, 1990). Finally, for the related but different construct of working memory (that is working with memory), aging researchers emphasize that this modality has consistently been revealed to be exceptionally sensitive to aging effects (Craik & Jennings, 1992; Hultsch & Dixon, 1990; Light, 1990; Rabbitt, 1986).

There is ample normative data reported on neuropsychological tests in which the mean score trends are commensurate with the above summary of memory aging effects from the cognitive aging literature. For example, marked age decrements are noted in short-term verbal and visual immediate and delayed recall for the Logical Memory, Verbal Paired Associates and Visual Reproduction subtests of the Wechsler Memory Scale - original version (Mitrushina & Satz, 1989; data in Lezak, 1983; Spreen & Strauss, 1991), as well as for same subcomponents of the Wechsler Memory Scale - Revised version (Cullum et al., 1990). In contrast, data which are presented for the Wechsler Memory Scale as a whole (Albert et al., 1987), do not reveal a clear aging effect. This serves, as with the observations on the composite Digit Span score above, to highlight the limited utility of using composite scores of this nature in a neuropsychological context due to the averaging effect across less and more age-sensitive subdomains within a multifaceted cognitive realm such as memory.

Further, steady age declines are noted for short-term immediate and delayed verbal memory, on the Buschke Selective Reminding Test, and the Rey Auditory-Verbal Learning Test (Mitrushina, Satz, Chervinsky & D'Elia, 1991; data in Lezak, 1983; Spreen and Strauss, 1991); for short-term immediate and delayed intentional visual memory on the Benton Visual Retention Test - Revised, and the Rey Visual Design Learning Test (data in Lezak, 1983;

Spreen & Strauss, 1991); and for short-term incidental visual recall on the Rey Complex Figure Recall, the Rey-Osterrieth Recall, and Digit Symbol Incidental Recall tests (Shuttleworth-Jordan & Bode, 1995b; Troyer, Graves & Cullum, 1994; data in Lezak, 1983; Spreen and Strauss, 1991; Vazzana, Bylsma, & Kawas, 1995). In contrast, age declines are noticeably absent for recognition memory on the Buschke Selective Reminding task for females (although not for males), and recognition on the Rey Auditory-Verbal Learning task (data in Spreen and Strauss, 1991).

With respect to new learning ability, on serial digit learning tasks there is greater sensitivity to aging changes than on simple digit span repetition tasks (Benton, Eslinger & Damasio, 1981; Shuttleworth-Jordan, 1992). Similarly for working memory, the trend of an age effect was somewhat more noticeable for Digits Backward than the very marginal decremental trend due to age noted on the same sample for Digits Forward (Shuttleworth-Jordan, 1992). Further confirming the age-sensitivity of working memory is the very severe age effect present on the WAIS Digit Symbol subtest, and Trail Making Test Part B of Reitan (1958), which is considered to be due to the contribution of a significant working memory component in both of these visuo-perceptual tasks (Cornfield & Shuttleworth-Jordan, 1996; Shuttleworth-Jordan & Bode, 1995b). (For remote memory, age-graded normative data are not presented in any of the three normative data collations or major normative data studies cited in the present review).

Language skills. As with memory, language is not a unitary phenomenon and consists of a highly complex array of skills which are differentially affected across the life span. Semantic and phonologic abilities are well preserved with age, and this has promoted the erroneous impression that generally language ability is maintained with age. Preserved abilities are evidenced by horizontal graphs of performance across age stages on traditional vocabulary tests (Lovelace, 1992), and on tests such as appreciation for face reading in comprehension (Obler & Albert, 1985). However the complexity of the situation is evidenced in the presence of a U-shaped curve for age effects in word definition and elaborateness of speech, as well as in metalinguistic skills, with optimal performance occurring around 50, and less optimal performance in the younger- and older-than 50 age stages (Obler & Albert, 1985). In addition, the more complex the language task, and the more it involves cognitive aspects which extend beyond pure semantic ability, the greater is the age effect. Thus language skills show age effects when a high memory load is involved (Nolan & Blass, 1990), and also when word finding or working memory components are involved (Lovelace, 1990). Generally word finding problems are a robust age-related finding (Light, 1990). However the ability appears to be differentially affected depending on the task: confrontation naming shows little or no age effect, whereas results on verbal fluency tasks show significant age effects (Huff, 1990).

Normative data on single neuropsychological language tests including older adult age groups, and with accompanying standard deviations, are presented only for a limited number of verbal functions. These data are broadly supportive of the above age-related observations from the cognitive aging literature. Central tendency scores from Albert et al. (1987) and Troyer et al. (1994) on WAIS Vocabulary indicate a steady trend of decline across the adult age groups. However, the declining trend in these two studies does not reflect the presence of significant age effects in this modality which would stand in strong contradiction to the relative age-stability of this task as noted in the cognitive aging literature. This is because, first in Albert et al's study there is grossly disproportionately increased variability evident in the oldest age group, and the scores of the middle age groups all lie within one-third of a deviation of each other; and second, all the scores on Troyer et al.'s data are less than one-third of a deviation apart. On the Boston Naming Test (Albert et al., 1987; Au et al., 1995; Van Gorp, Satz, Kiersch, & Henry, 1986), only a mild to moderate age effect after the age of 70 is noted. On the Controlled Word Association (data in Spreen & Strauss, 1991), differential age-related effects are revealed depending on the educational level of the subjects: there is a strong trend of aging effects from 50 years for the low education sample, whereas only marginal age effects appear after 70 on the high education sample. The data from Albert et al. on a cross-section of educational levels suggest age declines on verbal fluency from age 50.

Visual skills. Visual cognitive skills other than within the memory-loaded context, are conceptualized broadly as visuo-spatial tasks, and incorporate the visuo-perceptual functions of recognition, searching and scanning, in addition to visuo-organizational and visuomotor aspects of performance. Whereas perceptual-motor slowing is strongly associated with aging, there is little in the current evidence to indicate fundamentally decreased ability in the elderly to *think* spatially. The fact that research findings on visuospatial tasks show decline after the age of 65, is considered a likely consequence of the visuo-perceptual and speeded nature of many of the tests, and the extent to which they often need to draw on memory systems (Ogden, 1990). Accordingly, performance on judgement of line orientation and copying complex figures which do not involve speed and are amenable to verbal organization, appear to be stable with age (see Lovelace, 1990); tests of visuospatial ability as measured by the WAIS Block Design and Object Assembly subtests show reduced (but not eliminated) decline with age when time limits are not applied (Nolan, & Blass, 1990). Taken together, the dissociations implied by these task findings suggest that rather than spatial ability per se, factors such as speed as well as visuo-perceptual and /or novel problem solving skills are contributing to the decrement on the two Wechsler visuospatial tasks.

Neuropsychological normative data studies provide clear support for the above dissociations in visuo-spatial tasks noted in the cognitive aging literature. Data in Spreen and Strauss (1991) on spatial orientation tasks of Right-Left orientation show marked stability with age in the untimed Benton (1959) form; in contrast the timed Culver (1969) form indicates a steady age decline. Data on the untimed Embedded Figures Test of Spreen and Benton (1969) which taps visual search ability reported in Spreen and Strauss (1991), and norms for the untimed copying task of the Rey Osterrieth Complex Figure presented in Van Gorp, Satz and Mitrushina (1990), are stable across older age groups. In contrast, pronounced age effects are evident on WAIS Digit Symbol subtest mean score norms (Shuttleworth-Jordan & Bode, 1995b), a speeded test which calls upon visuo-spatial searching skills. Further normative data on the Trail Making Test of Reitan (1958), a timed test which also requires visual search skills, show clear age effects on both Trails A and B, which are particularly pronounced on the more complex double-tracking Trails B task (Cornfield & Shuttleworth-Jordan, 1996; Van Gorp et al., 1990). As noted earlier, the very severe age effect which is present on the Digit Symbol and Trails B tasks can probably be accounted for by the additional contribution of working memory components in both of these tasks.

Problem solving. Problem-solving refers to a multifaceted higher order cognitive mechanism, which broadly defined involves to a greater or lesser extent (depending on the task in hand), the following functions: reasoning, abstraction, perception of the problem, the ability for perceptual shifting, the generation of hypotheses, the recognition of solutions, retrieval of information from long-term memory and storage in short-term memory, manipulation of the material in short-term memory (i.e. working memory), and the executive mechanisms of rational ordering, planning and programming (delineation after Cronin-Golomb, 1990). Intricate qualitative analyses of problem solving styles reveal that older adult reasoning is 'different' and not necessarily worse than young adult reasoning, and that reasoning may improve in early old age, and only show true decline in very old age. Overall, however, in spite of the many tasks which are used to explore problem-solving mechanisms in older individuals, the general finding from traditional cognitive research is that problem-solving deteriorates with age (Cronin-Golomb, 1990; Lovelace, 1990). The ineffective spontaneous use of organizational strategies is identified as a probable source of poor performance in other areas such as memory acquisition, search and retrieval (Arenberg, 1977; Mitrushina & Satz, 1989; Poon, 1985), and all novel (that is fluid) intellectual tasks (Reese & Rodeheaver, 1985).

Due to the higher order and multi-faceted nature of the problem solving construct, neuropsychological tests involving problem-solving ability fall into each of the attention, language, visual and memory modalities which have already been reported on above. Thus in the interpretation of age-related findings, tests need to be evaluated in their own right for the hypothesized contribution of age-sensitive problem solving skills in the particular task performance. In particular, the significant age effects as cited above from normative data studies on verbal learning tasks (which involve organizational strategies for effective execution), and a test such as the Trail Making Test (which involves rational ordering), can probably in part be accounted for due to decrements in problem solving ability. Normative data presented in Spreen and Strauss (1991) on the Stroop test (Stroop, 1935), which calls upon perceptual shifting skills, show marked age effects from the age of 50 and again at 70; on the Wisconsin Card Sorting Test (Grant & Berg, 1948), which tests abstraction and ability to maintain perceptual set, aging effects appear erratic until 80 years when decline is noted. A clear age effect is evident on the Category Test as reported in Fromm-Auch and Yeudall (1983).

Albert et al. (1987) have reported aging data on a number of tests to evaluate functions which fall into the problem-solving modality as defined here. From their figures age-related decremental trends are evident on Gorham's (1954) Proverb Interpretation Test from age 60, and marginally from age 60 on Feldman and Drasgow's (1951) Visual-Verbal Test of concept formation. Stronger aging decrements are noted on the latter test in Troyer et al., 1994. On the Competing Motor Programs task, used by Albert al. (1987) to test shifting of perceptual set, there was no evidence at all of an aging effect. Rather than providing contradictory evidence for the overall relatively robust finding of decline in problem-solving ability with age, this result suggests that sustained attention which tends to hold in older age, may be a significant overriding function in the effective execution of this task.

Motor activity. Slowed reaction time is probably the most replicated effect of aging, and it has been proposed that underlying this effect, are *cognitive-motor* deficits in response planning, programming of motor activities, and in the control and execution of movement (Lovelace, 1990). Further it appears that older people show greater caution than younger persons when monitoring and inspecting signals prior to responding. Thus in other words, rather than deteriorated muscular factors, it is cognitive and emotional factors involved in the speed of decisions which limits the initiation and enactment of everyday motor tasks (Welford, 1977). Consistent with this is the fact that when older people receive warning to allow preparation for a response, reaction times are greatly reduced. However, when response complexity is increased, the age effect increases, and older people find it disproportionately more difficult to execute two oppositional movements. A recent report (Fozard, Vercruyssen, Reynolds, & Hancock, 1994) on cross-sectional and longitudinal analyses of reaction time data from 1,265 community-dwelling subjects aged 17-96 years, indicates that there is slowing of simple reaction time and relatively greater slowing of disjunctive reaction time across the decades for both males and females.

The above findings from cognitive research laboratory experiments are based on tests geared for the investigation of fine speed reactions, whereas the hand-motor responses required for many neuropsychological tests (for example the timed performance tasks on the WAIS), are relatively gross, and such tasks will be predominantly affected by higher-order cognitive functions, including the *speed* of higher-order cognitive processing, rather than the efficacy and speed of pure motor functions. Although as described above, the speed of motor processing and cognitive processing are intricately linked and thus will have interactive effects. Again neuropsychological normative data are highly commensurate with the observations from the cognitive aging literature. For the speed of hand-motor functions using the Grooved Pegboard Test (Tiffin, 1968) and Finger Tapping Tests (for example, Reitan, 1969), generally there is support for significant slowing with age, including fairly consistent, slightly superior performance with the preferred hand (Shuttleworth-Jordan & Bode, 1995c; data in Spreen & Strauss, 1991). When both hands are used together for the Grooved

Pegboard Test, and when the task is complicated by assemblies, the age effect is disproportionately large compared to the regular performance with either hand on its own (Strauss & Spreen, 1990, in Spreen & Strauss, 1991).

Summary. In summary, the examination of patterns of deficit within specific functional modalities is highly consistent with the more general patterns described above in the broad terms of stable 'crystallized' versus unstable 'fluid' functions, and the cascade model of increasing decrement with aging from 'verbal' abilities, through 'reasoning' skills and down to 'perceptual speed'. Each of the six delineated functional modalities ('attention', 'memory', 'language', 'visual skills', 'problem solving' and 'motor activity') consist of functional sub-modalities and levels of complexity, which are differentially effected due to aging processes. The cognitive and neuropsychological research literature taken together, consistently reveal normative decremental mean score trends across these six functions as follows (see summary of these trends, Table 3.1, p.69): The domains which appear relatively stable with age are those of sustained attention, recognition memory, semantic and phonological language, relatively pure unspeeded visuospatial skills, and unspeeded visuospatial copying tasks; those domains that reveal significant aging effects are general and divided attention, verbal and visual short-term memory (immediate and delayed), verbal new learning, working memory, word naming, verbal fluency, visuospatial processing and search, abstraction and reasoning, and hand motor activity. Overall this review has demonstrated a robust, domain-specific, normative pattern of cognitive decrements which occurs in association with aging.

3.2.3.3 Neuropsychological integration of neural and psychometric findings

Inferences from the cognitive neurosciences and neuropsychology, about the nature and localization of neural attrition in aging, have come from the attempt to link cognitive deficit patterns from psychometric findings (as reviewed in the previous subsection) with specific neurological findings. Hence taken together, the aging decremental trends that have been presented here with respect to intellectual functioning in general, as well as for the delineated specific functional modalities, can now be considered further in that light with the goal of identifying implications for inter-individual variability. For the purposes of this discussion the integrative frameworks are divided into, 'right hemisphere', 'frontal/subcortical', 'single versus multifactor' postulates, and 'brain reserve capacity' theory.

Table 3.1. Relative age-sensitivity of cognitive tasks: Domain-specific average ability trends for relatively high functioning healthy older adults.

Less Age-Sensitive Tasks	Average Trend	More Age-Sensitive Tasks	Average Trend
GENERAL INTELLECTUAL FUNCTIONING Crystallized Abilities (Long established habits)	**Stable**	**GENERAL INTELLECTUAL FUNCTIONING** Fluid Abilities (Novel problem solving, perceive new relationships)	**Decrease**
ATTENTION Sustained Attention	**Stable**	**ATTENTION** General Attention	**Decrease**
		Divided Attention	**Decrease**
MEMORY Recognition	**Stable**	**MEMORY** Short-term verbal + visual, Immediate	**Decrease**
		Short-term verbal + visual, Delayed	**Decrease**
		List Learning	**Decrease**
		Working Memory	**Decrease**
LANGUAGE SKILLS Semantic & Phonological Abilities	**Stable**	**LANGUAGE SKILLS** Vocabulary/Verbal Meaning	**Decrease**
Crystallized (eg synonyms/antonyms)	**Stable**	Speeded verbal tasks with problem solving	**Decrease**
		Word Naming	**Decrease**
		Word Fluency	**Decrease**
VISUAL SKILLS Pure spatial ability	**Stable**	**VISUAL SKILLS** Speeded visuo-perceptual	**Decrease**
Unspeeded visuo-spatial copy	**Stable**		
PROBLEM SOLVING Nil identified	-	**PROBLEM SOLVING** Abstraction (eg Wisconsin Card Sorting)	**Decrease**
	-	Inductive Reasoning	**Decrease**
	-	Classification (eg Similarities)	**Decrease**
MOTOR ACTIVITY Nil identified	-	**MOTOR ACTIVITY** Reaction Time	**Decrease**
	-	Hand Motor Activity	**Decrease**

Right hemisphere hypothesis. The age-related decline in performance IQ has been debated as an indication of selective right hemisphere dysfunction in older adults. However there is no neurological evidence of note to support this notion, and to the contrary researchers have found focal slowing in the left temporal region on EEG readings (see discussions in Nolan & Blass, 1992; Hinkin et al., 1990; Zarit et al., 1985). Further contradicting the 'right hemisphere hypothesis' of aging as argued in Nolan and Blass (1992), is research that has indicated some improvement in performance when the speed criterion is removed from the performance IQ tasks, and research that has systematically eliminated differences between neuropsychological tests of lateralized and focal function (familiar versus unfamiliar material, timed versus untimed tasks), which suggests that cognitive declines are not lateralized. As with alcohol dementia (see Walsh, 1991), there appears to have been an evolution away from right hemisphere models based fallaciously on equating the low multi-faceted performance IQ with pure hemispherically-related visuo-spatial dysfunction, towards more frontally based models.

Frontal/subcortical hypotheses. Support for the less lateralized, and more frontally-based hypotheses of aging comes from a number of sources. The similarity between the neuropsychological performance of normal aging and that of frontally impaired patients has been noted (Albert & Kaplan, 1980; Cronin-Golomb, 1990; Moscovitch & Winocur, 1992), although it is pointed out that there are differences in the aging subjects which confirm the involvement of additional brain regions including the hippocampus and locus coeruleus which have also shown cell loss in normal aging. Moscovitch and Winocur, for example, have conducted an intricate analysis of the expected memory deficits from frontal and hippocampal brain damage and linked this to memory deficits identified with aging. They conclude that psychometric findings support the presence of aging patterns which accompany damage to both frontal and hippocampal systems, which is consistent with the evidence for neural deterioration noted in each of these structures with aging. Importantly, these authors postulate further, that severe aging deficits noted in the 'working-with-memory' component, (a term they prefer to working memory), are explicable in terms of the dual involvement of frontal and hippocampal structures, in that the frontal lobes are required to work with the memory input and output from the hippocampal system.

Further, Hartley (1992) suggests that a source of attention decrements in aging may fundamentally be due to a breakdown of organization at the neurochemical level. He cites evidence for impairment in the norepinephrine system in older organisms, and hypothesizes that this would cause slowing on tasks involving brain areas such as the prefrontal and temporal cortices with projections from the locus coeruleus. Woodruff-Pak (1990), link spatial memory deficits to hippocampal damage, short-term memory deficits to impairments in the cholinergic neurotransmitter system involving nuclei in the basal forebrain and prefrontal cortex. Satz and his co-researchers (Hinkin et al., 1990; Mitrushina et al. 1994) present an argument for frontal-subcortical changes based on parallel patterns on neuropsychological test findings representative of all the functional modalities delineated above, between older adults and a young AIDS population with postulated subcortical pathology.

Thus taken together from the standpoint of the neuropsychological localization of function, the neural and psychometric evidence tends to support non-lateralized, relatively localized brain changes with aging, which include frontal, temporal, hippocampal as well as subcortical involvement. From this review it is clear that the neuropsychological model, which links psychometric test patterns with neural mechanisms, has provided some illuminative indications about the nature of neural aging. Hence, as noted in the introduction, its importance as an approach which has significant heuristic potential for future research in aging is strongly endorsed (Craik & Jennings, 1992; Rabbit 1993b). At the same time, however, the usefulness of this approach should not be exaggerated and its limitations need to be acknowledged. As Salthouse (1991) has emphasized, the fact that the biological and physiological aspects of aging in themselves are still relatively ill-defined, renders attempts to correlate neural aging changes with complex cognitive behavioural correlates necessarily tenuous. Further on this theme, Huppert (1990) imposes an important note of caution about the soundness of neuropsychological localization hypotheses when pathological changes are diffuse. She emphasizes that whilst the neuropathological changes in aging are relatively localized, they are fundamentally also known to be diffuse. Thus due to the immense complexity of a mixture of diffuse pathological processes, attempts at highly specific localization of function have dubious validity. In similar vein, Hartley (1992), has highlighted the extreme difficulties involved in making inferences about attention and its neural connections, in that different aspects of attention are hypothesized to be involved in a variety of possible brain networks.

It is clearly valid to *exclude* an over-localized inference such as the right hemisphere hypothesis by demonstrating a spectrum of deficit which cannot be explained on the more localized basis. In addition, validity of neuropsychological inferences are enhanced when clinical dissociations can be established for patterns of deficit, as for example Mitrushina, et al. (1994) have done: these researchers describe a parallel pattern for AIDS and normal aging, which was different to the pattern of deficit in patients with Alzheimer's dementia, and on this basis argue that subcortical involvement such as characterizes the neuropathological conditions associated with AIDS patients is true also for normal aging. However, the important point to be taken here, is that even in the demonstration of such diagnostic dissociations, there are limitations on the extent to which inferences can be drawn about specific neuropathological processes in neural aging on the basis of central trends in cognitive test performance: firstly, decrements on one function, or set of functions, may represent a number of different pathological processes; and secondly, it is not possible to predict how the interaction of different stages and combinations of neural attrition resulting in complex cognitive behavioural patterns will be reflected on psychometric test performance. Thus from the perspective of this thesis, it is considered that the study of variability will serve significantly to refine the description of cognitive aging decrements as they occur across functional modalities and between age groups on neuropsychological tests. However, in terms of the degree of specificity with which valid connections can be made between patterns noted in cognitive test performance and the complex neuropathological processes associated with aging, the same limitations will apply as for inferences that can be drawn from the study of central trends.

Single versus multifactor hypotheses. From the angle of less localized and more thoroughly diffuse conceptualizations of neural aging, has evolved the slowing with age theory of Salthouse (Salthouse, 1985a; 1985b; 1992b). As indicated above, a highly robust general finding on tests of intellectual functioning has been slowing of information processing with increased age, and this has been elaborated by Salthouse into a speed hypothesis as an explanation for cognitive decrements in aging in general. The postulation is that there are changes in the central nervous system which result in slowing down in general, and consequently there is an overall decline in intellectual level. However, whilst the strong *association* of slowed information with intellectual decline cannot be disputed, the subsequent deduction of speed as a *causally* related single mechanism which can explain all intellectual decline in aging has been strongly challenged as fallacious, and in its place a network model of multiple interactive mechanisms has been proposed as the more logical explanation for aging processes (Cunningham & Tomer, 1990; Maylor & Rabbitt, 1994; Nettelbeck & Rabbitt, 1992; Rabbitt, 1990; Rabbitt, 1993a; 1993c; Rabbitt & Goward, 1994; Robbins, James, Owen, Sahakian, McInnes & Rabbitt, 1994).

Research indicates that speed alone is insufficient to account for all age changes in cognitive functions (Cunningham & Tomer; Heaton et al., 1986), a point which is also clearly acknowledged in Salthouse's more recent theoretical treatise (Salthouse, 1991, p. 83). Rabbitt and his cognitive research associates provide elegantly designed research investigations into the Salthouse (1985b) speed hypotheses by means of which they demonstrate dissociations in age effects in memory function (Maylor & Rabbitt; Nettelbeck & Rabbitt, 1992), differential factor structures (Robbins et al., 1994) and qualitative dissociations in choice reaction times (Rabbitt & Goward, 1994) which cannot be attributed to processing speed. In this way they provide support for a *network model of parallel processing*, rather than a single factor model of serial processing, to conceptualize the cognitive aging process. The network model includes the speed factor as an important mechanism in the operation of cognitive processes, but it is seen to work in interaction with other independent performance parameters. In final support of a multiple mechanisms framework to account for intellectual decline in aging Rabbitt suggests that it is not logical from a neuropsychological perspective to support a single factor theory (Rabbitt, 1990; 1993c). He reiterates the well-known fact that the brain as a highly differentiated system in which distinct anatomical structures support separate cognitive functions, and that, as has been documented earlier in this chapter (section 3.2.2), the established evidence is that aging brain changes are not completely diffuse and global, but rather affect "to different degrees, separate sub-systems and so, presumably, the cognitive functions they support" (Rabbitt, 1993c, p. 23).

In the final analysis, Rabbitt's arguments for a multifactor network model of cognitive aging appear compelling, even without the fine dissociations of cognitive functions from the speed factor which he and his associates have demonstrated. This is a model which can account for the multifaceted nature of the relatively independent, yet diffuse partial processes involved in neural aging (described in section 3.2.2 above), which result in the spectrum of differential cognitive changes (reviewed in section 3.2.3 above). Together these neural and cognitive processes as conceptualized in a multifactor network model, inherently speak of multiple sources of variability in cognitive test performance that may occur between individuals.

Rabbitt's model however, whilst broadly supporting the theme of heterogeneity in aging, is fundamentally located in the cognitive, information-processing theoretical domain, and thus is not useful in terms of specific hypothetical indications for the present study. For this purpose it is possible to call upon a theory of brain reserve capacity, as formulated by Satz (1993). Whilst not specifically focused on the issue of the aging process itself, Satz's fundamentally *neuropsychological* set of postulates is of direct relevance to the present thesis.

Brain reserve capacity theory. The variable nature in which neuronal changes are expressed in behaviour links directly into the concept of *brain reserve capacity* (BRC) as formulated into a theory by Satz (1993) from the fundamentally neural perspective, and supported significantly by research from the neuropsychological/brain-behaviour relationships, perspective. It is necessary to differentiate the term as used by Satz (1993), and the brain reserve capacity construct as delineated by cognitive aging theorists (for example, Kliegl, Smith, & Baltes, 1989; Parks, 1992). The latter construct refers to detailed mechanisms of mental energy reserves and processing resources, whereas Satz's concern is focused rather on underlying neural processes. From the neural perspective, Satz acknowledges that the concept of brain reserve is not new. However, he advances for the first time a comprehensive review of the literature in the area, and an integrative formulation of the concept into a theory with postulates which can be empirically investigated. *Thus for the purposes of this thesis the term 'brain reserve capacity (BRC) theory' will be used to refer to the construct of brain reserve capacity specifically as formulated into a theory by Satz (1993).* Satz points out that the formulation of general threshold theory has received little attention to date, and he poses his own support for the theory at this stage tentatively.

In terms of Satz's formulation, the BRC concept is linked to the notion of a *threshold factor* which exists prior to the presentation of symptoms due to disease in the central nervous system. Of particular interest to Satz is the idea that *individual differences exist in terms of brain reserve that account for variable instances of protection from or vulnerability to symptom onset, and hence in the pre-clinical stages of disease serve to alter the symptom threshold between individuals*. This supposition, which clearly is of core relevance to the present thesis, provides an explanation for individual differences in risk of morbidity, including individual differences in risk of impaired cognitive test performance, in association with neural damage.

Satz treats the term brain reserve capacity (BRC) somewhat broadly, as a hypothetical construct that is related to adaptive behaviour, and for which the two psychosocial factors of general intelligence and educational level represent indirect measures. He notes, in addition, that the concept implies a measure of neuronal efficiency which can conceivably be operationally defined and measured by means of overall brain size. He cites literature to support this notion such as that for example of Katzman, Terry, et al. (1993), in which 10 patients with incipient Alzheimer's Disease diagnosed on autopsy, but who had evidenced preserved cognitive and functional performance, had higher brain weights and larger pyramidal neurones than a non brain-injured control group. Further reflecting on the potential for a neural operational link with BRC, he makes reference to literature in which there are beneficial cortical effects (dendritic arborization of distal branches of neurons) due to enriched environmental conditions (Greenough, Juraska, & Volkmar, 1979) and in association with higher cortical functions (Scheibel, et al., 1983; 1985). On the basis of an extensive review of autopsy, PET, MRI, and CT scan studies, and studies of the effects of education in interaction with a variety of phenomena such as Alzheimer's dementia, AIDS, traumatic head injury, chronic alcohol and drug abuse, Satz has formalized a number of postulates which form the basis of his brain reserve capacity theory, and for the present purposes these are synthesized as follows:

In the event of neuronal disease,

(i) The greater the brain reserve capacity and presumed redundancy of neuronal networks, the less likelihood there is of showing neuropsychological impairment because greater BRC acts as a protective factor and decreases the risk of functional impairment;

(ii) The less the brain reserve capacity, the greater likelihood there is of showing neuropsychological impairment because less BRC acts as a vulnerability factor creating greater risk of functional impairment;

(iii) The aggregation effect of disease and sufficient challenge will lead to a greater likelihood of showing neuropsychological impairment since reductions in BRC due to neural mechanisms in conjunction with appropriate challenge will serve to increase vulnerability to functional impairment.

For the most part Satz has structured his discussion of the above hypothetical formulation of his theory around symptom onset in head injury. However, the threshold concept has central applicability to the behavioural presentations of neural loss in association with age changes, in that neuronal attrition due to aging can be conceptualized as serving to reduce BRC, and hence, depending on an individual's availability of brain reserve, will result in variable behavioural manifestations of aging. Apart from the fact that age clearly represents a central risk factor in the onset of Alzheimer's dementia, the notion of reduced brain reserve and lowered disease threshold due to aging, is strongly promoted by Satz on the basis of a number of studies. For example he cites the study of Gedye et al. (1989) who showed in a retrospective study of probable Alzheimer's cases, that individuals who sustained a head injury after the age of 65 but before the onset of Alzheimer's dementia, showed particularly rapid symptom progression; and that of Corkin et al. (1989), who noted enhanced cognitive deterioration following head injury in young adulthood that did not present until ages 40-50, suggesting the delayed effects of prior neuronal attrition as a function of age. Thus a clear theme that emerges from Satz's review is that the neuronal attrition due to normal aging constitutes one of the key factors which serve to reduce brain reserve capacity, and to increase the risk of functional impairment.

Further with regard to risk factors and the lowering of the brain reserve threshold, Satz cites studies which support the aggregate effect of diminished neuronal reserves and lower levels of education (and by implication IQ), gender effects and the effects of what he terms 'challenge'. The reader is referred to Satz's (1993) review article for a full rendering of these studies, and only a number of examples will be isolated here for illustrative purposes. With respect to *education*, for example, he describes the studies of DeObaldia and Parsons (1984) and Grant and Adams (1986) which indicate that low education probably aggregates with alcoholism to cause more severe neuropsychological outcome; and the study of Satz et al. (in press) who report a higher predicted prevalence of cognitive abnormality in AIDS patients with less than 12 years of education compared to those with higher education. Regarding *intelligence,* although Satz does not report any studies which specifically implicate this variable as a risk factor in relation to symptom outcome, for the purposes of his BRC theory (see above), he does use the psychosocial factors of education and intelligence as roughly equivocal. Further he highlights the "ubiquitous correlation of .70" between the factors of education and IQ, and draws attention to the "long known" fact that "lower educational level and IQ are related to a shorter life span and increased psychological and physical morbidity" (p.281). *Gender* influences are not discussed by Satz in great detail. In this regard he only makes reference to studies which support the fact that gender is a risk factor in tardive dyskinesia although he does not specify in which direction, and that female gender is a risk factor in Alzheimer's dementia (Zhang et al., 1990; Gerlach & Casey, 1988). Whilst not addressed by Satz, it is of relevance here to note the commonly-known fact of male disadvantage relative to females in terms of longevity (Dulbecco, 1991; Kristof, 1996). This is attributed to genetic as well as behavioural factors, and the shorter life span for males is enhanced among blue-collar male populations (Dulbecco, 1991). Together with Satz's point noted above, in which he links shorter life span for those of lower educational/IQ level with increased psychological and physical morbidity, the implication is that there would similarly be an increased risk of such morbidity for males. Finally Satz introduces the notion of increased *challenge* which serves to lower the brain reserve capacity threshold and reveal evidence of functional impairment. In order to illustrate this phenomenon he describes the 'challenge' tool of the stress electrocardiogram used in cardiovascular medicine. This technique incorporates an exercise tolerance test, which lowers physiological reserve capacity thus increasing sensitivity to symptomatology, and hence serves to predict risk in myocardial infarction. Similarly in aggregation with subclinical conditions of neural disease, Satz makes the point that a cognitive task of appropriate challenge could serve to increase vulnerability to the demonstration of cognitive impairment.

In sum, therefore, it emerges from the Satz review, that low education and IQ, and appropriately high task challenge will serve as threshold lowering factors and enhance vulnerability to symptom onset. With respect to gender, Satz indicates that there may be differential effects, but he provides no specificity about the direction of such effects. Tentatively, by inference, however, it appears that in addition to low education/IQ and high challenge, male gender membership may serve to enhance vulnerability to symptom onset.

Hypothetical indications for the present study. Thus in terms of the postulates from Satz's brain reserve capacity (BRC) theory which are supported on the basis of research studies such as those cited above, it can be argued, not only that there is likely to be significant inter-individual variability in cognitive test performance with aging in the general terms as implied also in Rabbitt's multifactor model, but more particularly that this variability will be differentially affected in interaction with the brain reserve capacity *threshold altering factors* of lesser versus higher levels of *education and IQ*, different levels of *gender*, and lower versus higher levels of *task challenge* (such as presented via differing levels of complexity of cognitive tasks). Thus, by way of conclusion, it is possible to elaborate further on the Satz brain reserve capacity postulates as listed above, taken in conjunction with the central indications which have emerged from the broader review on neural and psychometric perspectives on aging, in order to produce a number of more specific hypotheses on inter-individual variability in cognitive aging as follows:

(i) From the neural perspective on aging it is hypothesized that due to the inherent inter-individual variability in the complex physiological mechanisms of neural attrition due to aging, as with all forms of brain impairment, there is likely to be increased variability of scores on cognitive tasks in association with aging;

(ii) From BRC theory, it is hypothesized that the inter-individual variability in cognitive test performance in association with the neural aging process will be influenced by inter-individual differences in levels of brain reserve capacity which alter the threshold of vulnerability to, and protection against, the onset of functional symptom presentation. In particular, from the Satz review it is expected that low education and IQ, high task challenge, and by inference male gender, will serve as threshold-lowering factors and will thus be associated with earlier symptom onset with aging; and conversely it is expected that high education and IQ, low task challenge, and by inference female gender, will act as threshold-raising factors and will be associated with later symptom onset with aging. (The Satz review did not provide a definitive direction for gender effects, however these were inferred from the well-known fact of female longevity in conjunction with general BRC principles).

(iii) Finally, from the psychometric perspective in conjunction with BRC theory, it can be hypothesized that differential sensitivity to aging processes <u>between</u> different cognitive tasks and functional modalities, as identified on central trends analyses, will have an associated differential effect on the patterns of variability in cognitive task performance. This is in that tasks which are less sensitive to aging on central trends analysis are evidently those that present less challenge to brain capacity thresholds in association with aging, and those tasks which have been demonstrated to be more sensitive to aging on central trends analysis, are evidently those that present more challenge to brain reserve capacity thresholds; hence differential effects between test modalities will be reflected similarly for both central trends and variability data in that for those tasks and modalities where central trends effects are more marked, variability effects will also be more marked and vice versa.

3.2.4 Specific perspectives on aging: Section summary

In concluding this section the following features have emerged:

(i) On the basis of neural investigations, it is apparent that the neural aging process is characterized by a collection of relatively independent partial neurophysiological aging changes that occur such that neural attrition due to aging has a combined localized and diffuse presentation in which there is extensive variability in the form and course it may take for particular individuals.

(ii) On the basis of psychometric investigations from both the cognitive and neuropsychological literature, it is apparent that neural attrition is reflected in a robust set of normative decrements which are highly domain-specific as follows: The domains which appear relatively stable with age are those of sustained attention, recognition memory, semantic and phonological language, relatively pure unspeeded visuospatial skills, and unspeeded visuospatial copying tasks. Those domains that reveal aging effects can approximately be ranked on a mild/moderate/severe aging-effect continuum as word naming, verbal new learning, verbal and visual short-term recall immediate and delayed, verbal fluency, abstraction and reasoning, visuospatial processing and search, working memory, and hand motor activity;

(iii) Taken together neural and psychometric investigations tend to support non-lateralized, relatively localized brain changes with aging, which include frontal, temporal, hippocampal as well as subcortical involvement. However it is highlighted that due to the relative diffuse nature of neural attrition, caution must be employed with respect to making highly specific inferences about cerebral aging processes from neuropsychological test data;

(iv) The multifaceted nature of neural attrition in association with its variable, and intricately domain-specific presentation in cognitive test patterns, appears to be conceptualized more satisfactorily in a multifactor network model of cognitive processes, than in a single factor model in which speed is seen to account for all cognitive changes in aging. The network model inherently speaks of multiple sources of variability in cognitive test performance that may occur between individuals.

(v) Finally, from the Satz BRC theory used in conjunction with the indications from the neural and psychometric central trends reviews, it has been possible to pose specific hypotheses for expected patterns of variability effects: it is postulated that due to brain reserve capacity threshold-altering influences of lower versus higher levels of education and IQ, male versus female gender, and lower versus higher levels of task complexity and age-sensitivity of tasks, there will be differences in the presentation of inter-individual variability on cognitive tasks in association with the neural aging process.

3.3 CHAPTER SUMMARY

In this chapter a theoretical basis was delineated for the present study. Metatheoretically the thesis was conceptualized as offering a *differential* perspective on the study of cognitive aging, to complement rather than replace the existing *normative* perspective. Accordingly, in terms of the relativistic principle of complementarity, it was posed that, as with the resolution posed for the dual wave/particle conceptualizations of light, the alternative views of 'differentiality' and 'normativity' should be considered in parallel. In this way it was argued that a more complete picture of the complex phenomenon of aging would evolve than would be possible on the basis of the exclusive use of either paradigmatic standpoint on its own.

It was noted that, from a theoretical perspective, the psychology of aging is not very advanced. The discipline is characterized by a plethora of disparate concepts on aging, and there are minimal indications of hypothesis-testing research which is generated from broad-ranging and systematically articulated theory. This problem is even more enhanced in the neglected area of research from a differential perspective. It was argued that at this undeveloped stage of psychogerontological investigation into inter-individual variability there is a *mandate* for descriptive-exploratory research, from which higher level theoretical formulations can subsequently be informed and developed. Thus it was considered appropriate to locate the present research on the descriptive-analytical-theoretical continuum as follows: (i) it was purposefully primarily descriptive, (ii) it incorporated a level of analysis with respect to the influence of education, I.Q., and gender, and (iii) it went beyond the mere 'shot-gun' data-gathering plane, in that it was located within the broad theoretical perspectives of neuropsychology and brain reserve capacity theory, and a specific model of variability, from which standpoints it was possible to generate hypotheses for further analysis and interpretation of the descriptive data.

Further in this chapter, the research findings from neural and psychometric perspectives were reviewed, and robust patterns of domain-specific functional deficit were found to characterize *normative* cognitive aging processes, which in turn appear to be consistent with the established nature of neural aging mechanisms. From a review of integrative models from the cognitive neuroscience and neuropsychological approaches to aging, the *multifactor network* and *brain reserve capacity* positions (as formulated by Rabbitt and Satz, respectively), were considered to be the most compelling. From a neuropsychological angle they both imply multiple sources of variability in cognitive test performance that may occur across functions and between individuals in the aging process. From the more neuropsychologically oriented brain reserve capacity theory of Satz it was possible to pose a number of specific hypotheses around the possibility of differential effects on the presentation of variability on cognitive tasks in relation to differing levels of gender, education, IQ, task difficulty, and task domain.

Apart from the work of Rabbitt, all the conceptualizations and attempts to link neural-aging mechanisms with their cognitive manifestations which have been reviewed in this chapter, have come from studies of *average* trends. Whilst from the Satz BRC theory it has been possible to elaborate further and propose hypotheses about causal influences in the presentation of inter-individual variability, these postulations require verification, and the more exact characteristics of variability patterns in association with the postulated causal influences from BRC theory, are still in need of demonstration. Clearly, it is the nature of individual differences in cognitive aging which should serve as the focus for further investigation, against the complementary background of the normative findings presented in this chapter. By this parallel mode of investigation it will be possible to achieve a more integrated picture of the aging phenomenon.

4

CHAPTER 4: INTER-INDIVIDUAL VARIABILITY

Having created a conceptual backdrop within the aging literature for a study on inter-individual variability in cognitive aging (Chapters 2 and 3), the purpose of the present chapter is to revisit the issue of variability in the manner in which it was initiated in the Introduction (Chapter 1). Broad issues which were raised to orient the reader to the central rationale and aims of the thesis will be expanded on in more precise detail. The chapter is introduced with a systematic review of the reasons why it is important to study variability. This is followed by a review of available research into heterogeneity in older populations. First research that describes the *pattern* of cognitive variability is presented; this is followed by a review of research findings on the *origins* of heterogeneity in older populations. From these reviews directions for further research from a methodological point of view are identified. In particular, the chapter lays the foundation for the formulation of a model of inter-individual variability for the adult cognitive aging process (which is developed in the following chapter), and from which specific hypotheses for the present study were delineated.

4.1 THE NEED FOR VARIABILITY STUDIES IN COGNITIVE AGING

In the previous three chapters there have been a number of references made to the need for taking heed of variability in aging research. The purpose of this subsection is to present a synthesis of these reasons, and to delineate them more precisely. For the purposes of this discussion they are differentiated along the following dimensions: 'psychometric', 'developmental', socio-political, 'clinical', 'models of cognitive aging', and 'aging research'.

4.1.1 The psychometric angle

From the psychometric angle (for example, Anastasi, 1958; 1991; Hoyer, 1974; Lezak & Grey 1984; Rabbitt, 1983; Verran et al., 1992), it has been argued that variability data should be studied as events in their own right. This is because to ignore them due to operation exclusively within a central tendencies paradigm, especially in a population which appears to be vulnerable to increased heterogeneity, is to risk arriving at artifactual inferences about general trends, since there will be significant numbers of individuals who are not well described by the mean (see discussion, section 2.1.1). The increased risk of Type II error when comparisons are made between groups with skewed distributions has been highlighted, and for this reason it has been suggested that non-parametric forms of analysis which take individual distributions into account, may be more informative (Hoyer, 1974; Lezak & Grey, 1984).

Rabbitt (1983) has discussed the smoothing out of the age curve that may occur artifactually and obfuscate a bimodal distribution of age-related trends across time in a manner that mimics the curve seen with continuous aging decline across all individuals. Thus should the true state of affairs across time be that with each decade there are a proportion of individuals who maintain a high level of functioning into late old age, and other individuals whose functioning falls off disproportionately early (see Figure 4.1a, below), this would be reflected in sharply increased within-subjects variance with group-mean age, but the plot of the mean score curve in the bimodal distribution (Figure 4.1a), would look exactly like that from a pattern of continuous decline which is true for all individuals (Figure 4.1b, p.84). Rabbitt emphasizes the need for studies on patterns of inter-individual variability in aging to elucidate the true nature of the aging curve.

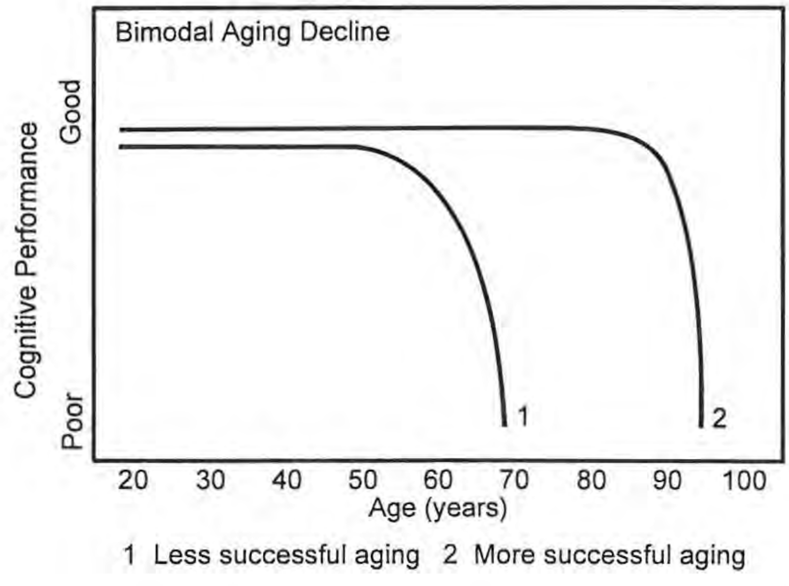


Figure 4.1a. Schematic representation of bimodal aging decline (after Rabbitt, 1983).

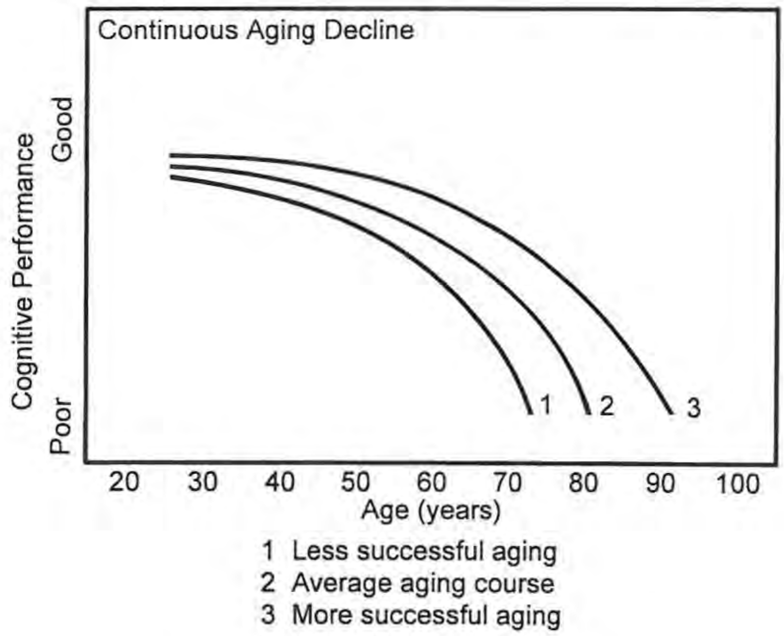


Figure 4.1b. Schematic representation of continuous aging decline (after Rabbitt, 1983).

Although Salthouse (1991) does not acknowledge the presence of increased heterogeneity with aging (except as a measurement artifact), he has also noted with reference to future directions in aging research, that there is a need to study the whole distribution of individuals at each age range, and not merely average performance. He suggests that it is conceivable, for example, that "age-related influences are more pronounced among the poorer-performing members of the population, perhaps because they are more vulnerable to a variety of factors contributing to impairments in performance", and further that "Information about age trends in the entire distribution may also help establish the generality of the results by demonstrating that the performance pattern of interest is not restricted to comparisons in particular regions of the measurement scale" (Salthouse, 1991, p. 357).

4.1.2 The developmental perspective

From a developmental perspective (Berg, 1992; Birren & Renner, 1977; Bidell & Fischer, 1992; Labouvie-Vief, 1985; Wohlwill, 1973), the general consensus is that traditional models of cognitive changes across time, which ignore the differential aspects of individual development, can no longer be supported. These researchers consider that there is the need for systematic attention to be paid to the sources of variability that are separate from age itself. On the basis of a review of studies, Labouvie-Vief (1985) argues that intelligence is an adaptive capacity, and that biological growth processes continue into adulthood given proper ecological conditions. Berg (1992) and Bidell & Fischer (1992) focus on the importance of a contextualized approach to understanding cognitive development which in effect provides the *expectation* of differences. These authors are critical of theories which are structured according to central trends such as that of Piaget, and Bidell and Fischer present an elaboration of Piaget's theory in order to take account of contextualized influences on cognitive variability. Furthermore, Birren and Renner (1977) suggest that in advanced years powerful phylogenetic factors may outweigh the effects of time, and hence variability will occur in relation to the standard of time.

Whilst developmentalist Wohlwill (1973) presents a discussion which is specifically directed towards variability in child development, he emphasizes the need to focus on individual differences in developmental processes, and to examine the stability of the 'invariance', (which is another way of saying that it is important to abstract the presence of significant variance), across developmental age stages. In order to study the differential aspect of development he identifies two aspects: The first involves the need to *describe* the characteristics and course of individual differences; the second concerns itself with addressing the *origin* of developmental individual differences. In cultivating a methodology for descriptive purposes Wohlwill presents a set of graphical illustrations to show how group variance trends might be conceptualized to differ across age stages depending on the functional modality (see Figure 4.2, p.86).

Trend 1 represents the situation in which there is increasing variability between individuals reaching a maximum at some intermediate point in development, and subsequently disappearing; Trend 2 shows a pattern of progressively increasing variability between individuals starting with the presumption of no differentiation at birth; Trend 3 reflects the situation of a stable pattern of individual variability across age stages. Wohlwill points out that this kind of descriptive graphical material is important because it can provide valuable hints as to the nature of variability trends in developmental elements, following which the sources of such differences can be investigated. Whilst he uses the illustrations to show possible forms of individuation processes that might occur across age stages in child development, such graphical representation could clearly be expanded to illuminate single function trends in variability across the adult age range.

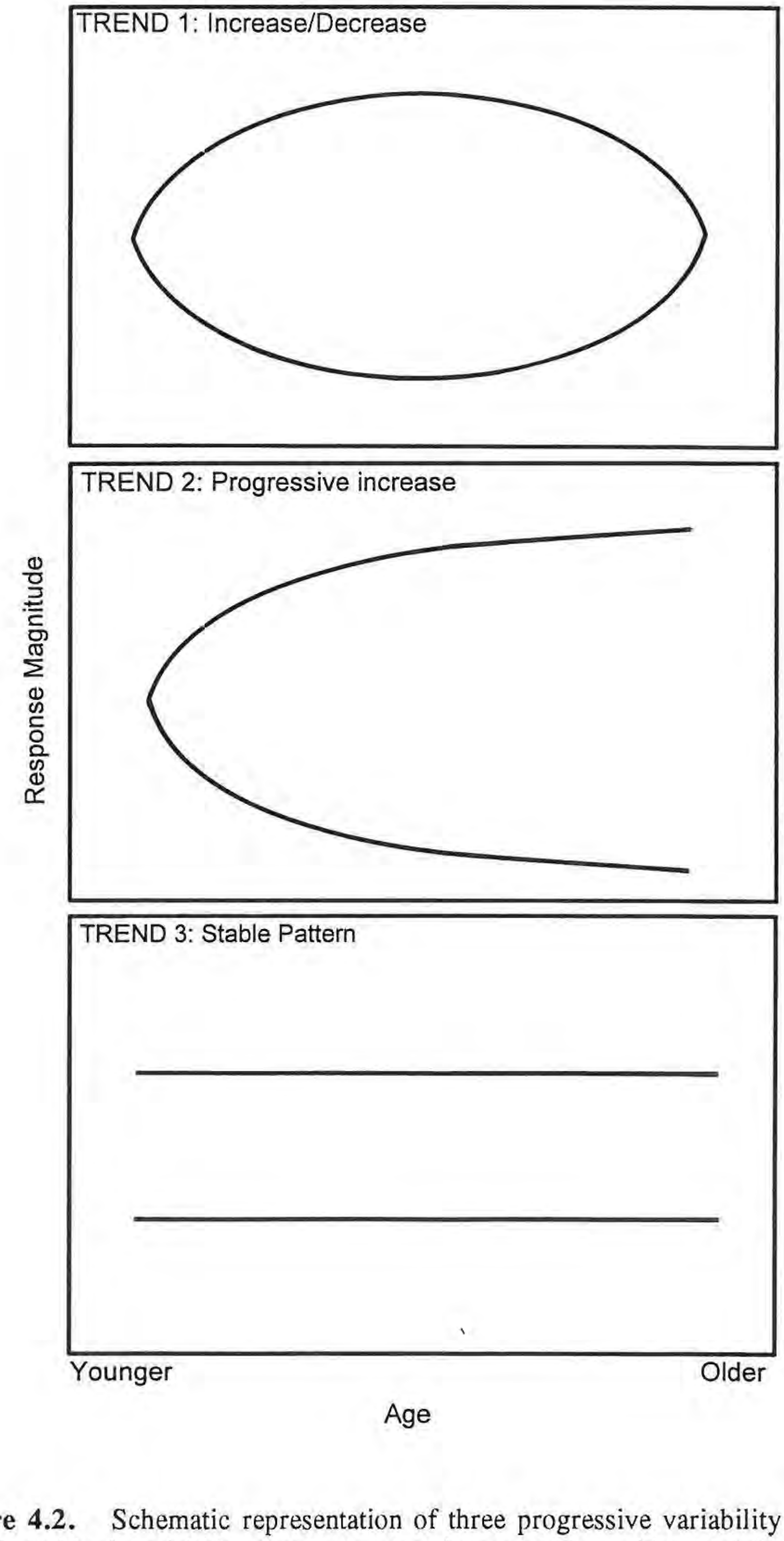


Figure 4.2. Schematic representation of three progressive variability patterns of inter-individual variability that may occur in association with age (after Wohlwill, 1973).

4.1.3 The socio-political view

From the socio-political angle, there has been a growing trend in the literature for researchers to highlight the presence of increased heterogeneity in aging populations with respect to social, psychological and health factors. In particular emergent signs of more overlap in the cognitive range between younger and older adult individuals than was initially envisaged is focused on, and hence emphasis is laid on the pressing necessity to take more account of this phenomenon in aging research (Aiken, 1995; Cunningham, 1990; Dannefer, 1988; Dolen, 1982; Gold & Arbuckle, 1990; Heron & Chown; Holland, 1990; Krauss, 1980; Labouvie-Vief, 1985; Lehr, 1980; Lovelace, 1990; Moscovitch & Winocur, 1992; Nettelbeck & Rabbitt, 1992; Owsley, 1991; Plomin & McClearn, 1990; Poon, 1985; Rabbitt, 1995; 1983; 1990; 1993a; 1993b; 1993c; Rabbitt et al., 1993; Rediess & Caine, 1996; Rowe & Kahn, 1987; Schaie, 1983; 1985; 1988a; 1988b; 1990; 1994; Schaie & Willis, 1986; Sheppard, 1978; Spirduso & MacRae, 1990; Sterns & Miklos, 1995; Thomae, et al., 1981; Willis, 1985).

The urgency of these authors' comments revolves around two central socio-political issues: Firstly, there is concern that the over-generalization of normative trends in the face of aging heterogeneity can create very skewed impressions about older populations which result in ageist discriminatory practices, and this has crucial implications for political and social policy making decisions for occupational retirement, and with respect to the nature of continuing educational programmes and recreational facilities for older people (see in particular Aiken, 1995; Hayslip, 1989; Rabbitt, 1991; Schaie, 1988a; 1988b; 1994; Sheppard, 1978; Sterns & Miklos, 1995; Thomae et al., 1981; Willis, 1985). Secondly, it creates an erroneously pessimistic view which ignores the plasticity of the aging brain and the possibility which has been revealed for intellectual improvements in older adults through the mediation of stimulating environmental situations (for example Cunningham, 1990). Thus the overriding message is that aging should not be over-endowed with a universal flavour. It has been pointed out that the *reasons* for successful versus unsuccessful aging amongst individuals on the wide continuum from successful aging to dementia need to be very carefully explored as one of the most *crucial* areas of future investigation (for example, Rediess & Caine, 1996), and that this calls for a differential rather than the traditional normative approach to studies in aging.

4.1.4 The clinical perspective

In addition to the above socio-political indications, recognition of heterogeneity in older populations and the understanding of its characteristics, translates into a number of clinical implications. Rowe and Kahn (1987) have emphasized how failure to take account of aging variability on the medical dimension may cause illness characteristics to be prematurely and erroneously accepted as the inevitable result of aging (see in Rowe and Kahn, 1987). They provide the example of cardiovascular disease which was originally associated directly with cognitive aging. Subsequently, however, it was discovered that socio-economic conditions and adverse life-styles were strongly contributing intervening variables which accounted significantly to the mental decline originally attributed to heart disease. Further these authors note, that whereas in recent years there has been a thrust towards health promotion and corresponding increases in life span, an emphasis on the idea of 'successful' versus 'unsuccessful' aging needs to be implemented as a clinical thrust which aims to go further than mere health maintenance, and strive more ambitiously for the *fullest possible functioning* of older people until the end of the life span both physically and mentally. Thus a plea has been raised for improved assessment procedures that take individual differences into account, as a prelude to designing *individualized* intervention programmes for the elderly (Hoyer, 1974). For example individual differences need to be taken into account in the implementation of practical memory training for older people with everyday memory problems (West, 1989).

A further issue of clinical note, relates to the detection of cognitive dysfunction in the elderly for diagnostic purposes. Accurate detection of the nature and presence of cognitively-related disease (such as for example the differentiation between reversible versus progressive dementia) is vital to inform appropriate medical and psycho-social interventions. As is well-recognized, this facility is very dependent on the establishment of appropriate normative criteria for older adults on psychiatric instruments (Clarke & McKenzie, 1994) and for neuropsychological tests (for example, Lezak, 1995; Kaszniak, 1990). However, in the face of large individual differences in cognitive aging, it is *also* necessary to highlight the significance of variability differences on the neuropsychological data derived across age groups. This is in order that the diminished reliability of mean scores for age stages where there is enhanced heterogeneity can be appropriately taken into account (for example Bode & Shuttleworth-Jordan, 1993; Cornfield & Shuttleworth-Jordan, 1996; Shuttleworth-Jordan

& Bode, 1995a). To date scant attention has been paid to this task as evidenced in the major sources of neuropsychological data in which variability trends are barely commented upon (see Lezak, 1993; 1995; Spreen & Strauss, 1991). In these 'bibles' of neuropsychological assessment, which merely reflect the research orientation of the studies they report on, the extent to which the central tendencies paradigm predominates in a central sector of clinical aging literature, is forcefully demonstrated.

4.1.5 Theoretical and research indications

Finally the need for research into variability emerges with respect to *theoretical indications* about cognitive aging, and out of the status of *cognitive aging research* itself.

4.1.5.1 Theoretical indications on cognitive aging

Perspectives on cognitive aging such as the multifactor network model described by Rabbitt (1993c), and the Brain Reserve Capacity (BRC) theory as formulated by Satz (1993), in conjunction with the findings from neural and psychometric research, provide compelling support for the presence of mechanisms which imply that there will be significant variability between individuals in the 'symptomatic' presentation of cognitive aging processes (see discussion, section 3.2.3.3). Rabbitt's conceptualization of aging has formed the inspiration for an energetic research drive into the investigation of variability trends in aging (for example, Maylor & Rabbitt, 1994; Nettelbeck & Rabbitt, 1992; Rabbitt, 1990; Rabbitt, 1993a; 1993c; Rabbitt & Goward, 1994; Robbins et al., 1994), with some illuminative results to be reviewed subsequently. However, this type of dynamic research endeavour, which systematically focuses on the issue variability, appears as a rare phenomenon in the literature, and the hypotheses that have been generated in this thesis from the Satz BRC theory (Chapter 3, section 3.2.3.3) have not yet been verified.

4.1.5.2 The status of cognitive aging research

In spite of all the above-listed admonitions around the need for aging researchers to take account of variability, as noted in Chapter 1, first Krauss (1980) and subsequently Nolan and Blass (1992) have commented on the trivial significance that is paid to the issue of variability *in effect* since there is a conspicuous absence of systematic research in this area. A few others for the same reason have challenged the very presence of heterogeneity as a phenomenon in older populations (Bornstein & Smircina, 1982; Montgomery & Borgatten, 1986; Salthouse, 1991), with some provocatively suggesting that this might be merely another of those illusionary doctrines that have tended to surround aging (Bornstein & Smircina, 1982). Although the comments contained in the Bornstein and Smircina three-page article are located in an extremely superficial review of the situation, without the use of any statistical analysis, the article has nevertheless been given sufficient weight to be accepted into a prime aging journal. Ironically, this article, partially due to its challenging nature, but probably in much larger part because of its total isolation as a journal contribution dedicated entirely to the topic of variability in aging, has become a much cited publication. The comments from both Montgomery and Borgatten (1986) and Salthouse (1991), for example, depend significantly on Bornstein and Smircina's review to make a case from the literature *against* the presence of increased variability in aging, in spite of the weighty literature cited in the present thesis which stands in contradiction to this. Curiously, much of the same literature in *support* of increased heterogeneity in association with aging has been acknowledged by Salthouse himself in his brief review of the issue (for example, Albert, 1988; Baltes, Reese, & Lipsitt, 1980; Baltes & Willis, 1977; Botwinick, 1967; Bromley, 1956; Kausler, 1982; Perlmutter, 1988; Rabbitt, 1982; Schaffer & Poon, 1982; Hertzog, 1985; Hoyer, 1974; Schaie, 1983; 1988; 1989b; Schonfield, 1974; Wechsler, 1952; Welford, 1957, 1958, 1959; Willis, 1985, 1989b, *all as cited in Salthouse, 1991*), and then discounted. Salthouse even presents graphical material from Schaie's Seattle longitudinal study which has been used by Schaie (1988b) and Willis (1985) to *support* the presence of increased variability, and makes no attempt to address this contradiction.

Thus a comprehensive and theoretically integrated review of research studies in the area of inter-individual variability does not appear to have been conducted. The only literature overviews available appear to be those of Bornstein and Smircina (1982), Krauss (1982) and Salthouse (1991), all of which taken together can only be described as 'half-hearted', relatively superficial and unsystematic, and none have provided statistical or theoretical support for their arguments. An urgent foray into this neglected area is clearly what is indicated to gain a measure of precision in this dispute. It appears that there are two key questions to be addressed:

(i) Is enhanced heterogeneity with aging a myth or a reality?

(ii) If it exists, precisely what form does it take?

4.2 REVIEW OF RESEARCH ON INTER-INDIVIDUAL VARIABILITY

The review of available research into inter-individual variability will be structured according to the two dimensions noted earlier that have been delineated by Wohlwill (1973) as being those aspects necessary for the study of inter-individual differential aging effects: First the findings from studies which contribute to a description of the *patterns* of inter-individual variability (that is the characteristics and course of inter-individual variability) due to aging will be reviewed; second, consideration will be given more generally to research which has relevance to the *origins* of age-related inter-individual differences.

4.2.1 Patterns of inter-individual variability in aging

> *'Demonstrations that variability between individuals increases with the ages of the groups sampled are consistent with actuarial statistics, which show that people have very different life spans and provide formal endorsement for folk wisdom that people "age" at very different rates.'*
>
> *(Rabbitt, 1993b, p. 389, Does it all go together when it goes? The nineteenth Bartlett Memorial Lecture).*

The state of investigation into specific inter-individual variability patterns in cognition in association with age can only be viewed as being in its extreme infancy. The indications from studies are few, unsystematic, and so inadequately synthesized to date, that to embark on such a review can be paralleled with the experience of initiating the mapping out of a new planet in a fog in which there are only mirage-like glimpses of odd disconnected landmarks. However, what can be gathered together from a highly disparate literature will be reported on, and an attempt will be made to synthesize the findings within the framework of functionally-related performance patterns as described by Lezak (1995) for use in clinical neuropsychological assessment. For a number of key variability trends, particularly where there are contradictory findings or dissociations, the hypothetical indications derived from the Satz brain reserve capacity (BRC) theory with specific reference to inter-individual variability in cognitive aging (see section, 3.2.3.3), will be called upon for explanatory purposes. Age-related variability data which appears in the literature without the specific purpose of serving as an investigation into variability will *not* be perused as part of this review. The emphasis rather will be on providing a conceptual integration of reports on studies or research overviews which *have* made particular reference to issues of inter-individual variability, to the extent that they exist. First broad overviews of variability trends that have appeared in the literature will be discussed, these being largely the work of Schaie, Rabbitt and Salthouse; second findings within specific functional modalities will be collated and discussed.

4.2.1.1 General variability trends

Whilst the prevalence of a general decline in mental abilities has been reliably established with advancing age on the basis of group data (as has been demonstrated earlier in the review of average trends), Schaie and Rabbitt emphasize that *this is not true for all older individuals* (for example Holland & Rabbitt, 1991b; Rabbitt, 1986; Rabbitt, 1993b; Schaie 1990; 1988b). They substantiate their argument on the basis of noting from aging studies: (i) the *degree of overlap* between older and younger populations and what *proportion* of individuals maintain their previous level of functioning, or at least sustain average functioning with respect to young adults; and (ii) *alterations in variability data*, as follows.

Overlapping distributions. Schaie (1988b) illustrates on the basis of an initial cross-sectional data analysis from the Seattle Longitudinal study (n=1629), for subjects from age 25 through to the late 80's, across a range of modalities (Verbal Meaning, Space, Inductive Reasoning, Number, Word Fluency, Immediate Recall, Life Tasks), that there is a tremendous degree of overlap in performance across the adult age range: until age 60 there is greater than 90 per cent overlap for all the intellectual variables tested which persists for a number of skills right up to age 88; only for inductive reasoning is there a complete lack of overlap noted across the older age groups.

In keeping with the theme of Schaie's findings, Rabbitt (1986) reports that on a sample of 998 healthy subjects in the age range 18 to 85 years, 10% (whom he labels the "super geriatrics"), performed as well as the most able individuals between ages 30 and 50 on a speed of reactions task. In a subsequent large cross-sectional analysis of older adults in the age decades from the 50's through to the 70's, Rabbitt graphically represents the 100% overlap in the distribution for the Vocabulary Mill Hill Test; for IQ, auditory and visual recall and learning tasks he notes in contrast, not simply a downward shift of a normal distribution in association with diminishing mean scores, but a markedly skewed distribution for the advancing decades from the 50's through to the 70's due to a disproportionate number of increasingly poorer performances across decades, but also because of a proportion of individuals who sustained a high level of functioning. He estimated that between 10 and 15% of individuals over 75 performed at the level of the best volunteers between the ages of 50 to 55 across this additional spectrum of cognitive measures. In another research endeavour, Rabbitt and his associates estimated that 7% of 70 year olds showed no detectable loss in overall IQ test performance with age when compared with an estimate of IQ based on vocabulary test scores (Holland & Rabbitt, 1991b).

Consistent with the findings of Schaie and Rabbitt, are also the reports from much earlier studies reviewed by Krauss (1980), which demonstrate that variance tends to increase with age (Anastasi, 1958; Sward, 1945), that there is a great deal of performance overlap among widely spread age groups (Anastasi, 1958), and that the overlap is domain-specific in that nearly 80% of older subjects scored at or above the mean of the younger group on a synonyms test, but only 4.6% on the symbol-digit subtest (Sward, 1945). Weintraub et al. (1994) report that in a sample of healthy volunteer physicians (aged 28-92 years), the best ten subjects in the over 75-year-old group outperformed a significant proportion of subjects in the under 35 group. These authors conclude that deterioration in memory and attentional abilities does not occur in all individuals, and that little fall-off in capacity may be seen in subjects even over the age of 75 years. Heaton et al. (1986) have focused on brain damage misclassification issues, and thus on percentages of older individuals who do *not* sustain the level of functioning of younger individuals. This is the kind of analysis, albeit important from a diagnostic perspective, that contributes to an overly pessimistic view of cognitive aging which is erroneously extended by implication to have universal application. On reframing their data in an inverse positive manner, it is apparent that at least 50% of individuals over the age of 60 (mean = 68 years) with 12^{+} years of education, and at least 30% of those with less than 12 years of education, would still be classified as normal by an average impairment rating on the Halstead-Reitan battery established for younger subjects.

Group variability data.

Data from Rabbitt. In terms of the analysis of variability data, Rabbitt (1993b; 1993c) has reported on indications from large-scale cross-sectional studies with subject numbers varying between 300 to 600, across successive age groups from the 20's through to the 70's. He notes that in association with declining means of scores on tasks of choice reaction time and the AH 4 IQ test, between-groups variability increases regularly with age as increasing age spreads the group across a wider spectrum of performance. He points out that these data demonstrate that differences in ability between the most and least able subjects steadily increases with the ages of the groups sampled (a differentiation effect), which in turn implies that individuals vary in their rates of cognitive aging. Rabbitt does not document the standard deviations in his reports, thus it is not possible to examine the exact nature of the variability trends across different modalities. Nor does it appear that he has made his global statement of age-related increases in variability on the basis of any statistical analyses. However Rabbitt's general observations are consistent with the hypothesis of expected increases in variability with increasing age in older adult groups in terms of his own multimodal network model of aging, as well as the hypothesis of expected increases in variability in association with neural aging extrapolated above (section 3.2.3.3) from the neurally based perspective of aging and the Satz BRC theory.

Data from Schaie. In reporting on the Seattle Longitudinal Study, Schaie (1983; 1988b) has examined the patterns of inter-individual variability for 7-year intervals from 25 through to the 80's across modalities including Verbal Meaning, Space, Reasoning, Number, Word Fluency, Immediate Recall and Educational/Real Life Tasks[9], for both a cross-sectional analysis of the initial data (Schaie, 1988b) and from the outcome on the longitudinal study (Schaie, 1983). He notes on the basis of non-statistical trend analysis of mean score standard deviations for the cross-sectional study, that there is a slight trend for increased variability at around 60 on Verbal Meaning (a recognition task) and for Immediate Recall (of a 20 item list) around 88. For Spatial Orientation (a rotation of objects task), Inductive Reasoning (a letter series completion task), Number (a simple addition task) and Word Fluency (in written letter category mode), variability remains relatively stable until there is a sharp decline for Spatial Orientation, Inductive Reasoning and Number from the 70's, and a sharp decline for Verbal Fluency from the late 80's. On Real Life Tasks (an educational test series mixture of interpretive and information selection verbal and visual tasks) there was an exponential increase in variability from the late 60's.

[9] Schaie (1983) has listed a measure as 'Educational Aptitude' for the longitudinal data set, which on the cross-sectional data set appears to be the same measure as that termed 'Real Life Tasks' in that, according to Schaie's detailed measurement descriptions, the 'Real Life Task' measure is made up from an educational aptitude battery. Hence for the purposes of this discussion the measure is termed Educational/Real Life Tasks with reference to both the cross-sectional and longitudinal data sets.

On a similar non-statistical trend analysis of standard deviations of difference scores from the longitudinal study, Schaie (1983) notes that the predominant pattern is for variability to be slightly inflated in the age range 25-32 over that noted between the ages of 32 to 60 where it remains relatively small. After the age 60 there is an increase in variability, and then from the 70's onwards with the exception of Verbal Meaning and Educational Aptitude Tasks which remain relatively high, there is a subsequent decline in variability, which is most marked for the Space modality. Schaie does not appear himself to have made any comparisons between his cross-sectional or longitudinal variability data, but on perusal of the two data sets it is clear that the dip in variability noted in the middle adult years from the 30's until the subsequent trend for enhanced variability in the 60's present for the longitudinal data set, is not clearly evident on the cross-sectional data. It is possible that this early reduction in variability for the longitudinal data set may be due to the confounding effect of test familiarity over the *initial* repeat-test instances in the longitudinal study. This was a trend which was identified as a hypothetical probability from the laboratory data of Anastasi (see discussion, section 2.2.1.2, note 5) for which on initial trials there was reduced variability, followed by an increase in variability in subsequent trials due to repeated practice.

However, for the years following 60 it is apparent from Schaie's two data sets, taking into consideration that they are presented in different statistical terms (SD's from mean scores, and SD's from difference scores, respectively), that the domain-specific variability trends across the cross-sectional and longitudinal approaches are in broad terms *virtually identical*. (Subtle indications of slightly more pronounced variability effects after 60 for the longitudinal data compared with the cross-sectional data may not be significant; on the other hand the trend is consistent with Anastasi's observation of increased variability following many trials of repeated testing - see section 2.2.1.2, note 5). Nevertheless, the overall consistency of variability effects from both research modalities provides extraordinarily compelling evidence that the alterations in variability approximate the 'true' nature of inter-individual variability patterns in association with aging. In both research instances the indication is that for Verbal Meaning in earlier old age there is a slight increase in variability around 60 which remains high through older old age, implying a spreading out of the gap between the most and least able performers which remains stable in later years (differentiation effect). For the other modalities there is in the main an initial tendency for increases in variability from the 60's (differentiation effect), followed by a marked decrease in variability in the later older years (dedifferentiation effect). These domain-by-age interaction effects on variability from Schaie's longitudinal data are represented graphically in Willis (1985) and are replicated in Figure 4.3 below, with a specific focus on the later adult age groups. Willis describes the shifts in inter-individual variability as 'notable', and suggests that they have important implications for the design of learning programs in older adults.

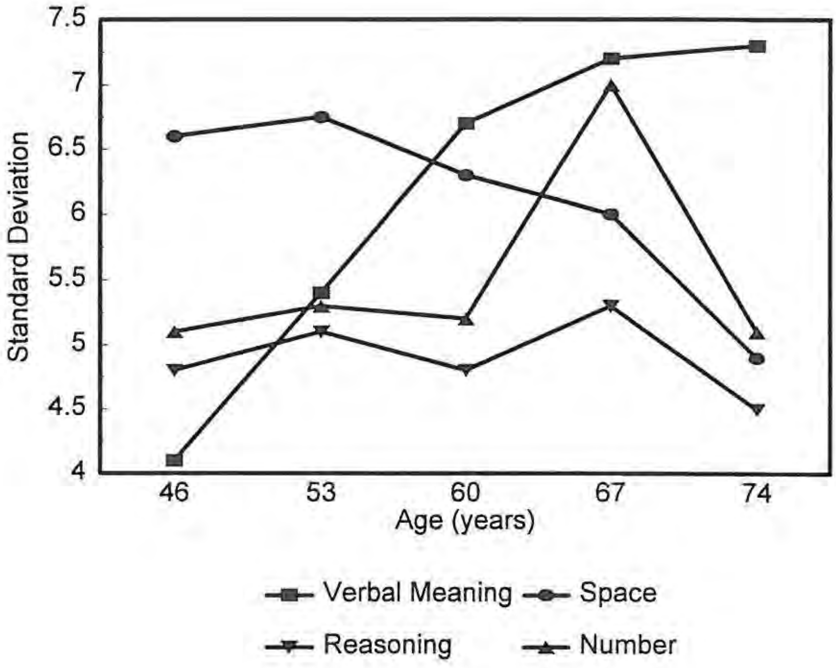


Figure 4.3. Variance across the older adult age range from Schaie's (1983) longitudinal 7-year change intervals (after Willis, 1985).

On detailed scrutiny of Schaie's particular measurement variables in terms of the differential age-sensitivity of tasks as demonstrated on the central trends analysis in Chapter 3 (section 3.2.3.2), the following is apparent. The three tasks of Space, Reasoning and Number, which reveal the initial age-related differentiation and subsequent older-age dedifferentiation effect from the 70's, all have an age-sensitive (as demonstrated on the central trends analysis in Chapter 3, section 3.2.3.2), speeded spatial or symbolic component, or a combination of these. The Word Fluency task, which also shows the dedifferentiation effect in older age, has similarly been isolated as one of the more age-sensitive tasks on the central trends analysis. On this basis these four tests (Space, Reasoning, Number and Verbal Fluency) can be functionally dissociated from the Verbal Meaning task (which is a less age-sensitive, more crystallized and educationally-related verbal recognition task), and from the Educational/Real Life Task measure (which is a mixed modality task, incorporating, as in an IQ test, both fluid and crystallized functions), both of which do not show the dedifferentiation effect in older-age, but rather a sustained differentiation effect.

The above non-linear initial differentiation and subsequent dedifferentiation effect can be explained in terms of the Satz BRC theory in the following way. It implies an initial spreading out of the most and least able in the early phases of the aging process as elements of the neural aging 'symptom presentation' starts to take effect, and individual differences in levels of brain reserve capacity produce differential protective effects (differentiation effect). This is followed by a dramatic closing of the abilities gap in the later stages of the aging process (dedifferentiation effect), when the neural aging process has advanced to such an extent that a significant proportion of individuals are no longer protected from its effects. As to be expected from the Satz BRC challenge hypothesis, the effect is task specific in that the tests representing the more age-sensitive, less crystallized, and less educationally dependent functional domains appear to be most susceptible to the subsequent onset of the more severe, dedifferentiation aging effect.

Thus, in sum from Schaie's data, not only is there a high level of consistency between the cross-sectional and longitudinal findings, but also there appears to be an overall increase in variability in the initial older adult years, which is not fundamentally contradicted by the fact that there is a subsequent dedifferentiation effect for some tasks in later years); there is a comprehensible functional consistency in association with the variability patterns that have been demonstrated across the data; and all of these observations are consistent with the hypotheses extrapolated from the Satz BRC theory, as delineated in Chapter 3.

Data from Salthouse. Salthouse (1991) presents a discussion as to whether or not the steady decline in average cognitive performance with aging is an artifact of a dramatic decline for a small segment of the population, with most, or even some individuals maintaining their original levels of performance. He goes on to reject this possibility claiming that the increases in inter-individual variance to be expected if this were true are not in evidence. First and foremost it is important to be clear that where Salthouse wants to make the point that the *majority* of older individuals do not sustain their previous cognitive ability with aging, this does *not* stand in fundamental contradiction to the position of Rabbitt and Schaie. Rabbitt and Schaie are making a very different point, which is that there are a proportion of individuals who sustain a disproportionately high level of cognitive performance across the adult years, and that the performance of these individuals is not well reflected by the central tendency data. Where these parties *do*, however, stand in opposition to each other, is where Salthouse intimates that there are not even *some* older persons who will show relative stability with age, and in that he supports his overall position with the claim that there is no evidence for increased variability in cognitive performance with aging. On the other hand, as already noted, Rabbitt and Schaie support their position with the contrary observation of increased variability in association with aging.

In order to investigate the question of whether the average cognitive decline which has been so robustly demonstrated in association with age is an artifact of vastly disproportionate decline between individuals, Salthouse (1991) examines the issue of increased variability in association with cognitive aging. He presents data for *variances by chronological age* in graphical form for (i) the eight Army Alpha subtests from the study of Jones and Conrad (1933), (ii) the eleven subtests of the WAIS and WAIS-R (Wechsler, 1955, and Wechsler, 1981, respectively), and (iii) for the five Prime Mental Abilities (PMA) of Verbal Meaning, Space, Reasoning, Number and Word Fluency, reported from Schaie (1985). The latter are the cross-sectional data from Schaie, that have already been discussed above. Salthouse's discussion of this wealth of material is extremely undifferentiated taking up the space of one short paragraph in which he states as follows:

> "The data .....indicate that the mean level of performance can decrease without concomitant increases in variance. This pattern is evident in *several* of the Wechsler Performance tests, and in *most* of the Army Alpha and PMA tests. An implication of these data is that it is apparently not the case that some people remain stable while others decline markedly, because that would result in increases in variance that are not observed." (Salthouse, 1991; p. 68, emphasis added).

It is apparent even from his own language here (as emphasized in the above text), that Salthouse has glossed over this material, and that there probably were some indications of increased variance that he noticed but chose not to focus on for the sake of his overall argument. Whereas the fact is, that *most* of the WAIS and WAIS-R subtests as depicted in the Salthouse graphs show a marked overall trend of increases in variability between younger and older years, and most also show an initial differentiation and subsequent dedifferentiation effect as follows: An overall differentiation effect is apparent for Object Assembly; an initial differentiation and subsequent dedifferentiation effect is observed on Vocabulary, Similarities, Information, Comprehension, Picture Arrangement, and Digit Symbol. The pattern of variability changes tends to be the same across subtests for both WAIS and WAIS-R versions of the test, except that *in each instance* the differentiation and subsequent dedifferentiation effect occurs *earlier* for the WAIS subtests, which indicates a cohort effect. For Picture Completion there is no variability effect noticeable for the WAIS, whereas there is an exponential increase in variability on the WAIS-R after the age of 40; only for Digit Span, Block design and Arithmetic do there appear to be no clear variability effects. For illustrative purposes, data have been extracted from the Salthouse WAIS/WAIS-R graphs and appear in Figure 4.4 (p.100). Importantly, these graphs highlight the consistent *earlier dedifferentiation effect in each instance for the WAIS, compared with the WAIS-R data.*

The material in the form in which Salthouse has presented it in his own text, certainly does appear at first glance to be very patternless, especially in that multiple functional modalities across all three sets of data, and different studies in the case of the Wechsler IQ tests, are all depicted together in the same graphs. However, in terms of the hypothetical and research spectacles provided thus far in this thesis, differential patterns between functional modalities *are to be expected* and will *necessarily* appear contradictory and patternless until modality effects are teased out. Further, the results on the same test from different cohorts such as the Wechsler 1955 (WAIS), and 1981 (WAIS-R) studies, may reflect lower versus higher educational/IQ[10] levels, and lower versus higher levels of test sophistication, respectively (see discussions in Lezak, 1995; Shuttleworth-Jordan, 1995). Thus in terms of the hypothesis of expected differential effects for variability due to lower versus higher educational/IQ levels as extrapolated from the Satz BRC theory, differences in variability patterns between these two studies can also be expected. Conceptually the *earlier* onset of increased variability in the cohort of arguably lower education/IQ (that is the WAIS), is *entirely* commensurate with what could be expected in terms of BRC theory. In such a cohort, there would be relatively reduced brain reserve capacity, and hence earlier symptom presentation than in the cohort of higher education/IQ.

---

[10] For the purposes of general discussion in this thesis, education and IQ are grouped together when describing the relative characteristics of samples. This is *not* on the assumption that they are entirely equivalent, but rather because educational and IQ levels have been shown to be *broadly* comparable in the populations of normal aging studies (for example, Cornfield & Shuttleworth-Jordan, 1996; Ivnik et al., 1992). Further, Satz (1993) has argued that educational level and IQ level are roughly equivocal measures of adaptive functioning which are equated with higher levels of BRC (see section 3.2.3.3).

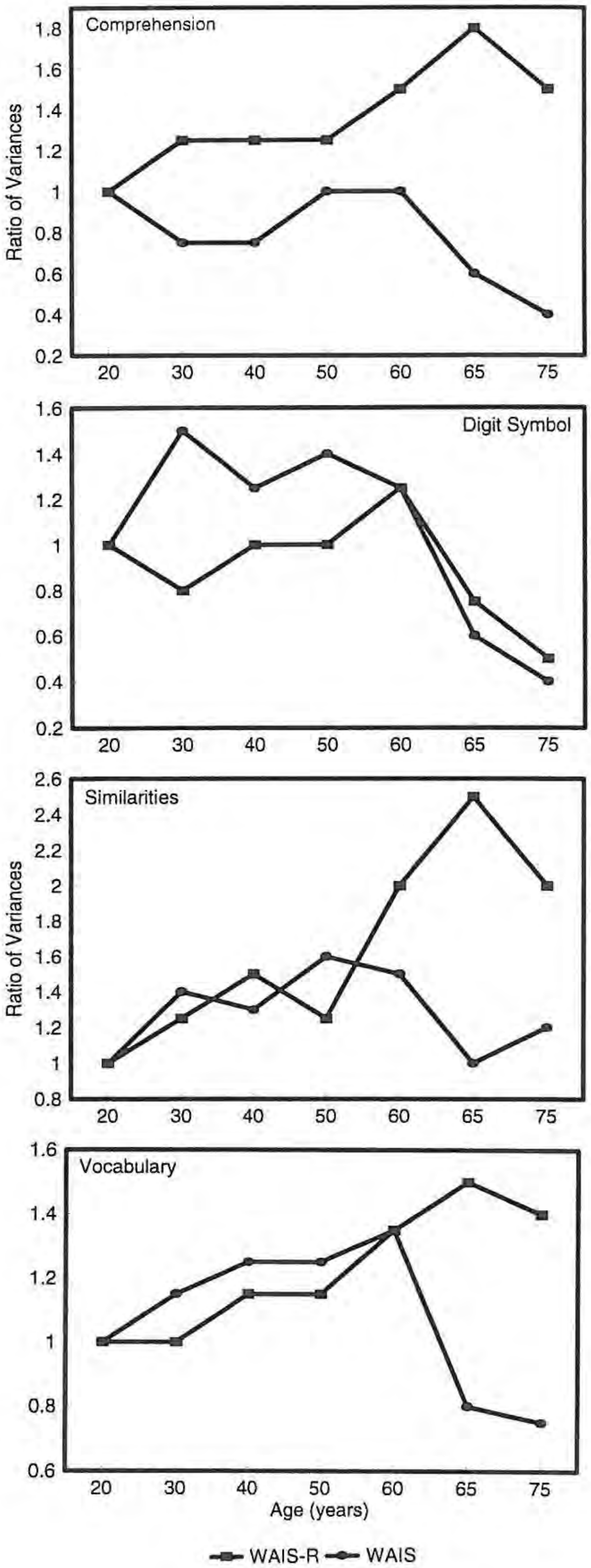


Figure 4.4. Ratio of variances for the WAIS and WAIS-R subtests of Comprehension, Digit Span, Similarities and Vocabulary (after Salthouse, 1991), depicting in each instance the earlier peak, and earlier older-age reduction in variances for the WAIS data.

Further still, the data will appear patternless if the expectation is for a *linear* increase across the adult years, and it appears that in his analysis Salthouse has ignored the *non-linear* patterns across the full spectrum of the adult years which appear for many of the tasks on his graphs. His observations seem to be based on bimodal comparisons between the very early and very late older adult years which do in fact reveal only a few indications of increased variability in this comparative sense, due to the late decreases in variability which occur in many instances on the test material he presents. However, such an undifferentiated comparison which ignores the entire curve of the graphs, presents a distorted view of the situation in that it fails to depict the full picture of a commonly occurring *initial* increase in variability in older age, and subsequent older-age decrease, which has already been identified as a characteristic from the Schaie data as discussed above (see Figure 4.3, p.96).

Finally, it is important to be aware that patterns that appear unusual and are not conceptually coherent may be artifactual, and such explanations should be considered before coming to the conclusion that there is no particular pattern to variability in association with the aging process. For example, the bizarre-appearing exponential increase on Picture Completion after 40 years which is apparent for the WAIS-R but not for the WAIS in the data reported by Salthouse, could possibly be explained as the consequence of an artifactual ceiling effect which has occurred for the younger age groups in the presumably more highly educated, and test-sophisticated WAIS-R cohort. Further, results in the graphical material presented by Salthouse on the Army Alpha subtests tend to suggest an artifactual phenomenon. On the basis of the hypothesis of *differential* variability effects which are expected across functional modalities, when the *identical* trend occurs across all test modalities within a study, the indication is that the data are reflecting the artifact of a powerful sample effect. For the Army Alpha data (reported in Salthouse), regardless of the functional modality, there is a highly consistent reduction in variability for the 50 year old age group for every one of the eight subtests, and no consistent trend occurs in this way across all the Army Alpha subtests for any other age group. This hypothesized artifactual occurrence is illustrated in Figure 4.5 upper graph (p.102) in which the pattern of three of the eight subtests (Antonym Vocabulary, Arithmetic, and Disarranged Set), have been extracted from Salthouse (1991) to demonstrate the point. If the 50's age group is eliminated from the Army Alpha graph as in Figure 4.5 lower graph, a rather bizarre and inexplicable *double* aging differentiation/dedifferentiation effect disappears to reveal a cluster of effects with comparability to those from Schaie (see

Figure 4.3, p.96): a *single* differentiation/ dedifferentiation effect in the later adult years for two tasks (Disarranged Set and Arithmetic), and the contrasting absence of any dedifferentiation effect for the relatively more crystallized task (Antonym Vocabulary). A sample effect of this type could result from such phenomena as a vastly disproportionately high or low number of subjects, or unequally high or low educational/IQ level in the 50 year old age group, relative to the other age groups. However it would be necessary to go back to the original data in order to unravel this effect clearly.

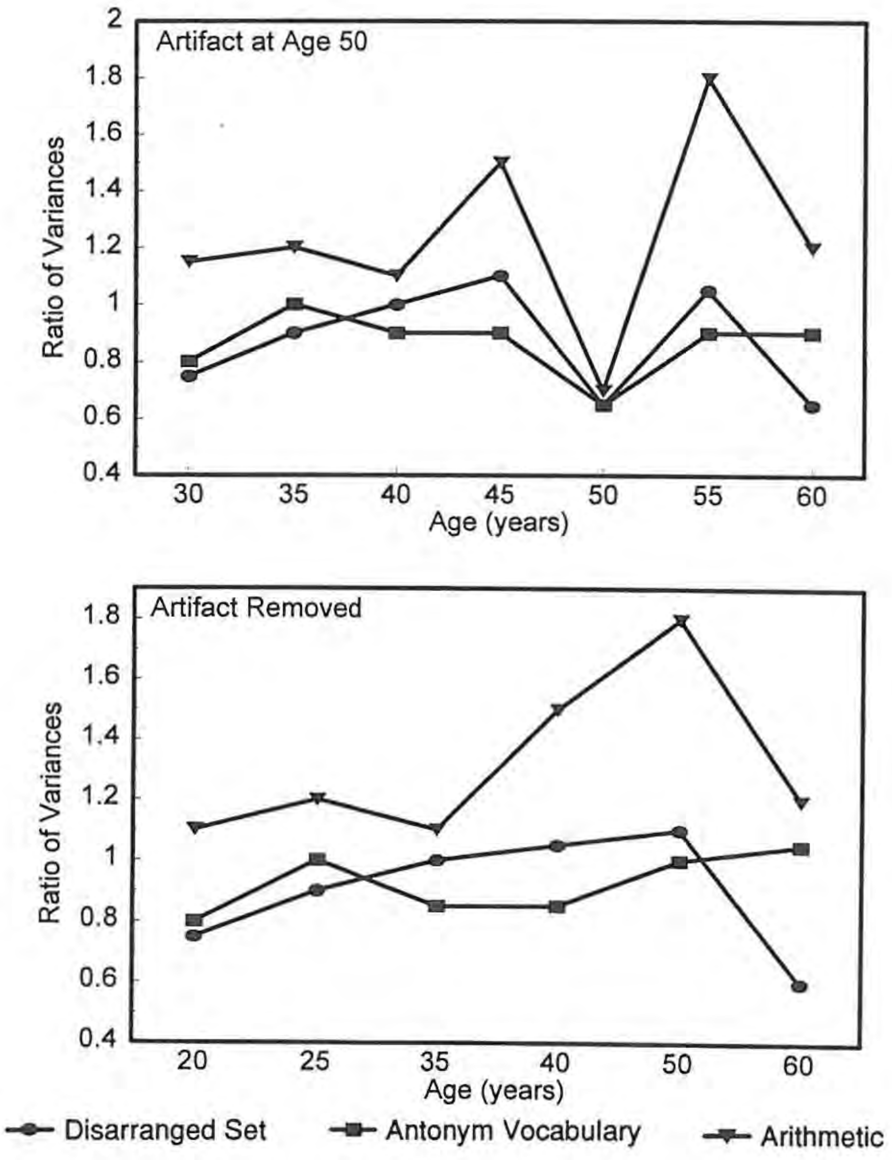


Figure 4.5. Schematic representation of an hypothesised artifactual decrease in variance at the 50's age stage for the three Army Alpha subtests of Antonym Vocabulary, Disarranged Set and Arithmetic. The upper graph (after Salthouse, 1991) depicts the presence of the artifact; the lower graph depicts the variance pattern with the proposed artifact removed.

Thus on the re-examination of the data presented by Salthouse through the theoretical prism of neural and psychometric perspectives on aging, and the Satz BRC theory, albeit not on the basis of statistical analyses or access to the original data, there do appear to be marked *conceptually* coherent trends (which are commensurate also with the indications from Schaie's research) of increased variability as age advances, and of a subsequent domain-specific dedifferentiation effect in the older-old age groups. Furthermore there appears to be support for the hypothetical indication from the Satz BRC theory of differential variability effects due to low versus higher levels of education and IQ. From the data reported in Salthouse, it appears that this difference in variability effects consistently takes the form of an *earlier* onset of age-related variability increases in association with a cohort of *lower* educational/IQ level, which is conceptually coherent with the basic tenets of the Satz BRC theory.

Individual variability data. In more intricate descriptive studies on variability in older aging, only a very limited number of research endeavours have set out to examine the actual distribution patterns of individual scores with respect to elements of cognitive performance (Maylor & Rabbitt, 1994; Rabbitt, 1993c). Maylor and Rabbitt use a comparison of the distributions of individual variation across letter coding and visual search to demonstrate that there is task specific effects of aging across 60- and 75-year-old average age groups, which would not have been revealed by an analysis using only group means. From this research it appeared that letter coding was more affected by age than visual search, and these authors suggest that this might be explained in that the former requires greater involvement of working memory.

Rabbitt (1993b; 1993c), further reports on an aspect of a large cross-sectional analysis involving around 300 subjects in which composite memory test scores were plotted against IQ scores across the three age decades of 50-, 60-, and 70-year-olds. It was found that the numbers of subjects that fell beyond 2 SD below expectation from their current IQ test scores steadily increased with each of the 50- to 60- to 70-year-old age decades from 3, to 7, to 19, respectively, indicating that within a large population of healthy older adults the probability of a large dissociation between memory and IQ increases with age. Rabbitt uses this dissociative finding on outliers, not as an item of 'error variance' to be discarded, but as an observation with prominent scientific significance which supports his premise from the network (as against single factor) model of aging, that speed of information processing does not serve as the exclusive underlying cause of all decrements due to cognitive aging.

General variability trends: conclusions. Taken together, the above analyses of overlapping proportions of individuals between older and younger age groups, and the scrutiny of variability data from the large-scale aging studies of Schaie and Rabbitt and associates, have led these leading aging researchers to claim in summary that many older persons have cognitive abilities which are well within the range of younger groups, and that in some cases older people may sustain their previous level of intellectual functioning into late old age. Thus they note that this is expressed in great variability in age-related cognitive decline such that some people show little or no impairment, whereas the performance of others covers a spectrum from mild to severe decrement. Although neither Schaie nor Rabbitt have used statistical analyses to verify the broad patterns of alterations in variability between age groups, their observations gain considerable conceptual validity when considered in conjunction with their demonstration of overlapping trends between younger and older groups, and the isolation of a proportion of individuals who appear to remain cognitively intact relative to 'pre-aging' estimates of functioning even in the oldest-old age group.

Whilst Salthouse (1991) claims that existing data do not support increases in inter-individual variance with increased age, a more differentiated analysis of the large-scale data bases he reviews does not support this observation. Rather it provides significant support for the view of Rabbitt and Schaie and the variability trends which have been revealed on the basis of their (Rabbitt and Schaie's) data. It even serves to reveal new conceptually coherent variability trends with respect to the effects of higher functioning samples. Thus overall, the review thus far provides consistent indications that variability changes across age groups are not necessarily linear, and that the patterns of variability changes in aging are intricately domain-specific. It further produces the indication that low education and low IQ may result in the earlier onset of age-related variability effects regardless of functional modality.

4.2.1.2 Specific variability trends

The research review from this perspective will be presented first in terms of what has served as an area of high focus in the area - and that is fluid versus crystallized functions. Second variability data from IQ tests, and on tests which can be categorized according to the functional modalities delineated earlier under normative trends, will be discussed.

Crystallized/Fluid functions. Four studies comparing variability changes across adult age groups for crystallized and fluid functions appear in the literature, three of which (Christensen et al., 1994; Hayslip & Sterns, 1979; Morse, 1993), indicate that *variability increases with age on fluid abilities but remains stable on crystallized functions*. Christensen et al.'s study examined the older-old age decades from 70 through to 90 cross-sectionally on a healthy, community-based Australian population; the other two studies investigated the differences on the bimodal young versus old dimension also from a cross-sectional perspective. Christensen et al.'s finding of greater variability in fluid as against crystallized functions was in association with declining mean scores in both instances; Hayslip and Sterns' similar finding was in association with declining means with age only on fluid functions; the study of Morse, which is by far the most comprehensive variability study to appear in the literature to date, and comprises a meta-analysis of functional trends on multiple studies from two leading aging journals, does not include associated mean trends with the variability data. Taken together, these three studies which have all derived their findings on the basis of statistical analyses, do appear to provide a significant measure of cross-validation for the finding of age-associated increases in variance with fluid abilities across the broad adult age span, in the absence of increased variability for crystallized abilities.

The fourth study (McArdle, 1984), is an unpublished conference presentation reported in Horn and Hofer (1992) which provides compelling graphical representation (although not supported with formal statistical analysis) for the apparently complete *opposite* finding to the above three studies of entirely *stable* variability on fluid functions in association with a continuous average decline in this ability between the ages of 40 and 80, whereas across the same period crystallized functions show a relatively small average decline starting after age 50 in association with a marked *increase* in variability. It is of dubious validity to attempt an explanation of such a startling apparently contradictory finding on the basis of a second-hand unpublished report. For instance, differences might become clear on the basis of how the fluid and crystallized functions were operationally defined, in that, for instance, a direct translation of Performance and Verbal IQ into fluid and crystallized functions, respectively, would be erroneous and might go some way to explaining such an apparent transposition of effects (see discussion to follow below in the next subsection on variability effects for IQ). However, on the presumption that no such conceptual misrepresentation is the case, it may be that an explanation for McArdle's totally opposite finding lies in the area of sampling differences as follows.

It is reported that McArdle's study was based on a sample drawn to match census data with respect to demographic factors such as age and sex, which is likely to be much more generally representative of the full spectrum of elderly persons. Such a sample (see discussion on generally representative elderly populations, section 2.2.1.3) would incorporate a much wider range of socio-economic, educational and intellectual levels than the other reported studies that have investigated the fluid/crystallized dimension for variability trends. Moreover a census sample, such as that of McArdle's, might not have been scrupulously screened to exclude non-hospitalized individuals with incipient disease, including brain disease, the presence of which would in anycase be more prevalent in lower socio-economic populations. In contrast to McArdle's study, the other studies in question were all based on the healthy, non-diseased population classically used in aging research. Morse in fact specifically warns that none of the samples used in her meta-analysis were randomly selected, all being derived from the healthy spectrum of the elderly. Thus it is possible to understand the phenomenon of these opposing trends between that of the first three studies on relatively high functioning samples, and that of McArdle on a postulated lower functioning sample, in terms of the Satz BRC theory in the following way (see Figure 4.6, p.107).

In a generally representative population (compared with a restricted high functioning population) the risk factors of a lower levels of education and IQ, and incipient disease in association with lower socio-economic status, would aggregate and serve to reduce brain reserve capacity. This would result in the failure of a significant proportion of individuals to sustain the hold-pattern on crystallized abilities as occurs for most higher functioning, healthy individuals, and which is reflected in a stable pattern of variability in crystallized functions across the younger to older adult years. Consequently, a differentiation effect in the lower functioning group is likely to occur in association with aging, and be reflected in the opposite trend of an *increase* in variability over time for crystallized abilities. Further due to reduced brain reserve capacity in the lower functioning population, an exceptionally early differentiation/dedifferentiation effect may occur on fluid functions which are particularly vulnerable to aging effects, and in the study window of McArdle's study this

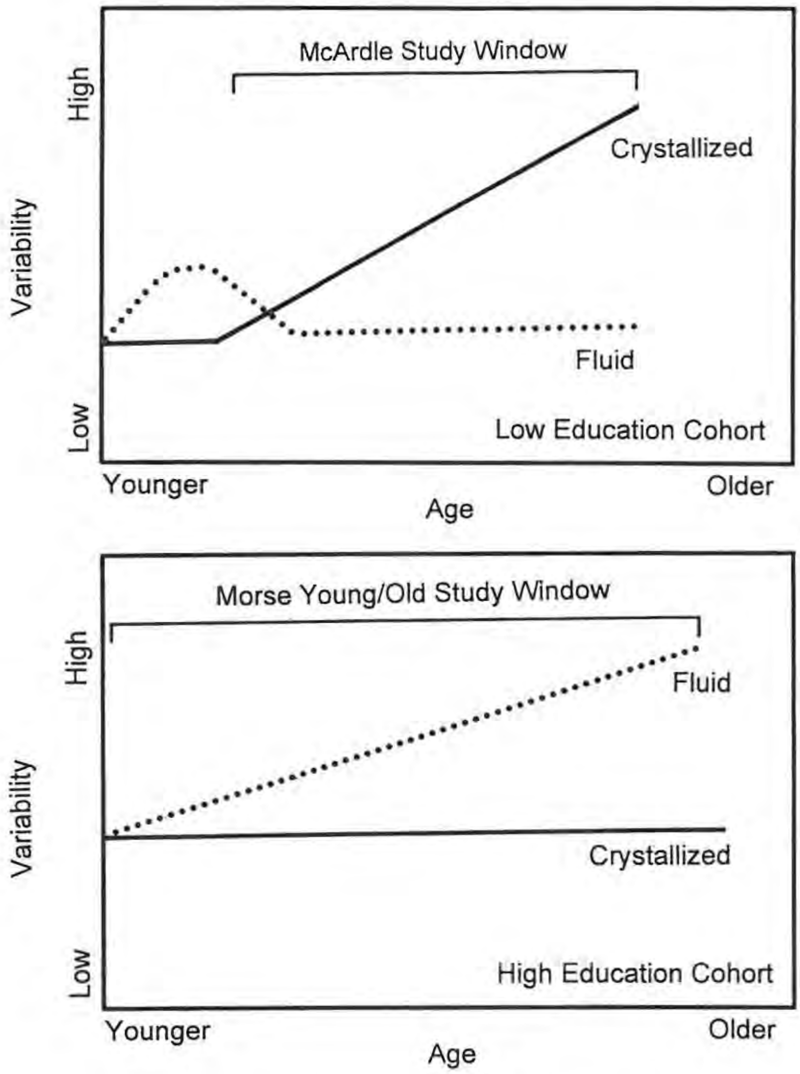


Figure 4.6. Schematic representation of an hypothesised apparently paradoxical effect consequent on the use of a low functioning cohort (upper graph), and a high functioning cohort (lower graph).

would be reflected in an opposite trend of no apparent increase in variability for fluid functions in the lower functioning group. The latter explanation is consistent with the very early differentiation effect noted for the WAIS Digit Symbol subtest (see Figure 4.4 above, p.100). However, an equally plausible alternative explanation for the absence of increased variability for fluid functions on McArdle's study, needs to be considered. The aggregated effect of a group with low education/IQ in conjunction with the challenge of highly age-sensitive tasks, may reduce brain reserve capacity to such an extent, that there will be almost no differentiation effect at all on the fluid tasks due to the fact that an insidious decline with age occurs across all individuals across the entire spectrum of the adult years. Such an occurrence would also be reflected as an absence of increased variability on fluid functions as was apparent in McArdle's research.

Verbal/Performance IQ. In a related area to that of fluid versus crystallized abilities, are the indications of variability trends across older adult groups for Verbal and Performance IQ which have appeared it seems in only two studies. Maddox and Douglass (1974), in a longitudinal study on a generally representative population of North Carolina (age range 60-94), report no significant changes in variability between the first and sixth testing trials over a period of 13 years, although the trends are in the direction of an increase in variance through time, and are slightly more pronounced on the verbal as against the performance scale. Riegel et al. (1967) present cross-sectional variability data on a generally representative population from North Germany, which on informal scrutiny reveal equal and stable variability trends for Verbal and Performance IQ from between the ages of 50 and 75, after which there is a noticeable trend towards increased variability which again is slightly more pronounced on the verbal scale. Thus in broad terms this cross-sectional study of Riegel et al., and the longitudinal study of Maddox and Douglass, are confirmatory of each other's variability trends.

At first glance these findings with respect to Verbal and Performance IQ might appear to be yet another set of contradictions in terms of crystallized versus fluid patterns, respectively, with respect to variability patterns. However it is important to be clear that these two pairs of modalities are *not* strictly transposable and hence have been separately demarcated for this discussion. For example, subtests such as Arithmetic Reasoning and Digit Span which are contained in the WAIS Verbal IQ measure represent fluid abilities, and tests such as Picture Completion and Picture Arrangement, which are contained in the WAIS Performance IQ construct, incorporate non-fluid, acquired cultural components. On examination it emerges that the fluid component in the Morse study, for instance, comprises the Verbal IQ Digits Forwards and Backwards subtests in addition to the Performance IQ subtests of Block Design and Digit Symbol. Thus it can be argued that the composite mixture of fluid and crystallized skills such as occurs to some extent on an IQ subtest distribution such as the WAIS, might obscure the generally more crystallized VIQ and the generally more fluid PIQ effects for variability data, even though these relative weightings are not obscured in central tendencies data. Thus for central trends analyses these terms can be roughly equated, whereas for *variability* data they can not. Whilst composite VIQ and PIQ tests appear to elicit equivalent variability effects, distinctive opposite variability effects occur for the more pure compositions of crystallized and fluid functions.

Generally with respect to tests of intelligence, the consistent trend of increasing variability for IQ in both the Riegel et al. and Maddox and Douglass studies across the older adult age groups, is broadly commensurate with Rabbitt's (1993b) graphical representation of steadily increasing variability across the 50's, 60's and 70's age groups on the AH 4 (1) IQ test; Schaie's longitudinal demonstration of increased variability in the above 60's age groups for a task labelled 'Intellectual Ability'; and Schaie's parallel cross-sectional and longitudinal demonstration of marked increased variability in the older age groups on an Educational/Real Life Task measure which represents a mixture of visual and verbal tasks. Thus, taken together the results from all of these studies, there is very strong support for the fact that in older age groups there is enhanced variability for tests of composite cognitive abilities such as the classical IQ test in the older age groups, and this implies increasing differentiation between the most and least able individuals in general intellectual functioning across advancing age groups. In terms of the consistent observation described above (derived from published research on healthy older populations), of increased-fluid versus stable-crystallized function effects on variability, this finding suggests that the speeded fluid function is predominant in general measures of intelligence. This is a hypothesis which is also entirely commensurate with the well-documented Salthouse effect which emphasizes the powerful contribution of slowed information processing in association with intellectual functioning in general (Rabbitt, 1993b; 1993c; Salthouse, 1985a; 1985b).

It is of relevance to note that neither Riegel et al., nor Maddox and Douglass report the mean scores across groups in conjunction with their variance figures; and as mentioned in the subsection above, nor did Schaie (1983) report means in association with his longitudinal variability data. Thus it appears that all three of these groups of researchers have fallen into the trap of an over-enthusiastic differential focus as an antidote to the traditional over-exclusive normativity focus. Maddox and Douglass concede this as a limitation of their study which has focused purely on variance data, and reflect that it is very necessary to include an analysis of means in addition to variance in aging studies, in order to provide a valid depiction of the process.

Attention and Memory.[11] In the area of *attention* the only findings available appear to be those of Mitrushina and Satz (1991a). These researchers report a lower stability coefficient on the general construct of speed of mental processing and attention in their old-old group, and suggest that this reflects that some individuals over the age of 70 show a more remarkable decline in attentional capacity than others. In the area of *memory*, Morse (1993), without differentiating between any separate components of memory, reports significant increases in variability between her bimodal distribution of young and old groups. Christensen et al., (1994) on his regression analysis of residuals across the older-old age range of 70-94, also reports a strong age-associated effect on a composite memory score which is stronger than that of both fluid and crystallized intelligence. Christensen et al., poses that this may be in part due to ceiling effects in the younger subjects, but is unlikely to be the total explanation, and he considers that the more likely cause is that memory is more sensitive than other functions to change at older ages. Christensen et al.'s evidence for a dissociation between memory and IQ functioning is consistent with Rabbitt's (1993c) similar finding derived on the basis of regression analyses across the older age decades of 50, 60 and 70 (described above), in which Rabbitt demonstrated increasing proportions of individuals with this IQ/memory dissociation with advancing age.

With respect to the verbal memory modality, Christensen et al. demonstrates a significant variability effect in association with age for word recall, address recall, information, and word recognition between the ages of 70 to 94; Schaie (1988b)'s cross-sectional analysis reveals a slight trend for increased variability on an immediate list recall task; Rabbitt (1993b) presents a graphical representation of increasing distributions of scores for a list learning task across the age decades of 50, 60 and 70, in particular noting that these effects cannot be explained away on the basis of either floor or ceiling effects. For visual memory, Christensen et al. similarly report significant increasing variability in association with age for visual reproduction and face recognition tasks.

Taken together all of these findings on memory consistently suggest relatively global increases in variability in association with older age. However, this must be cautiously stated

[11] Poon (1985), has addressed the issue of variability in memory at length in a chapter dedicated to the topic. He focuses on the way in which age-changes vary across memory components (as reported on in section 3.2.3.2) as well as the origins of inter-individual variability (which will be discussed in Chapter 4, section 4.2.2.5). He does not, however, provide any descriptive data on the patterns of variability changes due to age in themselves within memory components.

in that the Christensen et al. study, which is restricted to the older-old age group, is the only one which differentiates between several components of both verbal and visual memory.

Language skills. Within the domain of language, Rabbitt (1986) indicated identical distributions for individuals in increasing decades from the 50's through to the 70's on the Mill Hill Vocabulary test, and commensurate with this Christensen et al., (1994) found no significant age-related increases in variability for his older-old 70 to 94 cohort on the Vocabulary subtest. In seeming contrast Schaie noted a trend for increasing variability for Verbal Meaning which is quite marked on his longitudinal data, and still present but less marked on the cross-sectional data. On the WAIS and WAIS-R Vocabulary data reported in Salthouse (1991), there is an increase in variability with age around 40 and 50 years subsequently levels off, and then moves into a very late dedifferentiation phase which occurs earlier for the postulated less educationally sophisticated WAIS cohort (see Figure 4.4, p.100).

Whilst these variability effects for the general area of vocabulary and verbal meaning are not all on the identical tests, the differential findings probably reflect contrasting windows in terms of the age of the cohorts of the various studies rather than differential test effects. This postulated study-window occurrence is demonstrated in Figure 4.7, upper graph (p.112), in which it is suggested that these studies cumulatively reveal that on a vocabulary task for a relatively highly educated healthy population, there is an increase in variability with age around 40 and 50 years, which subsequently levels off in earlier old age, and then moves into a very late old-age dedifferentiation phase. Hence the effect of increased variability is not detected in the studies of Rabbitt and Christensen et al. which are restricted to relatively narrow age ranges within the higher age groups of 50 to 70, and 70 to 94, respectively. In the case of Christensen this could be due to tapping into a levelled-off effect either at the top of the differentiation curve which would be more likely in a very high functioning sample (Figure 4.7, upper graph); however it could be reflecting the bottom of the dedifferentiation curve, an effect which would be more likely to occur in a less high functioning sample (as illustrated in Figure 4.7, lower graph).

In more differentiated fashion, Riegel et al., (1967) have examined the effect of age in five-year age periods from 55 through to 75^{+} on what he terms as levels of redundancy in language functions, with synonyms/antonyms being the most redundant as well as the most

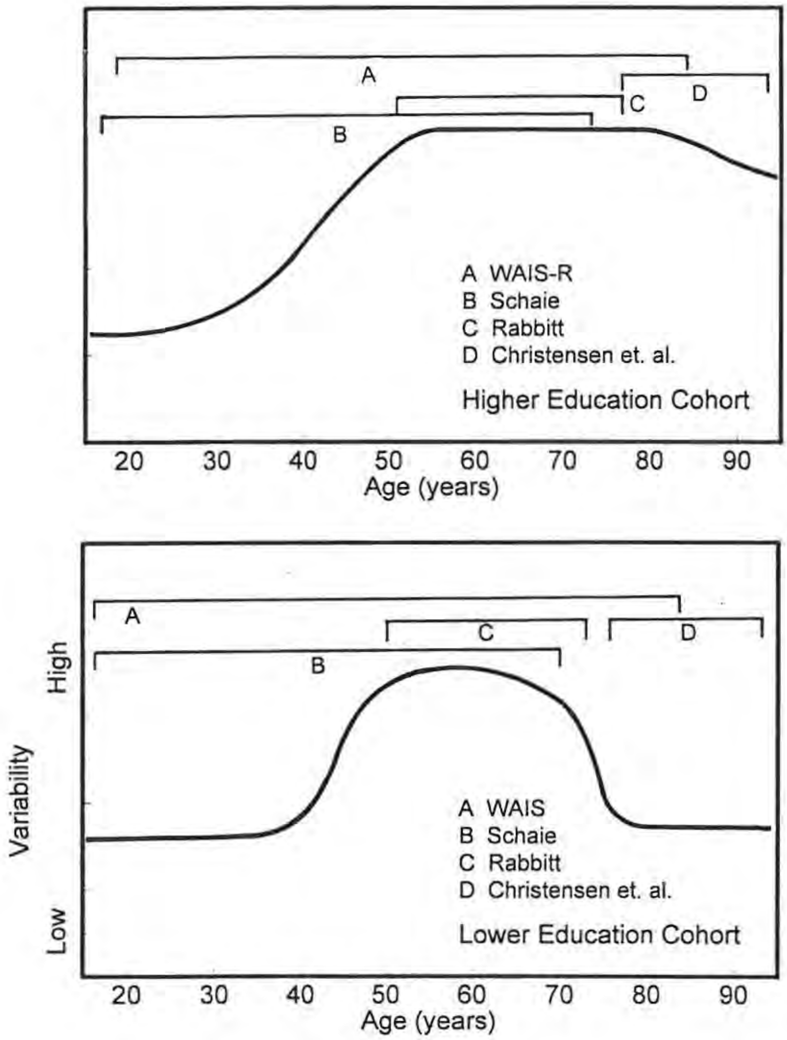


Figure 4.7. Schematic representation of postulated contrasting study window effects on Vocabulary/Verbal Meaning tasks for a cohort of individuals with relatively high education and an associated late dedifferentiation effect (upper graph), and for a cohort of individuals with relatively low education and an associated early dedifferentiation effect (lower graph). (Additionally, for cohorts of *equivalent* educational level, a late dedifferentiation effect might be revealed for a less age-sensitive task as in the upper graph; and an early dedifferentiation effect for a more age-sensitive task as in the lower graph)

crystallized function, and classifications being the least. He found that in association with older age, means declined most and variability increased most for classifications, and vice versa for synonyms/antonyms. With respect to variability, Riegel et al.'s finding of least increase in variability for the synonyms/antonyms task is clearly consistent with the lack of variability changes for crystallized ability noted in the three first studies described above (Christensen et al., 1994; Hayslip & Sterns, 1979; Morse, 1993; see section 4.2.1.2) on

Crystallized/Fluid intelligence. The direction of *increasing variability trends* from the more crystallized verbal ability of synonym/antonym production through to the classification/reasoning type ability has a high level of *inverse* consistency with the direction of *diminishing average performance trends* from crystallized down to reasoning ability as described in the cascade model (see section 3.2.3.1). In turn this effect is conceptually consistent with the increased differentiation that would be expected between the least and most able individuals across older age groups for tasks which are more sensitive to the normal aging process.

Further describing effects of language on variability, Van Gorp et al. (1986) informally note increased variability with increasing age from 59 through to 80^+ in association with a slight decline in normative trends on the Boston word naming test; Hartley (1988) documents increased individual differences in text comprehension in older than younger adults; as noted earlier the WAIS and WAIS-R data in Salthouse (1991) show the pattern of increased variability followed by decreased variability for Comprehension; and Schaie (1983; 1988b) demonstrated a trend for increasing variability on written verbal fluency in early old age followed by a sharp reduction in mid to late old age. In apparent contrast to Schaie, Christensen reports no significant effect for verbal fluency tasks. The absence of a significant increase in variability for verbal fluency on Christensen et al.'s data is difficult to explain on a functional basis since it would appear from normative data findings that this task, as with word naming, is sensitive to aging and thus should presumably show age-related increases in variability. Thus it seems highly probable that the explanation lies again in the restricted older-old range study window of his sample (70 to 94 years), and reflects earlier dramatic decline in variability for verbal fluency out of Christensen's study window, (as demonstrated on Schaie's data), which has levelled off into a stable state of dedifferentiation at the late age-stage of Christensen's sample. (Such a postulated study window effect is illustrated in Figure 4.7, lower graph, p.112).

Overall with respect to language, therefore, again a remarkably coherent picture has developed of no alterations in variability across age groups for the more crystallized language functions, whereas increases in variability do occur with the less crystallized verbal abilities which involve word definition, reasoning, and word naming components. In the case of verbal fluency and comprehension it appears that in addition to an initial increase in variability, a mid- to late-old-age dedifferentiation effect occurs similar to that demonstrated on Schaie's spatially related tasks.

Visual skills. In terms of visual skills, by means of a statistical comparison of standard deviations across three age decades from 60 through to 80 for the Trail Making Test, Shuttleworth-Jordan and her research associates confirmed significantly increased variability from 80 years on the visuo-spatial search and sequencing task of Trails A, and from 70 years for the double tracking task of Trails B in association with a highly significant decrement in average performance due to age which was more pronounced for Trails B (Cornfield & Shuttleworth-Jordan, 1996). This finding of an earlier onset of increased variability for Trails B relative to Trails A is entirely consistent with the challenge postulate extrapolated from the Satz BRC theory. It indicates that the more complex double tracking and working memory aspects of Trails B compared with Trails A together provide an enhanced level of task challenge. This results in earlier symptom onset and an associated earlier differentiation effect for Trails B, which is reflected in an earlier increase in variability for the more complex and more age-sensitive Trails B relative to Trails A.

In contrast to the overall linear increases in variability detected for the Trail Making Test for both parts A and B, in an analysis of variability across 20-39, 40-59, 60-69, 70-79 and 80-89 year-old age groups for the visuo-spatial search task of the WAIS Digit Symbol subtest, Shuttleworth-Jordan and associates found, also in association with a highly significant continuous decrement in average performance due to age, that variability increased at 60, remained stable across the 60's and 70's age groups following which there was a sharp significant reduction in variability from age 80 (Bode & Shuttleworth-Jordan, 1993). These contrasting test-by-age interaction variability effects for the Trail Making Test and Digit Symbol studies of Shuttleworth-Jordan and associates are graphically represented in Figures 4.8a and 4.8b (p.115). The *linear* pattern of results for the Trail Making Test as depicted in Figure 4.8a, is consistent with the many studies noted above, especially with respect to the more crystallized and educationally-based verbal functions and overall intellectual functions, which in linear fashion, show an inverse relationship of increased variability in relation to diminishing average scores with age (for example, Rabbitt, 1993b; Riegel et al., 1967). On the other hand, the *non-linear* initial increase and subsequent sharp decline in variability after 80 for the Digit Symbol subtest as depicted in Figure 4.8b, is *highly* consistent with the initial differentiation and subsequent older-age dedifferentiation pattern in the Digit Symbol

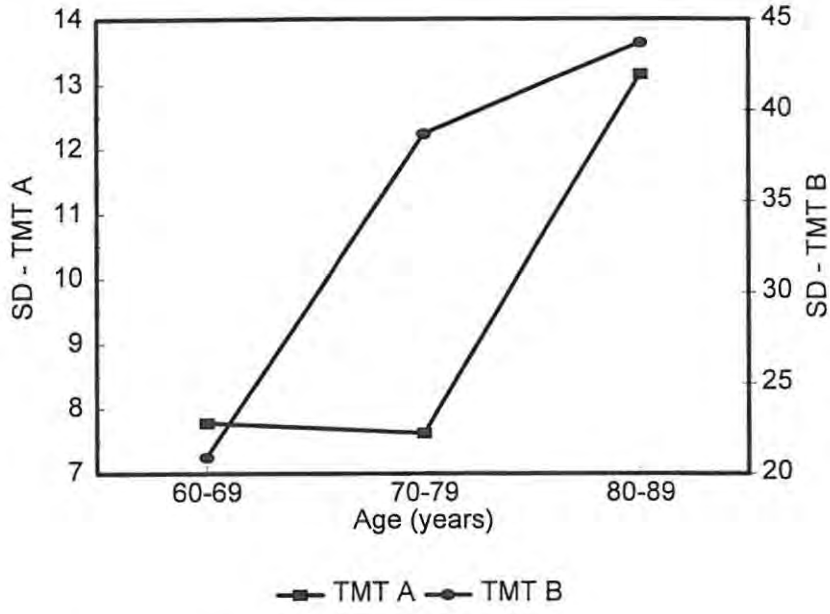


Figure 4.8a. The later differentiation effect in the 80's for the Trail Making Test Part A (TMT A), compared with the earlier differentiation effect in the 70's for Trail Making Test Part B (TMT B) (data from Cornfield and Shuttleworth-Jordan, 1996).

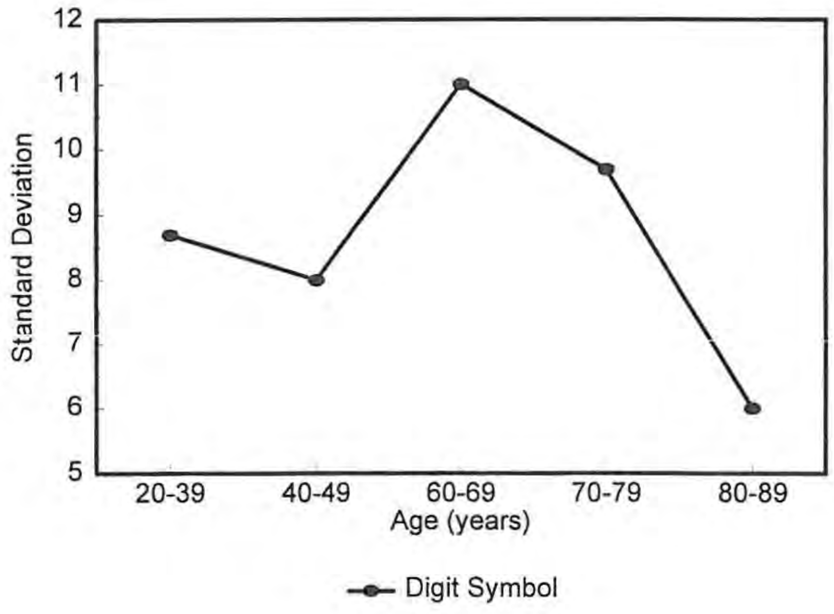


Figure 4.8b. The differentiation effect in the 60's and subsequent dedifferentiation effect in the 80's for the WAIS Digit Symbol Subtest (data from Bode and Shuttleworth-Jordan, 1993).

variability data illustrated earlier from Salthouse (1991) for both the WAIS and WAIS-R (see Figure 4.4, p.100), and commensurate with those of Schaie for his Space, Reasoning and Number abilities (see Figure 4.3, p.96), all of which have in common, visuo-spatial, numerical and time components in their execution.

The finding that dedifferentiation has occurred after 80 for Digit Symbol in the Bode and Shuttleworth-Jordan (1993) study, indicating that few people are spared an age effect on this task at this stage, presents an interesting dissociation with the Trail Making Test, for which by contrast it appears as though a significant number of 80-year olds *have* been disproportionately spared the overall age effect. The samples involved in these two Shuttleworth-Jordan and associates' studies comprised fundamentally equivalent relatively highly educated samples. Hence detailed comparisons between the studies is viable, and reasons for the dissociations in variability findings between the Trail Making Test in general compared with Digit Symbol, can be posed from the angle of the Satz BRC theory (as elaborated for the purposes of this study) as follows.

Although both the Trail Making Test and Digit Symbol tasks involve similar speeded visuo-spatial search components, and both are considered to be relatively equivalent in terms of age sensitivity (see Heaton et al., 1986), there is a clear working memory component, and probably stronger educational component for the Trail Making Test compared with Digit Symbol. This is in that the Trail Making Test requires rational ordering of the material in addition to the speeded visual search involved in both tests. Furthermore, from the perspective of Birren and Cunningham's (1985) cascade model of aging (section 3.2.3.1), which designates reasoning ability as less age-sensitive than perceptual skills, the rational ordering component in the Trail Making Test, which is absent for the Digit Symbol task, arguably provides a holding factor which renders it the less age-sensitive test especially for a highly educated cohort.[12]

---

[12] The significant effect of education and IQ on Trail Making Test performance has been reported on in Shuttleworth-Jordan and Cornfield (1996). In Heaton et al., (1986, p.114), the graphical representations on the WAIS and Halstead-Reitan battery of the effect of education highlight the extreme effect of lower versus higher levels of education on the Trail Making Test, although for comparative purposes there is no graph presented for Digit Symbol. The tables of age versus education effects (Heaton et al., 1986, pp.105-106) denote Trails B and Digit Symbol as falling in almost the identical position. Thus clearly on empirical grounds there are strong functional similarities between these two tests. However there are also clear differences between the two tasks in terms of their execution which are posed here on theoretical grounds.

Thus in terms of the hypotheses extrapolated from the Satz BRC theory, it is suggested that in this particularly highly educated Shuttleworth-Jordan and associates cohort, education can be brought to bear as a protective factor for a longer time for the Trail Making Test than for the Digit Symbol task, since it is more dependent on educationally-based skills and it is the less age-sensitive test. Consequently a proportion of very high functioning individuals will *not* succumb to the general age decline as soon for the more educationally based, less age-sensitive Trails task, and this is reflected in the *absence* of the dedifferentiation effect on this task. On the other hand, for the less educationally dependent, more age-sensitive Digit Symbol test, when the neural aging effect becomes very severe in the over 80's age group, the protective factor of education would not be of consequence in the way it would be for the Trail Making Test, and hence a significant proportion of more highly educated individuals would *not* be saved from the aging effect as for the Trail Making Test; dedifferentiation would occur and be reflected in the opposite trend to that for the Trail Making Test, of a plunging Digit Symbol variability figure in the 80's age group.

This same explanation can be extended to apply to all the four relatively pure, and as argued above, broadly visual, symbolic and spatially-related tasks of Schaie (Space, Reasoning, Number), which show the dedifferentiation effect; none of these four tasks involve the clear dual function (working memory) complexity of timed visuo-perceptual search in conjunction with rational ordering to the extent that this is called upon for the Trail Making Test, and all of these tests can thus be conceptualized as being less educationally dependent and more age-sensitive than the Trail Making Test. As illustrated earlier (see Figure 4.3, p.96), Schaie's Verbal Meaning test revealed a tendency for a sustained differentiation rather than dedifferentiation effect in the later 70's and 80's age stages. As with the Trail Making Test, Schaie's Verbal Meaning task stands in contrast to the Space, Reasoning and Number tasks in that it is arguably a more educationally dependent task. Thus taken together, these cumulative findings on the linear sustained differentiation effects, versus the non-linear initial differentiation and subsequent differentiation effects in association with aging, provide a strong element of conceptual consistency in favour of lawful patterns of age-related differences in variability in association with task complexity and/or the educational loading and age sensitivity across tasks. This degree of functional consistency across so many tests and several studies, tends to suggest that the patterns detected are *unlikely* to be the effects of measurement artifacts (for example, the possible artifact of different time mechanisms involved in the Trail Making Test and Digit Symbol tasks), although this cannot be ruled out and needs further verification.

Finally, in apparent contrast to this entire set of highly consistent findings of significant variability changes in association with age on symbolically and spatially related tasks, is the lack of significant change in variability associated with the symbol letter modalities test in Christensen et al.'s (1994) study. However, this opposite effect can again be explained in the same manner as for the negative findings for Verbal Fluency on Christensen et al.'s data. The older-old Christensen et al. cohort is not directly comparable with the broad range of age groups examined in the Schaie and Shuttleworth-Jordan and associates' studies. Hence the Christensen et al. cohort may already have reached a stage of dedifferentiation in variability and have levelled off at the point of their restricted higher age range, and the pattern of variability changes for visual symbolic tasks detected on the Schaie and Shuttleworth-Jordan and associates' studies would not be evident from the Christensen et al. narrow research window. (See the postulated early dedifferentiation study-window effect as discussed earlier and illustrated in Figure 4.7, lower graph, p.112).

Problem solving. With respect to problem solving ability, a number of findings can be gathered together as follows. As noted earlier from Riegel et al.'s (1967) study on different aspects of language, the classification test yielded the strongest age effect across his language modalities, which was accompanied by the most marked increases in variability of all the modalities, from age 50 continuing into the late 70's. Commensurate with this, on the basis of statistical analysis Beatty (1993) confirmed a highly significant increase in variability in older age for the three aspects of the Wisconsin Card Sorting task which also showed significant decremental average effects; Christensen et al.'s (1994) research revealed a significant effect for age on the Similarities task; and a marked trend of increased variability followed by diminished variability in association with older age was evident on the WAIS and WAIS-R data reported in Salthouse (1991). As also noted previously, Schaie's (1983) Reasoning task, which involved a speeded visual and symbolic component, evidenced a sharp increase in variability after 60, followed by a dramatic decline in the older years. Further, in that the Trail Making Test involves a level of rational ordering, it can also be subsumed to some extent under the broad rubric of problem solving, and as described above, the indication based on statistical analysis is of increased variability with age on this task from age 70 for Trail Making Test A, and from age 70 for Trail Making Test B (Cornfield & Shuttleworth-Jordan, 1996).

Taken together the data from these four studies provides strong evidence for a pattern of significantly increased variability for problem solving ability in the years after 60 in inverse relationship to decreasing average performance. This increase continues into the late 70's when language skills are involved, but with the likelihood of rapid dedifferentiation in the older-old age stages in some instances, especially when speeded visuo-spatial perceptual components are involved.

Motor activity and Reaction Times. With respect to motor activity Thomae et al., (1981) report on a study by Roth (1961) across three equal age groups (within the broad age spectrum from 25 through to 70), in which a range of tasks were tested under the generalized rubric of motor adaptability including hand-coordination, mechanical drawing, tapping, pendulum, speed of reaction, and scribbling. In association with a decrease in mean scores with performance, it is reported that for practically all tests, inter-individual variability increased with age. Thomae reports on similar results in a study by Schmitz-Scherzer et al. (1974) for a pursuit rotor task. Specifically with respect to reaction time the studies reported by Rabbitt (1993b) and Morse (1993) both support between-groups increases in variability from younger to older age groups. Similarly Horn and Hofer (1992) and Bornstein and Smircina (1982) each cite a number of different studies for reaction times which consistently support the indication of increased variance with older age. Most recently Fozard et al. (1994) report on cross-sectional and longitudinal analyses of reaction time data from 1,265 community-dwelling subjects aged 17-96 years which reveal age-associated increases in within-subject variability, as well as increases in inter-individual variability with age.

In contrast to all of these studies, Christensen et al.'s (1994) results on a choice reaction time task did not reveal a significant association with age. Christensen et al. have noted the contradictory finding themselves with respect to Morse's study, and suggest that the explanation is due to the greater complexity of the type of reaction time task in Morse's study. However, in the face of such strong general evidence of increased variability for reaction times across a whole series of studies as reported above, it appears more probable that the explanation lies rather in the higher and more restricted age range of the Christensen et al. sample. As posed for the Symbol Letter Modality and Verbal Fluency tasks in

Christensen's study, again there has probably already been dedifferentiation on their older-old cohort for the age-sensitive speed of reaction tasks (see Figure 4.7, lower graph, p.112). In that the speed of reaction tasks are known to be *particularly* sensitive to normal aging effects, and in that they have a very low association to education (as compared to a task such as Trail Making), it is unlikely that a significant proportion of individuals would be protected from steady normal aging decline on speed reactions tasks. Thus there would be no protective influence against the emergence of a global dedifferentiation effect in older age, as was enabled for the Trail Making Test.

Summary. In summary, on the basis of this review of specific functional modalities, variability trends in association with aging in the later adult years, can be viewed as progressing in intensity from stable (no effect), to increases in variability (differentiation effect), and at their most severe to increases followed by subsequent decreases in variability (differentiation/dedifferentiation effect). As predicted from the Satz BRC theory, patterns of variability occur in a domain-specific manner in association with the way in which average cognitive declines occur, in that as with central tendency age effects, *less* pronounced variability effects are evident on the more stable 'crystallized' versus unstable 'fluid' functions, and the variability effects become increasingly enhanced in accordance with the cascade model of average aging decrement which is the least for 'verbal' abilities, more for 'reasoning' skills and most for 'perceptual speed'. Thus, for example, pure language skills, which don't involve either reasoning or speed, show stable average performance across the age groups, and likewise no changes in variability. In contrast, those verbal skills involving reasoning and/or speed show decline in average performance in conjunction with increases in variability in association with aging. A visuo-perceptual task which also calls upon reasoning ability shows increases in variability in association with aging, whereas more pure visuo-perceptual speed tasks show the more severe initial increase followed by a subsequent decrease pattern of variability due to aging. A summary of all of these variability trends is presented in Table 4.1 (p.121). Table 4.1 is a replication of Table 3.1 (P.69) which was used to summarize the average trends from psychometric studies in Chapter 3, section 3.2.3.2, and which has been expanded to include the commonly identified variability trends in association with the onset of normal neural aging, from the present review.

In terms of specific domains, as with average trends, variability age effects are noted to occur within *all* the modalities delineated for the purposes of this thesis, of general

Table 4.1. Relative age-sensitivity of cognitive tasks: Domain-specific average ability trends, and variability trends, for relatively high functioning healthy older adults.

Less Age-Sensitive Tasks	Trend Ave	Trend Variab	More Age-Sensitive Tasks	Trend Ave	Trend Variab
GENERALINTELLEC-TUAL FUNCTIONING Crystallized Abilities (Long established habits)	**Stable**	**Stable**	**GENERAL INTELLECTUAL FUNCTIONING** Fluid Abilities (Novel problem solving, perceive new relationships)	**Dec**	**Inc**
ATTENTION Sustained Attention	**Stable**	-#	**ATTENTION** General Attention	**Dec**	**Inc**
			Divided Attention	**Dec**	-
MEMORY Recognition	**Stable**	-	**MEMORY** Short-term verbal + visual, Immediate	**Dec**	**Inc**
			Short-term verbal + visual, Delayed	**Dec**	**Inc**
			List learning	**Dec**	**Inc**
			Working Memory with rational ordering	**Dec**	**Inc**
			Working Memory without rational ordering	**Dec**	**Inc/Dec**
LANGUAGE SKILLS Semantic & Phonological Abilities	**Stable**	-	**LANGUAGE SKILLS** Vocabulary/Verbal Meaning	**Dec**	**Inc**
Crystallized (eg synonyms/antonyms)	**Stable**	**Stable**	Speeded verbal tasks with problem solving	**Dec**	**Inc**
			Word Naming	**Dec**	**Inc**
			Word Fluency	**Dec**	**Inc/Dec**
VISUAL SKILLS Unspeeded visuo-spatial copy	**Stable**	-	**VISUAL SKILLS** Speeded visuo-perceptual with rational ordering	**Dec**	**Inc**
			Speeded visuo-perceptual without rational ordering	**Dec**	**Inc/Dec**
PROBLEM SOLVING Nil identified	-	-	**PROBLEM SOLVING** Abstraction (eg Wisconsin Card Sorting)	**Dec**	**Inc**
			Inductive Reasoning	**Dec**	**Inc/Dec**
			Classification (eg Similarities)	**Dec**	**Inc/Dec**
MOTOR ACTIVITY Nil identified	-	-	**MOTOR ACTIVITY** Reaction Time	**Dec**	**Inc**
			Hand Motor Activity	**Dec**	**Inc**

Dec: Decrease; Inc: Increase; Inc/Dec: Initial Increase, Subsequent Decrease.
No Information available from the review.

intellectual functions, attention, memory, language, visual skills, problem solving and motor activity. Specifically the functions within these broader modalities which were sensitive to increases in variability with aging were, fluid functions, speed of information processing, short-term verbal and visual recall, verbal new learning, working memory, verbal fluency, vocabulary, word naming, visuospatial processing and search, abstraction, classification and reasoning, reaction times and hand motor activity. As is clear from Table 4.1, in the initial aging phases *increased* variability trends occur in a linear manner in association with *decreased* average ability. However, still in association with diminishing average age effects, in many instances an overall non-linear inverted-U effect emerges for variability, due to a subsequent decrease in variability in the later age stages. Very apparent also from Table 4.1, is that there are a number of gaps in the available research for variability effects, when compared with the information which exists for average age effects.

4.2.1.3 Artifacts in variance data

Measurement artifacts. In considering the overall validity of the above findings on inter-individual variability the possibility for test measurement artifacts that may contribute to the data must be considered (Christensen, et al., 1994; Salthouse, 1985; 1991; Rabbitt, 1993c). Salthouse points out that there are physical limits on the minimum and maximum times which are conceivably possible on speed of reaction tasks, and consequent floor and ceiling effects would produce artifactual indications about the true nature of heterogeneity between younger and older age groups. Ceiling effects would give rise to artificially inflated variability for older groups relative to younger groups in association with a decline in mean scores, and conversely floor effects would obscure increasing variability in the older relative to younger groups. Hence as Rabbitt (1993b) suggests, it is only when tasks are designed to prevent floor and ceiling effects that genuine differences in variability can be detected.

With respect to the above review of studies on inter-individual variability, both Rabbitt and Christensen et al. consider that the variability data they present are either not confounded at all, or cannot be entirely explained away on the basis of floor or ceiling effects. In that it is not possible to know the exact nature of all the tests used in the above studies a detailed scrutiny of findings in terms of measurement artifacts is not viable. However, brief consideration of the better-known neuropsychological and IQ tests included in the reviewed studies, suggests that many of them are suitably wide-ranging enough not to be subject to truncated effects (for example, Digit Symbol, Trail Making Test, long list learning tests especially in delayed recall form, recall of passages, extended motor activity tests, all relatively time-consuming tasks which call upon complex combinations of a number of functions). Whereas certain subtests on the WAIS IQ test have a propensity for ceiling effects (for example Picture Completion), in the main this IQ test would not have the pragmatic utility it does clinically, if it were heavily prone to floor and ceiling effects across all the subtests. Thus the most at risk tasks for measurement artifacts, are those laboratory type items such as simple time reaction tasks and brief memory span tasks, and the results of variability findings from such studies need to be interpreted with extreme caution.

Further, whereas ceiling effects might produce artifacts in variability comparisons between young and old age groups, it cannot explain variability increases across a range of *older* age groups (for example 50 through to 80 or 90 years) for tasks which show a continuous average decline with age, since in general all of these groups would be performing well below the ceiling and test measurement comparisons between them would thus be valid (see arguments presented in Shuttleworth-Jordan, 1995; Shuttleworth-Jordan & Bode, 1995b). Finally, if measurement artifacts were the prime factor contributing to the identified patterns of increased variability, the direction of trends should make sense conceptually according to test measurement characteristics. However, the patterns that have emerged from the above review have compelling conceptual validity, rather, in terms of functional modalities and their expected manifestation in terms of neural aging processes, and the hypotheses that have been delineated for the purposes of this thesis from the Satz BRC theory.

Thus, on an analysis of the broad nature of the test modalities used in the reviewed research, the range of the age groups studied, and the conceptual indications, it is considered that measurement artifacts cannot account in any significant manner for the overall findings on inter-individual variability that have emerged from this review.

<u>*Sample artifacts*</u>. There are two important aspects of sampling which may have a significant effect on variability between age groups, and must be considered in the interpretation of effects. These will be discussed further in terms of (i) the age ranges and subject numbers across groups, and (ii) the homogeneity across groups with respect to factors such as education and IQ.

(i) Age ranges and subject numbers. The importance of taking into account the age ranges used within and between different age-related studies for comparative purposes has been emphasized (Anastasi, 1958; Krauss, 1992; Lezak, 1995). In his review of variability studies Krauss (1992) points out the spurious interpretations that can result with respect to age-related inter-individual variability if a narrow age range is compared with a very wide age range, or, as in one of the studies he reviewed, there are vastly different numbers of subjects constituting the groups. (There were only 4 subjects in the young group compared with 24 in the old group of the study he uses as an example!). Overly inclusive age ranges will obscure gradations in change, as will the underinclusive bimodal old-versus-young comparisons such as that of Morse (1993).

In extracting the variability patterns from the above review of studies, careful account has been taken of the contribution of age-range effects, such as in particular with respect to Christensen et al.'s (1994) data which was on a very restricted sample of older-old adults. Further, the synthesis of research findings was focused on the identification of broad trends across the adult age group rather than the minutiae of decade by decade comparisons between studies. To this extent, the broad cross-comparison that has been conducted in this review, between multiple studies comprising in many cases well-represented and comprehensive sets of age ranges, appears to have been legitimate, and has yielded a description of variability patterns due to age that has a high level of conceptual validity. In terms of numbers of subjects in the respective studies, these were not always available for scrutiny (for example the data reported in Salthouse, 1991). Much of the review was based however, on major aging or norming studies with very large numbers (for example the Schaie and Rabbitt studies, and presumably the Wechsler studies cited in Salthouse, 1991), or on smaller studies with balanced subjects across groups (for example the studies of Shuttleworth-Jordan and associates). Thus overall, it is considered that differential age ranges or numbers of subjects between-groups will not have served to obscure the overall patterns isolated in this review.

(ii) Homogeneity of Educational/IQ levels between-groups. On hypothetical grounds in terms of the Satz BRC theory, the powerful effect that educational and IQ levels is likely to have on inter-individual variability in cognitive aging patterns have been discussed (see section 3.2.3.3). Thus it stands to reason that differences in educational and IQ levels between-groups may also produce variability patterns across the age groups which are an *artifact* of a lack of homogeneity between-groups on these influential factors. The reversal of aging trends which can occur due to significant differences in educational level is clearly demonstrated on a study such as that of Stuss, Stethem, & Poirier (1987), with respect to the expected *average* age-related trends on a number of cognitive tasks, and similarly such unmatched samples can be expected to distort variability trends.

The way in which an artifactual sample phenomenon could account for conceptually bizarre variability trends was demonstrated with the Army Alpha results reported in Salthouse (1991). Since educational and IQ levels for each comparative age group were not available for many of the studies reviewed, it was not possible to identify other possibly more subtle instances where this might have occurred. Thus the most reliable comparisons in the above review pertain to those of the Shuttleworth-Jordan and associates studies (Cornfield & Shuttleworth-Jordan, 1996; Shuttleworth-Jordan & Bode, 1995b) in which it is known that the comparative age groups were matched with respect to IQ and/or educational level and these studies supported the variability patterns identified generally from the review. Further, since this was a very broad analysis, across a large spectrum of studies, which nevertheless revealed a relatively consistent and conceptually coherent set of variability patterns, it is not anticipated that between-groups differences in educational or IQ level have confounded the overall indications to any significant extent.

In conclusion, as has been conducted in this review, the presence of both measurement and sample artifacts should be carefully considered as possible explanations for the particular pattern of variability effects, a point that has been repeatedly stressed particularly by Salthouse (1985; 1991). However, this is in as much as to confirm the presence of genuine effects (the issue that is of concern to Salthouse), as *also* to be sure, that artifactual aberrations in otherwise relatively robust patterns of effects do not blind the researcher to the presence of genuinely lawful variability effects (an equally important point that appears to have been overlooked by Salthouse). A clear theoretical perspective with hypothetical expectations as employed in the present review, is a crucial tool in order to see the presence of genuine effects through the smoke screen of powerful artifactual effects.

Intra-individual variability. A final methodological issue to be addressed is the degree to which the cognitive patterns that have emerged with respect to *inter-individual variability*, can be considered stable due to the influence of the closely related factor of *intra-individual* variability (see reviews on the topic in Nesselroade, 1991a; 1991b; Nesselroade &

Featherman, 1991; Wohlwill, 1973). An intricate covariate relationship exists between these two aspects of variability, and hence at any given time of measurement *inter-individual* variability is also composed of a measurement of *intraindividual* variability (*the way in which an individual is different from him/herself across time*). Studies in cognitive aging have demonstrated increases in intra-individual variability with age on verbal tasks such as vocabulary and word fluency (Stones, 1978), and in particular on tasks of reaction time (Hoyer, 1974). Nesselroade (1991a) makes the comment that with respect to the physiological attributes of the very old, "intraindividual variabilities might indicate increasingly errant homeostatic mechanisms" (p. 95). Research indicates further, that in the relative highs and lows of cognitive profiles there can be considerable intra-individual fluctuations, and that not all questionable profiles remain stable or decrease over time (Ivnik, Smith, Malec, Petersen & Tangalos, 1995). Such indications of intra-individual variability with age will necessarily contribute to the extent of inter-individual variability in evidence due to older age: as Nesselroade (1991b) points out, these two aspects of variability constitute the very 'warp and woof of development'.

However, what is of importance is to estimate the extent that the patterns which are being identified from the present review are stable and could be replicated, the chance of which would become illusionary if there was a chaotic proportion of increasingly unstable *test-retest* intraindividual variability due to older age, which in turn would be reflected in the chaotic presentation of inter-individual variability with aging. In point of fact it appears that this is very unlikely to be the case, in that research robustly supports very high correlation coefficients which generally attest to long-term test-retest reliability for tests of intelligence and specific cognitive abilities right into the 70's age group (Hartley, 1988; Maddox & Douglass, 1974; Riegel et al., 1967; Ryan, Paolo, & Brungardt, 1992; Salthouse, 1991; Swartzman, Gold, Andres, Arbuckle, et al., 1987). Salthouse (1991) points out that such high test-retest correlations suggest that individuals tend to preserve their relative positions in cognitive performance across time, in spite of a general decline in the average level of performance. Although Salthouse is using the evidence he cites for intra-individual positional stability to support a notion of relatively little change in inter-individual variability across time, he does this somewhat tentatively, conceding that this test-retest stability factor does *not* in point of fact exclude the possibility that some people could have declined disproportionately while others have retained their previous abilities at a relatively high level.

Further a number of researchers support a high level of stability in the factor structure of primary abilities over time through both the child years, and across the older age groups, (Mitrushina and Satz, 1991a; 1991b; Maddox & Douglass, 1974; Schaie, 1983; Wohlwill 1973; Anastasi, 1958). Cunningham and Birren (1980), however, highlight apparently contradictory findings in this regard, citing a number of studies which support factorial stability and a number of others with interpretations leaning toward factorial variability, suggesting that the former give rise to a *de-differentiation* hypothesis with respect to the progression of cognitive structures in association with older age, whereas the studies which describe diminishing correlations have pointed rather towards a factor metamorphosis, (that is a factor *differentiation*), in association with aging. Birren and Cunningham leave the impression from their article that the matter is somewhat confusing. However, what has become very apparent on the basis of the present review, is that the explanation for these two alternate sets of findings, may lie in the fact that *both* are occurring as part of the *lawful* pattern of intra-individual cognitive alterations identified on cognitive tests with aging. In the light of the inter-individual variability trends described in this chapter, which frequently show the non-linear initial differentiation and subsequent dedifferentiation pattern across the later adult years, this appears to be a highly plausible explanation. Thus in themselves, the opposite factorial trends identified by Birren and Cunningham from the literature serve as cross-validation for the non-linear pattern of inter-individual variability which has emerged as a common phenomenon on the basis of this review.

Overall, therefore, it appears that the domain-specific patterns of inter-individual variability which have emerged on the current review can be considered relatively robust, and that the findings are not likely to reflect such chaotic forms of intraindividual variability in older age as to render the results uninterpretable and not replicable. The most compelling measure of support for the relative stability of intraindividual cognitive performance across the older age groups is gained by the emergent signs from the synthesized data, of age-related variability patterns across a wide cross-section of studies which imply a surprising level of domain-specific conceptual consistency and replicability between studies, considering the uncertainty which appears to have surrounded the matter up to now.

4.2.1.4 Patterns of inter-individual variability: Summary

On the basis of this review on general variability trends from the broad-based data bases available from Schaie, Rabbitt and Salthouse, and a collation of specific variability findings across the six functional modalities delineated earlier for the average normative trends ('attention', 'memory', 'language', 'visual skills', 'problem solving' and 'motor activity'), a surprising wealth of indications has emerged with respect to the nature of inter-individual variability in association with aging. Whilst it is considered that measurement and sample artifacts may account for variability effects in some instances, these have been taken into account as far as possible in the overall analysis, and mainly it is considered that the age-related patterns detected on this review cannot be explained away on an artifactual basis.

There are still considerable gaps in the review in that it was dependent on the collation of available research findings on variability from a disparate set of studies which were most often not designed with the specific intention to investigate variability. Whilst it was possible to reflect on the influence of educational/IQ level, and level of task challenge, there was no research which allowed for comment on possible gender effects. With respect to the various functional domains, in the modality of attention minimal findings were identified, the many composite aspects of memory, language and visual skills were not comprehensively represented, and longitudinal studies were very under-represented. Furthermore, in many instances the variability findings were based on trend rather than statistical analyses. However, in a number of instances there *was* statistical support for the presence of typical patterns, and there was a high level of *conceptual consistency* revealed for the same tests and modalities across studies, which in turn were conceptually coherent in terms of neuropsychological patterns and the causal influences postulated from the Satz BRC theory. Also there were tentative indications of fundamentally parallel trends between both cross-sectional and longitudinal studies in the older adult age groups.

Thus considered overall, a number of conceptually significant characteristics with respect to inter-individual variability have emerged from this review of largely cross-sectional cognitive aging studies which can be posed as the basis for further examination and confirmation as follows:

(i) Generally there is a tendency for inter-individual variability on cognitive tasks to increase with the ages of the groups sampled. However, the extent to which variability is enhanced across older age groups is intricately domain specific, and there is the strong indication that *those domains which show the greatest age-related decline across groups on average normative data, in <u>inverse</u> manner, are the most* *prone to increases in variability*. Thus in accordance with the Cascade Model which has been postulated for the direction of normal decremental age changes, the domains which produce relatively stable variability across age groups are the *crystallized verbal abilities*, and those tasks that employ *reasoning* and *perceptual speed* are increasingly susceptible in descending order to variability age effects.

(ii) It is evident that in association with a continuous decline in normal age-related effects, differences in variability across the older adult years are not necessarily linear but rather are *non-linear*, appearing in the form of an *inverted U-shape*: in many instances following an initial inflation in variability (increased differentiation between the most and least able), an abrupt reduction in variability occurs (that is, dedifferentiation ensues between the most and least able individuals).

(iii) There is support for the hypothetical postulation from the Satz BRC theory, that educational and IQ level may contribute significantly to the variability patterns across test modalities and result in dissociations in these patterns across age groups. This occurs in association with the test- or sample-specific instances when higher educational/IQ level (in association with enhanced brain reserve capacity), can be brought to bear as a protective factor, in contrast with those occasions when it cannot. Specifically the more age-sensitive and less educationally based tasks and/or cohorts with arguably lower educational/IQ level, reveal *earlier* age-associated variability effects, and tend to reveal a dedifferentiation effect in the later stages of the aging process, in addition to a preceding differentiation effect.

In conclusion, the preceding review of variability studies has revealed the presence of increased variability in association with the aging process, not only as an indisputable fact, but also there is the strong indication that it occurs lawfully, in conceptually coherent patterns. It seems probable that confusion has arisen amongst researchers for a number of reasons. These include the non-linear nature of the variability effect which provides apparently contradictory findings depending on the study window adopted, the presence of measurement and sample artifacts, the susceptibility of variability to non-homogenous cohort effects, and the fact that the area has never been systematically reviewed and researched. Certainly, on the basis of the present more comprehensive and theoretically integrated review, it appears that the proposed 'myth' of enhanced inter-individual variability in older age groups can be exploded ----- as constituting a myth in itself! However, the full character of the specific variability patterns, and the causal ramifications of the heterogeneity in older populations, are still matters in need of systematic further investigation and verification.

The areas in need of further investigation are: (i) There is a need to document the characteristics of variability by means of a *more comprehensive investigation* which fills in the gaps in the under-represented functional areas, explores the influence of gender, and provides the facility for the replication of the trends revealed on the present review; (ii) There is the need for a more *systematic statistical analysis* of variability data in order to verify the pattern of trends, and to investigate the relationship between significant average decline in cognitive performance and variability effects; and finally, (iii) There is a need to explore the *characteristics of the individual distributions of scores* which contribute to the group variability changes that occur in association with age.

4.2.2 Origins of inter-individual variability in aging

> *'Many age differences that are reported in the literature, instead of being "caused" by aging, are more likely to be attributable to differences in demographic characteristics and cohort-specific experiences than to adverse maturational changes'.*
> *(Schaie, 1988a, p. 180, Ageism in psychological research).*

> *'I find myself concluding that the use-it-or-lose-it principle applies not only to the maintenance of muscular flexibility, but to the maintenance of flexible life-styles and a related high level of intellectual performance as well'.*
> *(Schaie, 1983, pp. 128-129, The Seattle longitudinal study: A 21-year exploration of psychometric intelligence in adults).*

Finally, in view of the fact that quite clearly there is significant inter-individual heterogeneity in cognitive task performance in association with older age, the possible contributing factors in this phenomenon need to be carefully scrutinized for consideration in the design of an aging research study in this area, and for explication of the results. Whilst chronological age has been used as the independent variable for all the studies used to describe the effect of inter-individual variability in the previous subsection, it is important to be aware that the construct is devoid of explanatory power in itself (see discussion, section 2.2.1.1). This becomes particularly so in the face of increased heterogeneity with aging, and accordingly it has been suggested that in older age, the variable of age itself becomes less relevant and particularly problematic as a predictor of what people can do (Offenbacher & Poster, 1985; Rowe & Kahn, 1987; Schaie, 1988a; Sheppard, 1978; Sterns & Miklos, 1995). Thus it is imperative to consider the interactive biological and cultural forces that are influential in causing the *variations* that occur around the more gross indicator of chronological age itself.

A very limited number of investigations have been conducted with the *specific* orientation of probing the origins of inter-individual variability in older adults (Poitrenaud et al., 1983; Rush et al., 1990; Schaie, 1983; Thomae et al., 1981). By means of correlational research, these few researchers have started with the inter-individual variability identified in their research, and followed this with correlational investigations to identify possible origins of such the variability along biological, psychological and social dimensions. In addition to these studies, a host of theoretical and research literature is available from within the normativity tradition, which explores the amount of variance that can be attributed to various aspects of the biological, psychological and social mechanisms in association with cognitive aging. Since research from the latter orientation does not set out to explicate the pattern of cognitive inter-individual variability in itself, it is not of methodological interest to the present thesis. However, the cumulative indications from these investigations have an important contribution to make in unravelling the significant dimensions which underlie heterogeneity in aging populations, and thus a synthesis of findings from of all of this literature is presented below.

4.2.2.1 Biological and health factors

At the most fundamental level, inter-individual variability in cognitive aging can be seen to begin with a *genetically* programmed biological aging process, consisting of the *complex integration of multiple cellular and organic agents* which result in changes in the human biological system and ultimately in death. Evidence for the extent of the influence of genetic factors on cognitive aging is available from part of the Minnesota Twin Study of Adult Development and Aging (Finkel & McGue, 1993). This study investigated memory and various cognitive and lifestyle variables obtained from 93 monozygotic and 67 dizygotic twin pairs aged 60-88 years, and revealed that 55% of the variance in memory performance could be attributed to genetic factors. Further, in the Swedish Adoption/Twin Study of Aging, the genetic and environmental contributions to high and low extremes of intelligence were explored in a sample of 302 twin pairs with a mean age of 65.6 years (Saudino, Plomin, Pedersen & McClearn, 1994). This research showed that genetic contributions were significant in the overall sample, and were significant and similar in magnitude for both high and low extremes of intelligence. Schaie (1994) reports that average family similarity of about .25 for all mental abilities investigated in his longitudinal study with the exception of perceptual speed, and that correlating relative performance with longitudinal subjects over three seven year intervals provided strong support for stability of family similarity with time and age.

Whereas heritability is reported to be relatively stable for cognitive abilities *within* individuals across the last half of the adult life span (Schaie, 1994; Finkel et al., 1996), the genetic contribution to life span characteristics is considered to vary significantly *between* individuals and thus in measure determine the differential rate at which neural attrition occurs and cognitive aging manifests across individuals (Holland & Rabbitt, 1991b; Swan, La Rue, Carmelli, Reed, & Terry, 1992). There is extensive individual variability inherent in the manner in which the independent partial neurophysiological aging changes occur, and these in turn will differ in association with the forms they may take when expressed by individuals in idiosyncratic interaction with the environment (for details see discussion, section 3.2.2). Rediess and Caine (1996), have drawn attention to an interesting phenomenon noted from a series of studies on the oldest old, which is that individuals in their tenth and eleventh decades show more favourable outcome on both neuropathological changes and cognitive performance than subjects who are 10 to 20 years younger. As with gradual reductions in intellectual capacity with aging, steady visual and auditory sensitivity declines occur, and whilst, as noted by Holland and Rabbitt (1991b), there is no necessary association between the time-course of sensory and cognitive deterioration, sensory losses clearly have radical effects on the varying presentations of cognitive competencies in the older adult population. Specifically, Lindenberger and Ulman (1994) examined relations among age, sensory functioning (visual and auditory acuity), and intelligence in 156 elderly subjects with a mean age of 84.9 years, and found that visual and auditory acuity accounted for 49.2% of the total and 93.1% of the age-related reliable variance in intelligence.

A central feature of the process of the physiological decline involved in aging, is the greater susceptibility it creates to declining *health* in comparison with younger populations, and the propensity for the onset of illnesses which interfere with central nervous system metabolism, thereby reducing cognitive efficiency (Holland and Rabbitt, 1991b; Hindmarch, 1993; Swan et al., 1992). Holland and Rabbitt (1991b) present a comprehensive review of disease-related research studies which support decremental effects on intellectual functioning as a direct consequence of such pathological processes which are common amongst the elderly including hypertension; blood flow circulatory disorders including arteriosclerosis, reduced cardiac output and low cerebral blood flow; chronic pulmonary insufficiency; and diabetes. These authors emphasize also the effects in association with ill-health of medication, fatigue, stress, anxiety and depression which in themselves and in interaction with disease processes, will all have varying degrees of impact on cognitive performance. Salthouse (1991) extensively reviewed the literature on the influence of health factors in cognitive aging studies, and suggests that age-related declining health status among adults typically recruited for such studies does not make a significant contribution to the age-related declines in cognitive performance. However this observation is not supported by a recent study (Uchiyama, Mitrushina, Satz, & Schall, 1996) in which the variable of medical history was found to have a direct effect on neuropsychological functioning in their cohort of normal geriatric persons.

With broad processes such as reduced cerebral blood flow and pulmonary insufficiency there is a borderland between what can be seen to occur as part of the normal aging process and what constitutes a pathological process, as is the case also with age-associated cognitive decline and Alzheimer's Dementia (see discussion, section 2.2.1.3), since differences relate to the *amount* rather than *type* of physiological process; medical diagnosis in such instances is thus a question of *degree* of impairment. In this regard Holland and Rabbitt (1991b), suggest that although there will be individuals who grow old without any definitively diagnosable pathologies, logic tells us that the chances that even such outstanding individuals will suffer no cognitive decrements at all is not possible. This is due to the age-associated cellular, metabolic and neural changes which will be certain to have effects on intellectual capacities, albeit minimized by the absence of distinct physical illness. Thus, in a manner of speaking, neural aging can be conceived of as a mild form of 'brain damage' which presents in varying degrees in all older individuals. That this will have the consequence of enhanced cognitive heterogeneity in older populations, is supported by the well-documented phenomenon of markedly increased inter-individual variability noted on cognitive test performance in brain damage populations (for example, Lezak & Gray, 1984; Ivnik, et al., 1994; Reitan & Wolfson, 1995 as cited in Shuttleworth-Jordan, in press).

The biological bases for heterogeneity in older individuals as explicated above, whilst clearly highly significant, is still only a portion of the total picture, and needs to be considered in conjunction with multiple additional social and psychological influences.

4.2.2.2 Education and intelligence

The powerful effect of education and intelligence on cognitive test performance across all adult age groups and in interaction with age has been extensively documented (for example, Heaton et al., 1986; 1991; Poon, 1985; Lezak, 1995; Reese & Rodeheaver, 1985; Reitan & Wolfson, 1995; 1996). Effects differ widely depending on the functional modalities called upon by the test (Heaton et al., 1986), with tests such as the WAIS Information, WAIS Vocabulary and Halstead-Reitan Trail Making Test B being amongst the most effected by education, in comparison with the WAIS Digit Span, WAIS Object Assembly, Halstead-Reitan Pegboard and Spatial Relations test being amongst the least effected. Although it is important not to treat education and intelligence as entirely equivalent, the two variables are very highly correlated and this has been demonstrated to be true also with respect to older populations (Cornfield & Shuttleworth-Jordan, 1996; Ivnik et al., 1992).

It has been argued that education enhances brain reserve capacity and serves as a significant protective factor against symptom presentation in neurological dysfunction, and consequently also against the effects of neural attrition in normal aging (Mortimer, 1994; Satz, 1993; see earlier discussion, section 3.2.3.3). In some measure of support for such a notion, Heaton et al. (1986) demonstrate a number of age by education interactions on cognitive tasks in which decline in the late 60's appears to be more dramatic for those with the highest educational level of greater than 16 years, compared with those groups with 12 to 15, or less than 12 years of education. They express difficulty in explaining this interaction but hint that it may relate in some way to increased reserve in the higher education subjects which keeps them functioning at a higher level for a longer period of time initially until they reach a threshold after which they decline revealing an apparent 'catch-up' effect by the less well-educated subjects.

Thomae et al., (1981), report on a large cross-sectional study by Rudinger conducted on a generally representative sample in the age range 20 to 90 years, in which it was found that age and education accounted for almost the same percentage of variance in the intelligence scores (approximately 20% and 23% respectively); health accounted for 9.8% of the variance. Thomae et al. contrast these findings with those of a study restricted to a "healthy" older population, in which an even greater contribution of age relative to education to cognitive functioning was demonstrated (0.5% and 19%, respectively), in addition to which the relationship between cognitive functioning and health become much less evident. Rabbitt (1986) similarly noted that when subjects in an older cohort were matched for IQ the effect of age disappeared. However Rabbitt stresses that this does not mean that high IQ will necessarily protect you from cognitive decline however old you are - "this would be only half the truth" (Rabbitt, 1986, p. 112), and he emphasizes that on the same cohort steady age-related declines were noted for IQ. Further as discussed above (see section, 4.2.1.1), Rabbitt (1993c) has demonstrated that in a significant proportion of older individuals memory declines markedly in spite of higher IQ. However, Rabbitt and McInnis (1988) also show on the basis of a large study that people in the age range 54 to 83 years with high IQ test scores remembered first and second memories at an earlier age than did those with lower IQ test scores.

A number of further observations have emerged with respect to the influence of education and intelligence, which reflect the powerful influence of verbal intelligence as tested by vocabulary, and strong links between higher intelligence, social factors and physiological resilience. Poitrenaud et al., (1983) assessed age-related changes in fluid and crystallized functions longitudinally in male subjects in the age range 60 to 79 years, and found disproportionate decline on a Vocabulary test was linked to cardio-arterial disease, as well as to the factor of not maintaining an occupational activity; in contrast on the tests of fluid ability no factor was found to be significantly linked with change in performance. Rush et al., examined different aspects of performance on the Stroop test in a range of older adults from 55 to 81, and found that those with lower verbal intelligence made more errors. Bolla-Wilson, & Bleecker, (1986) on a group between the ages of 40 to 84 years found that higher verbal intelligence was associated with better performance on the Rey Auditory Verbal Learning Test, and accounted for more variance than age. Related to the theme of variability in cognitive performance due to differences in verbal ability, is the broader issue of the highly complex set of variations that occur due to differential processing mechanisms in general, such as speed of processing and working memory (Hultsch, Hertzog, & Dixon, 1990; Light & Burke, 1988; Salthouse, 1991; Salthouse & Coon, 1994). None of these processes, in the view of these cognitive theorists, can in themselves account for all aspects of cognitive functioning, but do account for significant variability in cognitive expression.

Finally, a study reviewed by Thomae et al., unravelled the influential education variable and found that it was correlated with higher aspirations, better self-rated health, greater satisfaction with life, and more social, cultural and intellectual activities. Thomae et al. (1981) further report that in a comparison between survivors and those who died during their longitudinal study, the most impressive difference was in intelligence which was higher for survivors, whereas there was no difference with respect to physicians' rating of health and cardiovascular insufficiency. Survivors were also rated as more active, better adjusted, and to have higher morale than non-survivors.

In accordance with the above indications, aging researchers (for example Rabbitt, 1986; Riegel et al., 1967; Thomae et al., 1981) emphasize the intricate dialectical pattern of inter-relationships that exists between these social, health and intellectual variables, in that higher intelligence will result in higher aspirations, greater striving for educational opportunities, the creation of a higher socio-economic environmental climates, which in turn contribute to higher levels of health and intelligence. No single one of these can be seen to be causative in itself, and the interplay of all of these factors will contribute to varying influences on the heterogeneity of cognitive aging processes.

4.2.2.3 Gender

With respect to gender the findings are somewhat mixed. There is research which provides no support for differential gender effects in association with aging on measures of general intelligence and overall neuropsychological functioning (Heaton, et al., 1986; Huyck, 1990). Commensurate with this Heron and Chown (1967) report that in their aging study the Digit Coding subtest did not show gender effects, and nor did the Mill Hill Vocabulary; Stuss et al. (1987) report no gender differences for either the Trail Making or Paced Auditory Serial-Addition Tests. However in broad terms a study using a structural equation model revealed direct gender-related effects on neuropsychological performance in normal geriatric persons (Uchiyama et al., 1996). Furthermore a number of specific gender differences have been reported for various tests of cognition across the adult age groups as follows.

In Heaton et al.'s (1986) research males did much better on tests of motor speed including the Finger Tapping Test, whereas females did better on verbal modalities tapped by the Aphasia Screening Examination and the WAIS Comprehension subtest. In accordance with these domain-specific findings, Fozard et al. (1994) support male superiority for reaction time across age groups from 17-96 years; Bolla, Lindgren, Bonaccorsy, and Bleecker (1990) report female superiority on the Controlled Verbal Fluency (FAS) task, and women consistently perform better than men on the Rey Auditory Verbal Learning Test (Bolla-Wilson & Bleecker, 1986). Again broadly commensurate with these trends, cross-sectional analyses reported on by Schaie (1994) tentatively indicate an earlier decline for women on fluid abilities, and for men on crystallized abilities. Further data from longitudinal analyses (Schaie, 1994) reveal systematic gender effects in favour of women with respect to verbal meaning and inductive reasoning, and in favour of men for spatial ability. Whilst Heron and Chown (1967) found superior male over female performance on a series of spatially-related tasks across the adult age groups, all of these differences were no longer present from the 50's age decade. Similarly Robbins et al. (1994) report on a differential gender-related findings between earlier and later adult age groups: for females an advantage for information processing rate, which is not apparent in the earlier age groups, gradually emerges as age increases across the over-50's age stage.

Thus generally, where gender effects are noted, the reported domain-specific findings reflect the well-documented general pattern of male superiority in spatial abilities and female superiority on a number of verbal tasks (Huyck, 1990), although it appears that this occurrence *may not be apparent in older age groups.* In addition it appears that a pattern of male superiority for fluid type abilities and speeded tasks that has been noted, may also not be apparent, or possibly may even be reversed, in older age groups. Robbins et al. suggest that consistent evidence for female decrements to diminish relative to males in the later years provides support for a slower rate of neural attrition in women, in association with their known enhanced physical longevity.

Importantly, there are a number of indications which suggest that any negative gender effects with respect to cognitive aging and women must be interpreted with extreme caution due to the possible contribution of female social and educational disadvantage historically. Specifically this is a potential confounding factor in cross-sectional analyses. Schaie (1988b) in particular has warned about the potential for the erroneous interpretation of gender differences in studies which result in discriminatory sexist-cum-ageist attitudes towards women. Thomae et al., 1981, report on several studies on adults in the age range 60 to 90 years of age in which women did worse than men on a series of cognitive tasks but indicate that this is accountable for in terms of lower educational levels; Huyck (1990) suggests that gender differences are decreasing as young women show gains in mathematics and spatial abilities. Finally a striking finding is reported on by Schaie (1994) from longitudinal research which provides support for overall female advantage with respect to aging effects. In a calculus estimation of individual cognitive decrements in association with aging, Schaie lists *male gender membership amongst the most highly weighted variables that predict earlier-than-average cognitive decline.*

Thus in conclusion, although a few studies do not support gender effects on intellectual functioning, cumulatively the majority of research reveals strong support for the contribution of gender to variability in cognitive performance with fairly consistent domain-specific effects. In older age it appears that female sex may have an advantage not only in language skills which is true for all ages, but also in the latter years for information processing skills, implying the probability of slower neural attrition for women in association with known greater physical longevity. Any male advantages that predominate in studies, such as for example with spatial skills, need to be cautiously interpreted due to the lower educational advantage which historically has predominated for females.

4.2.2.4 Social factors

With respect to social structures and their influence on cognitive aging, robust *positive* influences have been found to include high levels of occupational status and income, and urbanization (versus rural) environmental origin. In particular these positive influences will contribute to higher achievement on tasks which involve language, to positive attitudes and motivation in the test situation and to greater test experience and sophistication (Anastasi, 1958; Cunningham & Tomer, 1990; Heron & Crown, 1967; Obler & Albert, 1985; Salthouse, 1992; Schaie, 1993; 1990; Thomae, 1991). More subtle influences demonstrated are high workplace complexity, intact marriage, exposure to stimulating environments, utilization of cultural and educational resources during adulthood, and lengthy marriage to an intelligent partner (see reviews in Schaie, 1990; 1994). Cumulatively a variety of socio-cultural effects in a positive direction contribute to cross-generational variability in the form of upward shifts in cognitive performance. This is illustrated in the analyses of Schaie from the Seattle Longitudinal Study (1983), and also Charness and Bosman (1990) confirm the presence of increasing thresholds on optimal levels of skills performance with advancing generations. Thomae (1991) identified the single most *negative* influence on overall patterns of aging to be the degree of chronic exposure to environmental stress which was closely associated in inverse manner with levels of activity and social competence.

Contributing most particularly to variable influential conditions in elderly relative to younger populations, is the phenomenon of retirement. Associated with retirement are the marked changes and variations that occur in social status and societal networks (Nesselroade & Labouvie, 1985), and vastly differential responses to the retirement phenomenon which occur in terms of activity and levels of satisfaction (Thomae, 1991). Retirement can result differentially in increased, decreased or alternative types of intellectual stimulation following a previous occupation which will all have an influence on cognitive aging processes. Furthermore creating increased heterogeneity amongst aging people, are the extremely diverse environmental conditions which are 'normal' in comparison to younger populations including institutionalization in an old age home (Hindmarch, 1993). Institutionalization has been shown to have largely negative effects relative to community dwelling individuals on Piagetian conservation tasks, and a wide range of neuropsychological tests including memory, learning, verbal responses, word naming and problem solving ability (Hindmarch, 1993; Moscovitch & Winocur, 1992; Obler & Albert, 1985; Reese & Rodeheaver, 1985). Institutionalization is often accompanied by a reduced activity level and a subjective sense of diminished personal control, factors which are associated with poorer performance on cognitive measures in the elderly (Moscovitch & Winocur, 1992; Winocur, Moscovitch, & Freedman, 1987).

Finally, adding to the potential for variability in older adults is the effect of environmental stimulation and its connection with postulated brain plasticity: positive effects on cognition have been demonstrated on older populations by virtue of educational training techniques and physical exercise programmes indicating that cognitive decline can in part be due to factors such as disuse and thus may even be remediable (for example, Black, Greenough, Anderson, Isaacs, & Krystyna, 1987; Clarkson-Smith & Hartley, 1989; Obler and Albert, 1985; Schaie & Willis, 1986; Schaie, 1994; Woods & Rusin, 1988). Huppert (1980) report the remarkable finding from an analysis of memory performance in the elderly in which the low IQ/Education group that participated in three or more leisure activities was found to perform as well as the high IQ/Education group. Ultimately, however, the inevitable reductions in social stimulation which must ultimately occur with advancing years, due to physical frailty and increasing sensory impairment, will necessarily also contribute (that is aside from neural or maturational processes), to differential degrees of decrement in cognitive faculties.

4.2.2.5 Personality/emotional factors

Moscowitz and Winocar (1992) suggest that taken together social and personality factors exert a more important influence than age on cognitive decline in elderly populations. However, whilst the social influences as discussed above have been robustly demonstrated to exert a powerful effect on variations on cognitive aging processes, the influences of personality and emotional factors are not as clearly evident. On the basis of their large aging study, Heron and Chown (1967) report that "the most notable thing about the correlations of personality with other variables is their smallness" (p. 114). Thus in their study the variables of 'Emotional Instability' and 'Unsociability' were of trivial significance in determining outcome on the psychometric test variable. Similarly from a review of aging studies, Kausler (1990) indicates that neither anxiety (whether state or trait), nor motivation were of importance as causative factors for age-related deficits on most cognitive tasks. In a review by Gold and Arbuckle (1990), they conclude that the communality of variance for studies on associations between personality variables and cognition was low, and where associations were found they did not vary with age. Generally, Neuroticism appeared as the most predictive factor of poorer cognitive performance (Gold & Arbuckle, 1990). Linking with this, Yesavage (1989) has demonstrated a significant main effect for Neuroticism for elderly subjects who did not improve in response to cognitive training techniques in memory.

However, despite the above broadly negative findings on the influence of personality and emotional factors, a number of subtle age-related considerations have been identified. Blanchard-Fields and Camp (1990) found that older individuals were much more differentiated in their responses on emotionally-loaded problem solving tasks than younger individuals, and varied between two extremes of cut and dry and/or highly effective emotional regulating responses. Wohlwill (1973) highlights the manner in which social environment fosters differential amounts of anxiety and fear of failure in performance on cognitive tasks developmentally, and this has particular pertinence for elderly populations. Not only are older individuals confronting diminished physical and mental capacities in reality, but in addition they live alongside a milieu of ageist attitudes about failing capacities in the elderly (Schaie, 1988a). Together these factors are likely to contribute to self-fulfilling prophesies around failure hence contributing to poorer performance on cognitive tasks. In partial support of this Poon (1985) reports on a number of studies that show how older groups are more sensitive to memory failure than younger groups. Dixon (1992) notes the variable manner in which older adults cope with perceived declines in everyday cognitive performance: some individuals are potential compensators and produce strategies designed to sustain a comparatively high level of functioning, whereas others feel out of control, reduce their expectations, and thereby limit their performance. Schaie (1983) maintains that the role of flexible personality style in midlife onwards is a major contributing factor to the maintenance of a high level of intellectual performance in the later years.

Further, attention has been drawn to the differential manner in which emotional characteristics will manifest in older people and influence performance on cognitive tasks. Reese and Rodeheaver (1985) emphasize that, as with children, the elderly are particularly sensitive to levels of anxiety, apprehensiveness, low self-esteem and cautiousness. Such emotional factors vary in the extent to which they have both negative and positive influences on cognitive competence (Gold & Arbuckle, 1990; Reese & Rodehenver, 1985; Rush et al., 1994; Salthouse, 1990). Nesselroade and Featherman (1991) have demonstrated how labile the phenomenon of depressed mood is in the elderly intra-individually, and warn about how this variation will influence the measurement of psychological variables at different points in time. Clinical depression is particularly prevalent in the elderly population as a whole, and its detrimental effect on cognition in the elderly has been extensively documented (for example Albert, 1981; Cunningham & Tomer, 1990; Bieliauskas, 1993; Lamberty & Bieliauskas, 1993; Zung & Zung, 1986). However, as suggested in recent literature, the effect of depression on cognitive test performance may have been exaggerated. It appears that it will largely only have a significantly decremental effect on cognition in cases of severe endogenous depression including the presence of vegetative symptoms (Bieliauskas, 1993; Bieliauskas & Lamberty, 1995).

Finally, as aging research reaches higher levels of sophistication, it is emerging that cognitive processing styles vary in relation to the differential demands of changing age stages, from more adaptive mechanisms in younger age groups, to more conservative and rigid processes in the older years grounded in a basis of accumulated wisdom (Labouvie-Vief, 1985). Thus it is emphasized that whilst cognitive modes in older persons may be qualitatively *different* from those of younger persons, they are not necessarily *deficient* relative to youth-centred criteria (Cunningham & Birren, 1980; Labouvie-Vief, 1985). Furthermore it has become clear that the expression of cognitive skills is not universally age dependent in terms of its peaks and valleys in that it may be domain specific (for example mathematical skills peak in the 20's and 30's decades, whereas verbal skills peak in the 50's and 60's). Additionally such peaks and dips will shift in interaction with the level of creative engagement which is consequent on the career stage of individuals. The latter can vary *considerably* in relation to age for a number of social reasons, such as late career development in women consequent on meeting maternal obligations, or second-career implementation for any number of reasons. Together these factors serve as significant further sources of heterogeneity of intellectual expression in older age groups. There will be those older adults who will evolve further away from adaptive cognitive mechanisms and adhere to more rigid mechanisms. In contrast there will be those more flexible persons such as identified by Schaie, who will be capable of greater differentiation via the use of both adaptive as well as more conservative cognitive styles in combination with years of accrued life experience, thus enabling considerable compensation for lowered neural resources. Finally, there will be marked variability in the expression of different combinations of rigid versus flexible cognitive styles, in interaction with the variable manner in which people are represented across a wide spectrum of differential stages of engagement in life's processes.

4.2.2.6 Origins of inter-individual variability: Conclusions

> *"Frank Trippett (Time, March 24 1980) has drawn attention to the popular misconception of the aged as helpless senility, when in fact not more than 8% of old people suffer from it ..... Perhaps the word 'retirement' is psychologically unfortunate, particularly as there is so much evidence that human potential is almost unlimited."*
> *(Shuttleworth, 1994, Preface, Age Need Not Weary You).*

> *"Many people reach the peak of their careers at an advanced age. Arthur Rubenstein was still drawing full houses to his piano concerts at the age of 90...... Sophocles wrote his most famous play Oedipus when he was 80 and George Bernard Shaw wrote his last play when he was over 90......Politicians tend to continue almost to the end of their lives ....Winston Churchill led Britain to victory when he was 71...... Perhaps we can claim that one of the most remarkable of all was our own intellectual giant, General Smuts..... He was still climbing Table Mountain when he was near his death at 80......"(Shuttleworth, 1994, pp. 57-58, Age Need Not Weary You).*

In conclusion, this delineation of the variable biological, social and psychological influences that occur in association with older age is entirely commensurate with the empirical evidence for enhanced inter-individual variability in cognition, which so clearly emerged from the review in the earlier subsection. In the final analysis, everyday observation of the wide variations that occur between older individuals cannot fail to provide further confirmation of aging heterogeneity. Whereas there are those who remain physically healthy into the late 80's with excellent physical, mental and cognitive health, enjoying quality of life and continuing to be creative (Shuttleworth, 1994), there are those that literally 'retire' from full participation in life and become steorotypically 'old' and frail decades earlier. Clearly these vastly discrepant consequences of aging come about via a complex interactive effect of at base level genetic and inevitable neural aging mechanisms, in a multiplicative interrelationship with the type of biological, environmental, social, personality, emotional and health factors noted above, each of which in themselves are subject to increased degrees of variability by virtue of the aging process, and all of which will have differential effects on aspects of cognitive functioning. In confirmation of this, the extreme complexity of the interactive and moderating influences on cognitive functioning in association with age has recently become a matter of some emphasis in the literature (Finkel et al., 1996; Reitan & Wolfson, 1996; Shuttleworth-Jordan, in press; Uchiyama et al., 1996).

The most clearly established *positive* influences on cognition in association with aging appear to be the closely related variables of high levels of education and IQ, and especially high Verbal IQ, high socio-economic status, and good health. Personality and emotional factors are more subtle and less clear in terms of their relative positive and negative aging effects, although it does appear that non-rigid personality attributes and a flexible cognitive style may contribute positively to cognitive aging. Gender effects are also subtle and somewhat mixed, although generally it appears that in the earlier adult years males show superiority in spatial skills and fluid skills including speed of information processing, whereas females tend to show superiority in verbal/crystallized skills. However, where education is kept constant, and particularly in the years after 50, it appears that female disadvantage in spatial skills and information processing disappears. Further male gender membership has been shown to be one of the weightiest predictors of earlier-than-average cognitive decline which indicates, by contrast, that female gendership is associated with more successful aging. The late dedifferentiation effect (as detected in the earlier review on variability patterns) suggests that ultimately negative biological influences will prevail almost invariably for all people, regardless of the initial wide inter-individual variations which take place due to the interaction of varying genetic, health, social, personality and gender influences.

The indications from the review on origins of variability are clearly consistent with the hypothesis posed from the Satz BRC theory (see section 3.2.3.3) that higher levels of education and intelligence will provide a protective effect in cognitive aging. Importantly the review also reveals indications of a more definitive indication for gender effects than was directly available from the Satz (1993) review. Thus it is possible to strengthen the second BRC hypothesis (which tentatively on the basis of female longevity alone, infers that gender effects are likely to favour females, see section 3.2.3.3), in that cumulatively across several sources it appears that where education is equivalent for males and females, *there is the strong likelihood that in the later age stages female gender membership will provide a protective effect*. Whilst this review of the origins of variability yields a comprehensive description of the many possible causes of heterogeneity in association with the aging process, the research almost exclusively focuses on group effects, and correlational studies which will obscure the details of outlying individual effects. With the exception of the research of Rabbitt (1993b; 1993c) in which he demonstrates that increasing proportions of individuals in older age groups have memory impairment in the absence of equivalent intellectual decline, there appears to be no other research which examines the individual characteristics of the extreme low and high scorers across older age groups. Thus there are a series of important hypothetical suppositions with respect to the specific characteristics of the exceptional 'successful' versus 'unsuccessful' agers which could usefully be further investigated in a study which focuses specifically on patterns of inter-individual variability in association with adult aging.

4.3 CHAPTER SUMMARY

This chapter comprised a comprehensive review of the whole issue of inter-individual variability in cognitive aging. The reasons for its importance as an area of investigation in aging research were identified from the psychometric, developmental, socio-political and clinical literature, from which an argument was developed to suggest that the neglect of a differential perspective in aging research will lead to a distorted view of the aging process. It was pointed out, however, that this widely held view is strongly challenged in the literature by a few influential, albeit scanty, reviews which lead their authors to claim that the presence of increased heterogeneity in association with aging is an illusionary doctrine which is not clearly supported on an empirical basis.

In order to gain clarity in this debate, a comprehensive review of existing research on the patterns of inter-individual variability in cognitive aging was conducted, using the neuropsychological framework and a set of hypothetical indications derived from the Satz brain reserve capacity (BRC) theory as the eye glasses from which to guide the analysis. On this basis a lawful and conceptually coherent pattern of domain specific increasing variability trends in association with aging was demonstrated, which occurs initially in inverse manner to the onset of decremental average trends, and subsequently may show a reversal of that trend, depending on the age-sensitivity of the task modality, or the educational level of the cohort. It was suggested that a number of factors may have previously prevented the identification of patterns from an apparently 'patternless' conglomeration of data, and serve to explain the confusion that has existed in the literature up to now: the non-linear pattern of age-related variability trends; the sensitivity of variability to artifactual effects; and in particular the lack of a systematic review of the literature conducted through an integrating theoretical prism. The latter has been necessary in order to expose the lawful relationships contained within the data.

Finally, a review of possible origins of heterogeneity with aging was conducted and revealed a highly complex set of interrelating biological, environmental, social, personality, emotional and health factors that operate in association with the aging process, and which can account for the evidence for the variability patterns noted in the earlier review. However, taken together, the reviews on the *patterns* of variability and the *origins* of variability require further investigation and replication, and leave a number of important questions unanswered. In particular there are significant gaps in the information acquired for variability within the functional modalities, and often the information was only on the basis of a single study and thus requires verification; the patterns reported were based largely on trend identification rather than formal statistical analyses; and there has been minimal investigation of the characteristics of the individual distributions across age groups, or examination of the extreme high and low scorers in order to comment more exactly on influences in association with 'successful' versus 'unsuccessful' aging.

5

CHAPTER 5: A MODEL OF VARIABILITY

Chapter 5 comprises the brief and highly focused culmination of the review of the cognitive aging literature from the perspective of inter-individual variability, which was contained in Chapters 1 to 4. The purpose of the chapter is to delineate a working model of variability in aging deduced from the indications that have arisen out of the research review. The model serves as the basis for hypothesis testing in the investigation to follow.

As argued in Chapter 3 (section 3.1.1), from the metatheoretical perspective of the principle of complementarity, there is a *particular* need for a clearly delineated theoretical frame of reference for the scientific investigation of a complex phenomenon such as the cognitive behavioural aspects of aging. In the sixties Schaie expressed the fact that it was timely to look for theoretical models in order to integrate the plethora of then available age-related research data, and thirty years later, Salthouse highlights the fact that if this was true then, "it is even more so now" (Salthouse, 1991, p. 350). Thus, the development of this model for the integration of patterns of variability in cognitive aging was with a view of meeting this challenge, rather than by means of the present research, to contribute yet another set of unrelated observations.

5.1 LEVELS OF THEORETICAL OPERATION

> *"Frameworks and descriptive generalizations are useful, and perhaps even essential, levels of theoretical discourse, but they are not substitutes for the specification of how the theoretical concepts are related both to one another and to empirical observations. These latter functions are served by theories and models, and consequently a theoretical perspective must contain speculations at the level of theories and models if it is to be considered a viable explanation for the phenomena of interest".*
>
> *(Salthouse, 1991, p. 14. Theoretical perspectives on cognitive aging).*

Salthouse (1991) has distinguished a hierarchical set of levels of theoretical discourse which will be adapted here for explanatory purposes in the development of the variability model. As depicted in Table 5.1 (p.147), classes of theoretical operation can be seen to range from the low level of pure observational data (such as the quantitative measures of performance on cognitive tasks), through the level of descriptive generalizations (which is the attempt to

Table 5.1. Schematic representation of five levels of theoretical operation as applied in the thesis (adapted from Salthouse, 1991).

LEVEL	AS APPLIED IN THESIS
Frameworks Concepts and Principles	**Neuropsychological Framework**
Theories Causal Mechanisms between Concepts	**Satz Brain Reserve Capacity (BRC) Theory**
Models Causal Mechanisms between Concepts in Specific Contexts	***Model of Variability***
Descriptive Generalisations Integrative Summaries of Empirical Data	**Table of Variability Effects**
Empirical Data Disparate sets of Research Data	**Variability Data Isolated for Review**

summarize patterns depicted from the observational data, often in the form of descriptive taxonomies and having no explanatory power), further up to the level of models (which account for the relationship of variables within a highly specific context, and link observational phenomena with theoretical concepts) and theories (which define the broad causal mechanisms and interrelations between sets of concepts), and finally up to the level of frameworks (which is the broadest level consisting of a set of concepts and principles, but which does not define causal relationships between concepts with any specificity). The levels can be seen to be hierarchical in the sense that each one is necessarily subsumed beneath the other: observations are necessary for descriptive generalizations, descriptive generalizations are needed to inform models, a number of models may participate in a theory, and a number of theories may be derived within a framework. A framework thus literally provides the 'frame' or 'umbrella' level of operation for all the others.

This set of hierarchical levels of theoretical discourse can be elaborated on with specific reference to the present thesis as follows. Starting at the top of Table 5, the dissertation was conceived of within a *neuropsychological framework* which provided the very broad, and relatively loose set of concepts involved in the description of cognitive performance in terms of functional modalities and their association with cerebral mechanisms, and the fundamental principle of needing to take account of demographic characteristics of testees in the interpretation of test data. At the next step down on the table, discourse occurred at the level of theory (falling under the umbrella of the broader neuropsychological rubric), and causal mechanisms for the onset of neural aging changes were derived from within the context of the *Satz brain reserve capacity (BRC) theory* specifically with reference to variability. Going now to the bottom of the table, the *disparate sets of empirical data* on variability available in the literature were reviewed, and further to this, moving one step up the ladder of the hierarchy, the results were summarized into a table which provided the level of *descriptive generalizations* of the data (as appears in Chapter 4, Table 4.1, p.121). Following this it is clearly evident that there is a *missing link* right in the middle of this hierarchical structure, and that is at the level of a *model of variability*.

The nature of this missing link is as follows. Operating from the position at the *top* of the Table 5.1 hierarchy, that is under a broad neuropsychological framework and from the perspective of BRC theory, it was possible to extrapolate a set of hypotheses with specific applicability to variability in association with neural aging as follows (see section 3.2.3.3):

(i) From the neural perspective on aging it is hypothesized that due to the inherent inter-individual variability in the complex physiological mechanisms of neural attrition due to aging, as with all forms of brain impairment, there is likely to be increased variability of scores on cognitive tasks in association with aging;

(ii) From BRC theory, it is hypothesized that the inter-individual variability in cognitive test performance in association with the neural aging process will be influenced by inter-individual differences in levels of brain reserve capacity which alter the threshold of vulnerability to, and protection against, the onset of functional symptom presentation. In particular, from the Satz review it is expected that low education and IQ, high task challenge, and by inference male gender, will serve as threshold-lowering factors and will thus be associated with earlier symptom onset with aging; and conversely it is expected that high education and IQ, low task challenge, and by inference female gender, will act as threshold-raising factors and will be associated with later symptom onset with aging. (The Satz review did not provide a definitive direction for gender effects, however these were inferred from the well-known fact of female longevity in conjunction with general BRC principles).

(iii) Finally, from the psychometric perspective in conjunction with BRC theory, it can be hypothesized that differential sensitivity to aging processes between different cognitive tasks and functional modalities, as identified on central trends analyses, will have an associated differential effect on the patterns of variability in cognitive task performance. This is in that tasks which are less sensitive to aging on central trends analysis are evidently those that present less challenge to brain capacity thresholds in association with aging, and those tasks which have been demonstrated to be more sensitive to aging on central trends analysis, are evidently those that present more challenge to brain reserve capacity thresholds; hence differential effects between test modalities will be reflected similarly for both central trends and variability data in that for those tasks and modalities where central trends effects are more marked, variability effects will also be more marked and vice versa.

From the review which focused on *patterns of variability data per se* which were available for discussion from the literature (see section 4.2.1), with the exception of gender effects for which there were no comparative variability data available, these hypotheses were all tentatively confirmed in that (i) in general there did appear to be enhanced variability in association with the onset of the aging process, (ii) the patterns of variability did appear to be differentially affected by educational/IQ level (as inferred from the effects on arguably high functioning cohorts), and task challenge, and (iii) the variability changes were domain specific and occurred in a manner which was related to decremental average age changes in cognition as established on the basis of central trends analysis. Further, an additional set of observations emerged about the particular *characteristics* of variability across later age groups, such as the domain-specific linear effect which stands in contrast to the inverted-U non-linear effect (contrasting patterns within Figure 4.3, p.96; Figure 4.8a versus Figure 4.8b, p.115), and the cohort effect (Figure 4.4, p.100; Figure 4.6, p.107), which, as discussed in each instance in Chapter 4, can fundamentally be *explained* in terms of BRC theory. However the problem is that these characteristics cannot specifically be *described* in terms of that theory. The conceptual underpinnings of BRC theory are too broad-based to delineate the intricacies of the particular aging *patterns* of variability which appear to be emerging. For this purpose a more specific model of variability is necessary.

Considering the position from the *bottom* of the hierarchy in Table 5.1. In an attempt to synthesize the patterns which emerged from the review of empirical data, the observations were summarized and tabulated in order to provide a more generalized level of description (Table 4.1, p.121). However, trying to capture the findings in this restrictive and inflexible taxonomy was not a satisfactory mechanism for their presentation. It was only possible to depict a synthesis of overall trends, and the table did not have the capacity to reflect the shifting patterns which produced apparently contradictory findings in some instances. In order to gain flexibility in the description of specific variability effects a more 'elastic' process model of variability is needed, rather than a rigid taxonomy. Thus from whichever direction on the hierarchy of theoretical discourse the problem is approached (top-down, or bottom-up), there is on the one hand too much flexibility at the uppermost level of framework and theory, and on the other hand insufficient flexibility at the lower level of generalizations based on empirical observations and taxonomy. It is clear that what is missing is the level of theoretical operation that would be offered by a specific *model* of variability.

5.2 THE EMPIRICAL BASIS FOR A MODEL OF VARIABILITY

In the process of building a model of inter-individual variability in cognitive aging, available modes of describing variability trends, and the key patterns that have emerged from the literature and research review, will be highlighted and finally synthesized into a single working model.

5.2.1 The overall older age inverted-U effect

As noted in Chapter 4 (Figure 4.2, p.86), Wohlwill (1973) presents a set of two-dimensional graphical illustrations of response magnitude to show how group variance trends might be conceptualized to differ across age stages depending on the functional modality. Wohlwill has identified three possible variability patterns in association with advancing age including

(i) a non-linear pattern of an initial increase and subsequent decrease in variability, (ii) a pattern of increasing variability, and (iii) a pattern of no change in variability. These trends have been used for explanatory purposes by Wohlwill with specific reference to child development. However, they can usefully be extrapolated in unidimensionsal form to describe the range of possible variability trends, as expressed in units of variance measurement, across the adult years (see Figure 5.1, p.152). In the review of variability patterns in Chapter 4, each one of the three variability trends (as schematically presented in Figure 5.1, p.152) was identified to occur in isolation or in combined form in association with cognitive aging. On an analysis of these effects, it appears that Trend 1, the *non-linear inverted-U effect of an initial increase and subsequent decrease in variability,* characterizes the *overall* pattern of variability effects which occurs in association with diminishing average performance on cognitive tasks due to the neural aging process.

This *overall* effect, which is further schematically illustrated in Figure 5.2 (p.153), is fundamentally explicable on the basis of BRC theory (as elaborated on in Chapters 3 and 4) as follows. It is expected that due to different levels of pre-existing cerebral reserve in association with the onset of neural attrition, the presentation of symptomatology will occur differentially between individuals, and be reflected in *increased* variability of cognitive test scores in association with increasingly poor average cognitive performance. Specifically, due to protective factors which *raise* the threshold of symptom onset, a notable proportion of individuals will not present with much cognitive fall-off; alternatively, due to vulnerability factors which *lower* the threshold of symptom onset, a notable proportion of individuals will show particularly marked fall-off. The overall consequence of this will be a significantly widened distribution of scores. Thus where poor performance is equated with low scores, it is expected that an initial inverse relationship between mean scores and variability ratings will occur as reflected in Figure 5.2. Later, as neural attrition becomes more advanced, it is expected that protective factors resulting in raised BRC thresholds for some individuals will become less effective such that these previously good scorers will perform closer to the norm. The result will be a reduction in markedly variable symptom presentation between individuals. The distribution of scores will narrow again and will be reflected in *reduced* variability of cognitive test scores in association with continued attenuation of mean scores.

Thus with reference to Figure 5.1 (p.152), it appears that Wohlwill's Trend 2 (the *increased variability effect*), when it is identified, represents the initial phase of Wohlwill's Trend 1

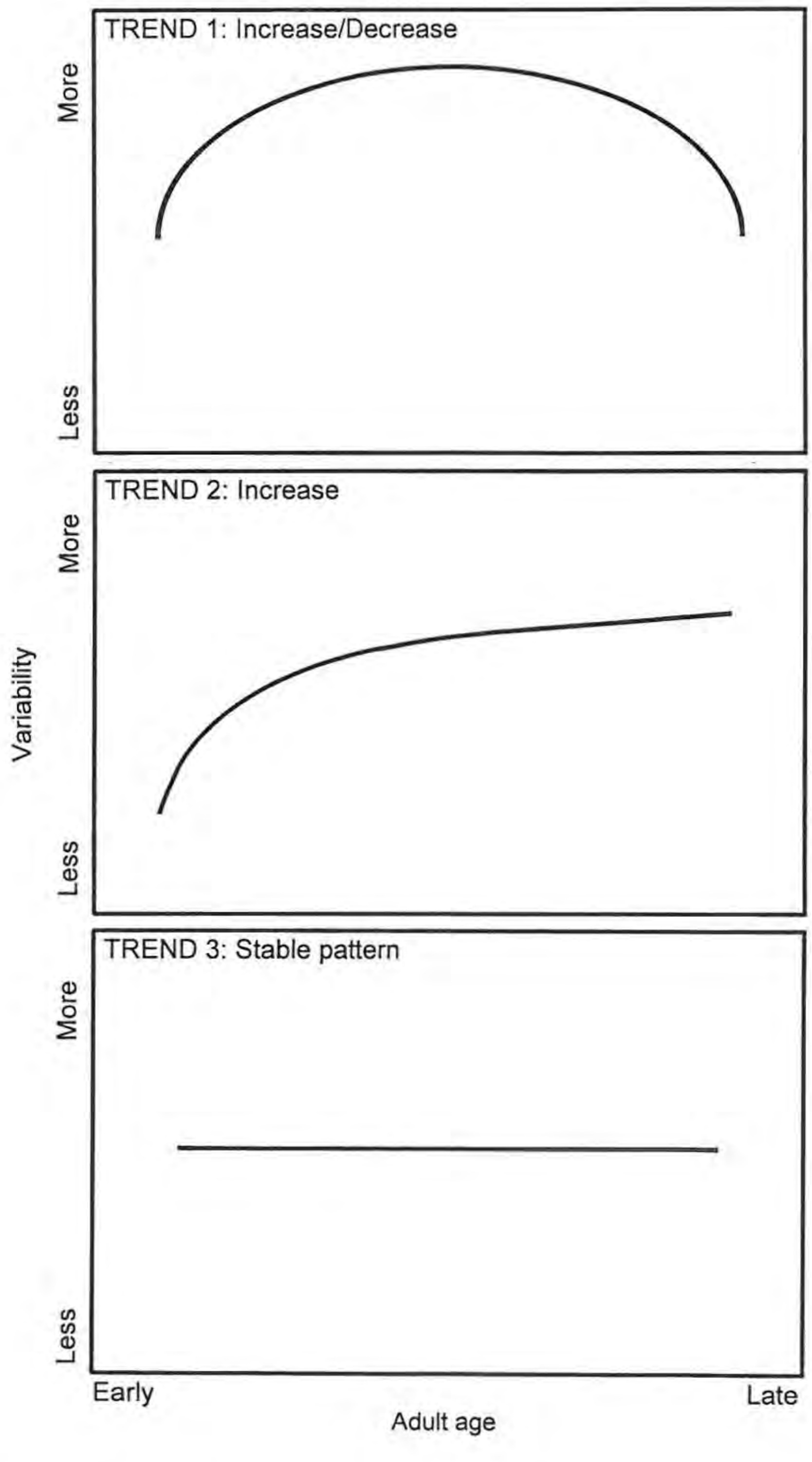


Figure 5.1. Schematic Representation (in units of variance measurement) of three variability patterns that may occur progressively in association with age.

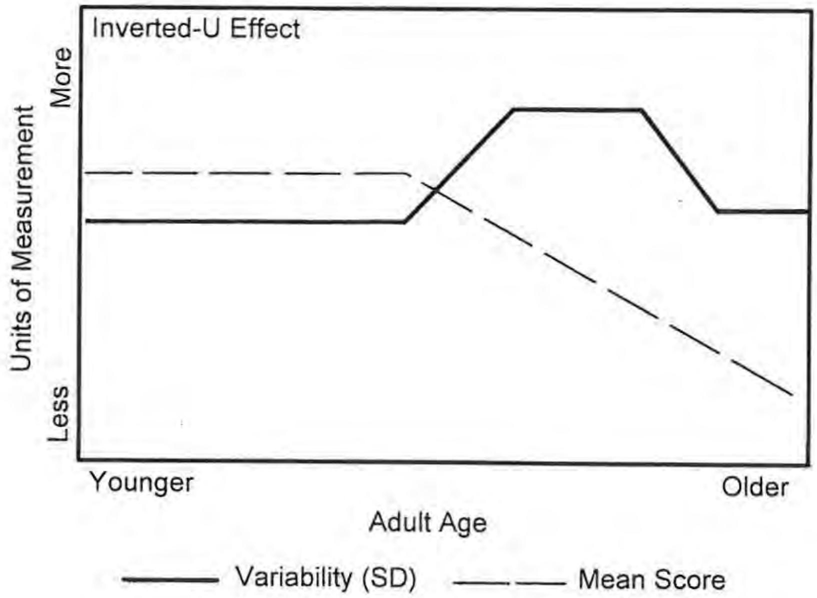


Figure 5.2. Schematic representation of the non-linear inverted-U variability effect as it occurs in association with declining average performance across the adult age range.

(the *overall increased/subsequent decreased variability effect*). Trend 3 (the *absence of variability effect*), is likely to be an artifact of a study window with an overly narrow age range in the very early or very late phases of the process, which would not serve to detect either increases or decreases in variability. Alternatively Trend 3 (the *absence of variability effect)* could be an artifact of a bimodal comparison of early and late old age which omits to take into account the interim bulge in the overall variability pattern, thus artificially revealing no change in variability in association with aging. Following on this it possible to designate Trend 2, which depicts increasing variability alone, as the less severe *initial aging effect*, and Trend 1, which reveals the non-linear inverted-U effect, as the expression of the more severe *overall aging effect*. Thus any other major increases or decreases in variability which occur especially in the later adult years, should raise the query of artifactual effects, such as was demonstrated with respect to the uncharacteristic late double differentiation effect which appeared on the Salthouse Army Alpha data (see Figure 4.5, p.102). It is now possible to proceed further and delineate two important influences which appear to occur in association with the overall inverted-U variability effect and neural aging, including cohort-specific effects, and domain/task-specific effects, as follows.

5.2.1.1 Cohort-specific effects

It appears that the age-related variability changes described above, of an initial increase and subsequent decrease in variability, will occur relatively earlier or later depending on particular cohort characteristics. The likelihood that this effect will occur has been explained in Chapter 4 in terms of the Satz BRC theory (which postulates threshold factors for symptom onset in neural disease) on the following basis. Specifically for the purposes of this thesis a fundamental notion was proposed, which is that in terms of the BRC framework *aging in itself can best be conceptualized as a form of neural impairment or 'disease'*. This is in that, in association with advancing age, a process of progressive neural attrition occurs which serves to reduce cerebral reserves such that there is the onset of deteriorated function. Consequently it is argued that aging (as with any neural disease) constitutes a powerful *risk* factor for lowered brain reserve capacity and the presentation of symptomatology. *Thus, in aggregated form with other factors which serve to raise or lower brain reserve thresholds differentially between individuals, the effect of aging is likely to show differential effects.*

As noted earlier from Satz (1993) it has been argued that lower levels of education and IQ constitute risk factors which result in the reduction of brain reserve capacity, and hence the lowering of the brain reserve threshold with respect to symptom onset in the progression of neural disease. Moreover, as noted in the discussion on the origins of variability (see section 4.2.2), lower levels of education and IQ are inextricably associated with each other and with a lower socio-economic level. These in turn are associated with an increased predisposition to incipient disease, including brain disease, the presence of which constitutes a further risk factor in itself in terms of the lowering of the brain reserve threshold and the enhancement of vulnerability to symptom onset. Conversely higher levels of IQ and education are linked to higher socio-economic status. This in turn is coupled with the greater likelihood of enriched environments and healthy life-styles which serve as protective factors against lowered brain reserve capacity and the onset of morbidity. Thus in a population with the risk factor of lower levels of education and/or IQ (which in turn are inextricably linked with a lower socio-economic level, more deprived environments, and the greater predisposition to developing disease), a significant proportion of individuals are likely to cross the brain reserve threshold sooner than would occur for individuals from a population with the protective factor of higher levels of education and/or IQ (with associated higher socio-economic status, relatively enriched environments, and lower predisposition to disease). Consequently such lower functioning individuals would fail to sustain a hold-pattern on cognitive abilities in association with the yet additional risk factor of neural aging (as with any other compromised neural state) for as long as would occur for individuals from a higher functioning population. Hence, due to the complex aggregation of risk factors specific to low functioning populations, in contrast to the complex aggregation of protective factors specific to high functioning individuals, the initial differentiation/subsequent dedifferentiation effect on cognitive tasks would occur sooner for lower functioning groups in association with aging than it would for higher functioning groups. (It is important to note that the underpinnings of BRC theory denote the positive influence of protective factors which serve to raise the threshold against symptom onset. Thus the gap between the onset of morbidity in the low functioning group compared with the high functioning group, is not merely due to negative influences in the more vulnerable low functioning group. For example, as noted earlier in the thesis, Satz (1993) draws attention to literature which reports beneficial cortical effects due to enriched environmental conditions).

The above explanation of BRC threshold effects for lower versus higher functioning cohorts formed the basis for understanding the paradoxical effect noted for fluid and crystallized abilities in the McArdle study (Figure 4.6, p.107). More importantly, the occurrence of *earlier* versus *later* variability effects, in association with arguably lower versus higher functioning samples, appears to be robustly demonstrated across most subtests on the comparative WAIS and WAIS-R data presented in Salthouse (Figure 4.4, p.100).

In addition to the influence of education and/or IQ, it appeared from the Satz (1993) review, that another possible source of cohort effects on variability might be inherent in gender differences. As noted earlier there were no clear indications from Satz (section 3.2.3.3), or from the review of specific patterns in variability data (section 4.2.1), which elucidated a particular direction for such gender effects. However, as discussed in section 3.2.3.3 (see subsection on brain reserve capacity theory), the shorter life expectancy of men relative to women, implies a lowered symptom threshold for morbidity generally in association with normal aging for men, and conversely raised symptom threshold for women. The natural corollary of this is that there would also be lowered symptom threshold and hence earlier presentation of impaired cognitive functioning due to normal aging for men; and conversely raised symptom threshold and later presentation of impaired cognitive function due to normal aging for women. Commensurate with this, from the review on the origins of variability (section 4.2.2), it emerged more strongly that female gender membership might operate as a protective factor, and conversely that male gendership would manifest as a vulnerability factor, with respect to the lowering of the symptom threshold (see section 4.2.2.3). In terms of BRC theory, the mechanisms of early versus late onset of variability effects for males versus females on cognitive test performance would operate in the same manner as described above for low versus high functioning populations. In a population with the risk factor of male gender membership, a significant portion of male individuals are likely to cross the brain reserve threshold sooner than would occur for individuals from a population with the protective factor of female gender membership. Consequently such vulnerable male individuals would fail to sustain a hold-pattern on cognitive abilities in association with the yet additional risk factor of neural aging (as with any compromised neural state) for as long as would occur for more protected female individuals.

5.2.1.2 Domain/task-specific effects

As discussed extensively in Chapter 4 (section 4.2.1), it appears that variability effects will also vary as a consequence of domain/task related effects. For demographically comparable samples, it appears that more age-sensitive tasks as identified on the basis of central trends analysis (which are those tasks which call upon fluid functions to a greater extent than crystallized functions, perceptual speed more than reasoning processes, and in turn both perceptual speed and reasoning more than relatively pure verbal abilities)[13], will result in the more severe overall inverted-U (differentiation/subsequent dedifferentiation) variability effect in contrast to the linear (sustained differentiation) effect. Further it appears that with comparable samples, more challenging levels of the same task will cause the earlier onset of increased variability, and the greater likelihood of the more severe overall differentiation/subsequent dedifferentiation effect, rather than the sustained differentiation effect.

---

[13] This hierarchy of age-sensitivity of functional domains is based on the discussions in Chapter 3 (section 3.2.3.1 and 3.2.3.2) which invoke the cascade model of Birren and Cunningham (1985) to encapsulate the direction of aging effects. In this model it is proposed that in descending order, the three cognitive constructs of 'verbal comprehension', 'reasoning', and 'perceptual speed' are *increasingly* differentially sensitive to aging. The discussion in Chapter 4 (section 4.2.1.2) on the proposed differences between the Digit Symbol and Trail Making Tests (in terms of the functions called upon to conduct these tasks) serves to emphasise that the contribution of different functional aspects to cognitive tasks is not an all-or-none phenomenon and may be extremely subtle. It is a matter of the *relative* contribution of functional components that is of relevance in the comparisons here. When tasks are viewed in relation to each other they may be conceptualised as containing 'more' or 'less' of a particular combination of Birren and Cunningham's three broad cognitive constructs (verbal comprehension, reasoning and perceptual speed), and in turn this is likely to have differential influences on the age-sensitivity of the tasks. Hence these proposed variability effects are being purposefully phrased in relative and not absolute terms.

Again these events are explicable in terms of BRC theory on the assumption that aging is conceptualized as a form of neural attrition which constitutes a risk factor for lowering the threshold of symptom presentation. Thus in association with other factors which contribute to lowering or raising the brain reserve capacity threshold within an individual, the progression of age is likely to produce differential effects. Such risk factors (as noted in the above subsection) might reside *within* the individual (for example differential levels of education, IQ or gender), and produce differential effects on cognitive task performance in interaction with aging. These factors can usefully be conceptualized as *internal risk factors*. However, bearing in mind that aging in itself is an *internal* risk factor which contributes to lowered levels of brain reserve capacity within individuals, there will also be differential effects of age on cognition function for individuals in interaction with *external risk factors*. Such external factors might be the amount of stress proffered by particular task-types, and the associated extent to which task-types *in interaction with aging* thereby serve to lower brain reserve threshold. Thus a task which presents strong challenge will produce increased risk of symptom onset in a situation of older age and consequent lower brain reserve capacity, than the same task presented in the protective situation of younger age and consequent greater brain reserve capacity. *Hence the aggregation of high challenge and lowered brain reserve capacity threshold due to older age would enhance the risk of symptom onset; conversely the combination of high challenge and raised brain reserve capacity threshold due to younger age would reduce the risk of symptom onset.*

Implicit in the concept of age-sensitivity of task (as established on central trends analyses), is the fact that a more age-sensitive task presents greater challenge to the aging brain than a less age-sensitive task. Thus, consistent with the above BRC understanding of events, it has been identified that, given a comparable cohort, a sustained increased variability effect is observed for tasks with a greater verbal or reasoning component (that is relatively less age-sensitive tasks), whereas the more severe overall non-linear inverted-U effect of initial increases in variability followed by sharp, rapid decreases in variability is revealed for tasks calling upon relatively pure visuo-perceptual speeded functions (that is more age-sensitive tasks) (see discussions in Chapter 4 with respect to data from Schaie, 1983, Figure 4.3, p.96; and Shuttleworth-Jordan and associates, Figures 4.8a and 4.8b, p.115). It has also been identified that the onset of variability effects will be earlier on the same cohort for a more challenging level of the same task (see discussions in Chapter 4 on the differential presentation of Trail Making Test A and B, illustrated in Figure 4.8a, p.115). Specifically from these examples it appears that due to reduced brain reserve capacity as the neural aging effect advances, a significant proportion of individuals fail to sustain a hold-pattern on cognitive abilities with more age-sensitive tasks (or with relatively higher levels of challenge within a particular task) for as long as occurs with less age-sensitive tasks (or with relatively lower levels of challenge within a particular task). Hence the full differentiation/subsequent dedifferentiation effect on more age sensitive fluid/perceptual tasks (or higher levels of challenge within a particular task) occurs sooner than for the less age-sensitive crystallized/verbal/reasoning tasks (or higher level of challenge within a particular task). These generalized domain-specific/task effects are schematically illustrated in Figure 5.3 below.

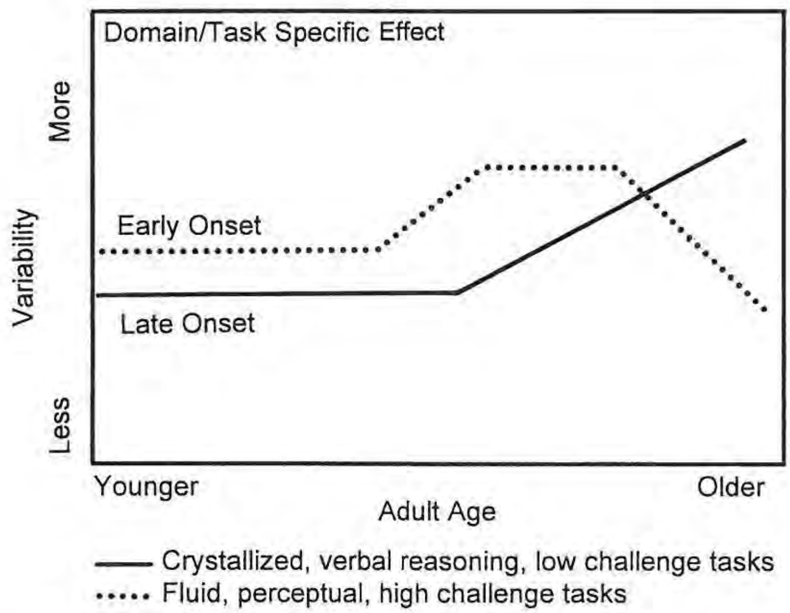


Figure 5.3. Schematic representation of the proposed contrasting domain/task specific effects: Tasks with high fluid/perceptual components and/or high challenge result in relatively early onset of increased variability and a subsequent decrease in variability; in contrast tasks with high crystallised components and/or low challenge result in relatively later onset of sustained increased variability.

On the basis of all the above empirical indications that have emerged from the literature review, about the characteristics of inter-individual variability in cognitive aging, it is now possible to go further and derive an integrated model of the causal mechanisms involved in the demonstration of these differential effects.

5.3 SHUTTLE MODEL OF INTER-INDIVIDUAL VARIABILITY

The proposed model comprises two key features which are well-described by the metaphor of a shuttle, used in its meaning both as an *object* (which describes the *form* of the variability curve), and as a *mechanism* (which describes the *operation* of the causes and effects of the variability patterns). Hence the model is termed the 'shuttle' model of inter-individual variability in cognitive aging and is delineated in terms of two central aspects as follows.

5.3.1 The shuttle shape of the variability curve

Up to this point all the graphical material presented in this chapter, has described the age-related variability effects in a single dimension, which is the unit of variance measurement. However, the variance measure actually reflects *the greater or lesser extent to which scores vary around the mean*, and hence can be described in two-dimensional form as Wohlwill(1973) has done (Figure 4.2, p.86), thereby illustrating the *magnitude* of the variation in scores. Thus it is possible to turn the overall inverted-U variability curve across the adult years as represented in unidimensional form (Figure 5.2, p.153), into a two-dimensional shuttle-shaped figure as illustrated in Figure 5.4 below.[14]

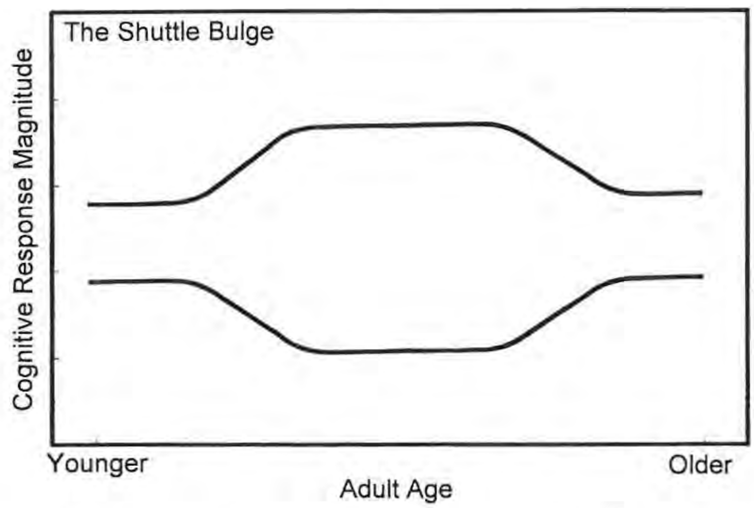


Figure 5.4. The Shuttle Model of Inter-individual Variability: Schematic representation of the proposed "bulge" in cognitive variability in association with adult aging, in the absence of any assumptions about alterations in the mean.

---

[14] The overall *shape* of the two-dimensional representation of the variability bulge in association with adult aging is comparable to the shape of a shuttle as per the following definitions: a weaving instrument with two pointed ends by which thread is carried (Sykes, 1982); a boat shaped weaving-implement on which weft thread is shot across between warp-threads (Fowler & Fowler, 1959).

In the absence of any assumptions about increases or decreases in average effects, this figure is a visual schema which encapsulates the increasing bulge in the distribution of scores that appears to occur regularly in association with normal aging, and the subsequent disappearance of the bulge which frequently occurs in the later stages of the aging process. *In this form the shuttle bulge is in effect a schematic mirror image of the overall non-linear inverted-U effect as illustrated in Figure 5.2, (p.153), which in turn was modelled on data presented in single dimensional form from Schaie and Salthouse (contained in Figures 4.3 and 4.4, pp.96 and 100 respectively), and Bode & Shuttleworth Jordan (Figure 4.8b, p.115).*

The bulge is given the term 'shuttle' not only because of its shape, but in that it is seen to reside on the age axis in a manner which is not fixed, and shifts backwards and forwards depending on a series of influences to be described in the next subsection. Further, however, it is tentatively envisaged that the *shape* of the shuttle itself may vary in that, depending on different circumstances, it may become more elongated, or fatter or very flattened such that it almost does not occur. The disappearance of the bulge, for instance, due to the aggregated effect of a low functioning cohort and particularly age-sensitive tasks was postulated in Chapter 4 (see section 4.2.1.2), as a possible explanation for the absence of any increase in variability for fluid functions which occurred on McArdle's study as reported in Horn and Hofer (1992).

Thus the model provides the flexibility for describing alterations in the shape of the variability bulge depending on different influences. However, the state of investigation into variability up to now has been at the stage of deciding whether or not increased variability occurs at all in association with aging, and has been *far* from the point of such fine analysis as the description of the *size* of a particular variability curve due to various influences. There is nothing identifiable in the literature in this regard, and hence the review was purely focused on the *progression of the curve* across the age ranges, rather than the *size* of the curve in and of itself in more absolute terms, which opens up another whole issue which could serve as a separate source of investigation. At the present stage of the development of this model, therefore, it is merely postulated that the propensity for the occurrence of a variability bulge *exists* in association with advancing age, and that this reflects the presence of a significant number of individuals with disproportionately low and/or high scores who are not well-represented by the central trend.

5.3.2 The shuttle mechanism of the variability curve

It has become apparent from the literature review that the variability bulge that occurs in association with aging as described above, *is predictable in the manner in which it presents as a whole in relation to the chronological age axis.* From this mechanical point of view, the bulge can be conceived of as shuttling backwards and forwards along the adult age axis depending on the constitution of the cohort, and the nature of the cognitive task. This occurrence can fundamentally be understood as a *shifting threshold effect* (as delineated in terms of the higher order Satz BRC theory within which the model is constituted), in that it is proposed that the protective factors of a high level of education and IQ (with associated high socio-economic status and better health), less age-sensitive and less challenging tasks, and female gender, in isolation or in aggregated form, will *raise the threshold* of symptom presentation, the variability bulge will take effect at a *later age stage*, and hence its position will *shuttle to the right* along the age axis; alternatively, the vulnerability factors of low education and IQ (with associated low socio-economic level and high levels of incipient disease), more age-sensitive and more challenging tasks, and male gender, in isolation or in aggregated form, will *lower the threshold* of symptom presentation such that the variability bulge will emerge at an *earlier age stage*, and hence its position will *shuttle to the left* along the age axis. These proposed earlier and later age-stage shuttle effects in the emergence of the variability curve, are illustrated in Figure 5.5 below (p.162).

The left shuttle effect will be more likely to reveal the full curvilinear variability pattern of an initial increase in variability followed by a subsequent decrease (Figure 5.5, upper graph), whereas the right shuttle effect will be more likely to reveal only the linear pattern of a sustained increase in variability (Figure 5.5, lower graph). This is not invariable, however, in that depending on the particular constellation of conditions, it is conceivable that two full variability curves may occur with the one shifted more to the left of the other, or two linear effects may be in evidence with the one shifted more to the left of the other. Thus on the basis of this model, it is possible conceptually to integrate differences in the patterns of variability data due to sample and task-related effects. Further it is feasible to integrate apparently contradictory findings across studies for the same task, in that for example, in instances where the bulge shifts far to the left (Figure 5.5, upper graph), no increases in variability may be detected in a study which only examines the 70 to 90 year old age groups, or the 20 versus the 80 year old age groups. If the curve shifts far to the right for any of the

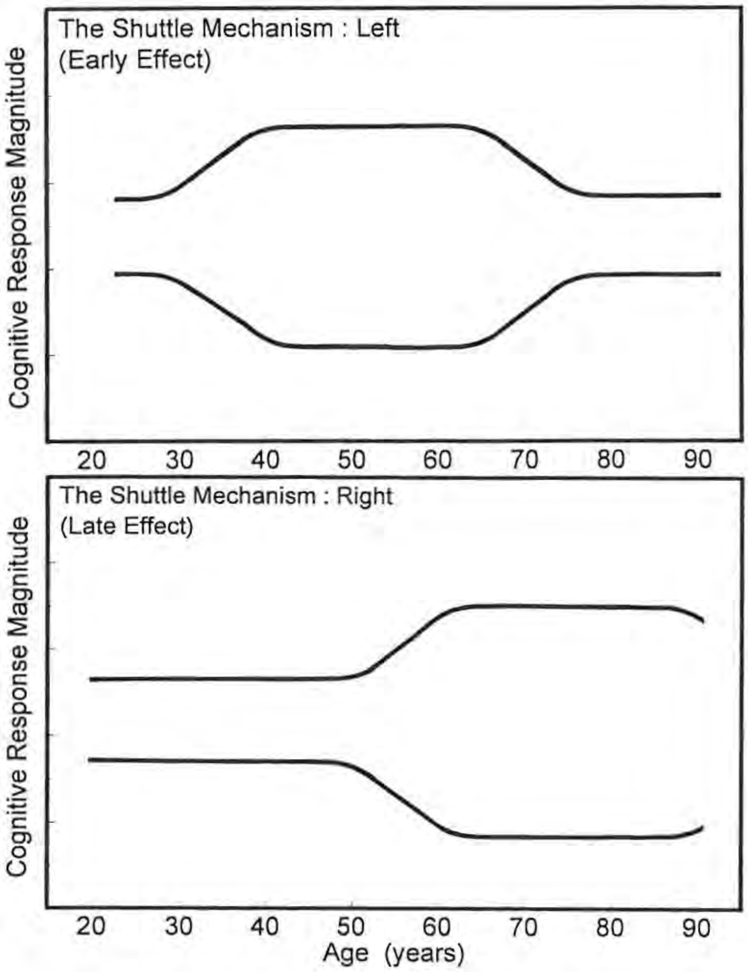


Figure 5.5. The Shuttle Mechanism: Schematic representation of the proposed left versus right shuttle effect of the variability bulge in relation to the adult age axis due to differential influences of education/IQ, task challenge, age sensitivity of tasks, and gender.

hypothesized reasons (Figure 5.5 above, lower graph), only the initial part of the overall bulge may be apparent in a study which stops with the 70 or 80 year old age stages, in the form of increased variability and no subsequent decline in variability.

Finally, it is postulated that regardless of whether the variability curve is shifted more to the left or the right on the aging axis, it will occur in association with the onset of steadily declining average cognitive performance across the younger to older adult years. This combined effect of declining average performance (represented for the instance where poor cognitive performance is equated with low scores) and corresponding standard deviations, for two different locations of the shuttle bulge, is schematically represented in Figure 5.6 below.

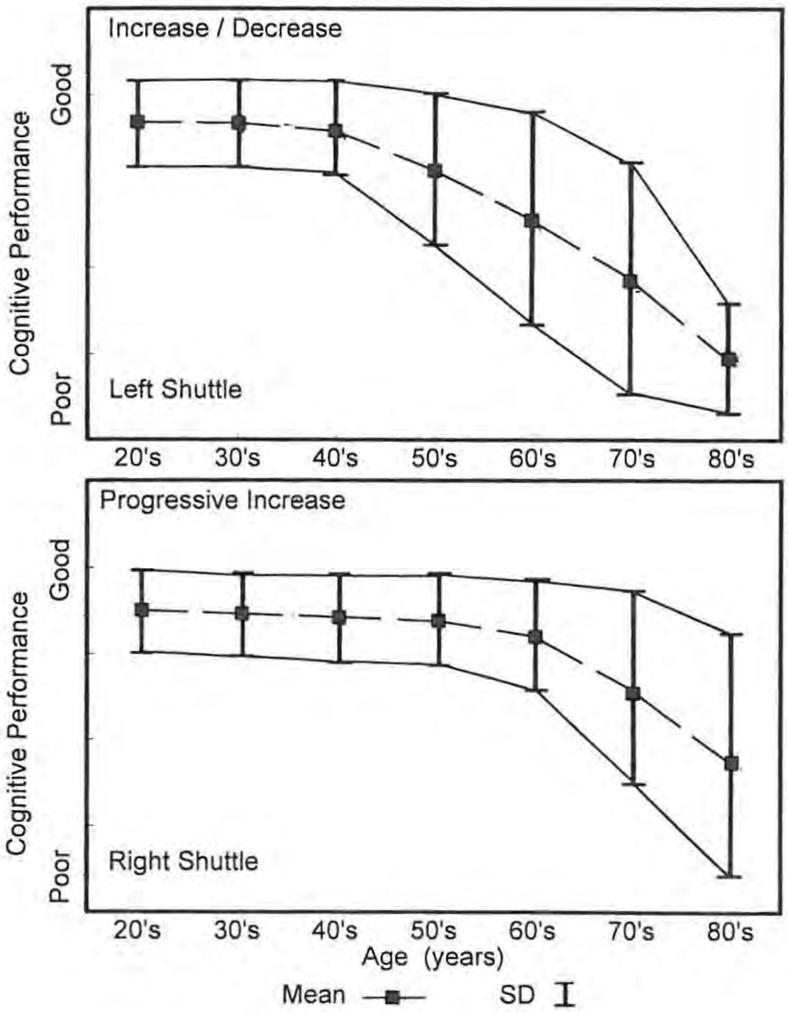


Figure 5.6. The Shuttle Effect: Schematic representation of declining Means and corresponding Standard Deviations for two proposed patterns of cognitive response magnitude: increase/subsequent decrease which occurs where there is early symptom onset and the bulge has shuttled to the left (upper graph); progressive increase which occurs where there is late symptom onset and the bulge has shuttled to the right (lower graph).

The earlier schemas of Figures 5.4 and 5.5 (pp.159 and 162) provide a mirror image representation of variability measurement *in itself* in association with age (in the absence of any assumptions about the slant of a particular central tendency), and thereby allow the opportunity for highlighting the shape and mechanism of the variability effect in an uncluttered manner. On the other hand the mode of illustration in Figure 5.6 (p.163) provides the opportunity to represent the shuttle bulge and shuttle shift mechanism as they occur more precisely in correspondence with the central tendency.

Thus in sum, the *same* variability trend can be illustrated in three different ways: (i) the unidimensional form of the variability data in themselves as plotted against the age axis (as, for example, in Figures 4.3, p.96; 4.4, p.100; and 5.2, p.153); (ii) the two-dimensional mirror image of variability data in themselves as they occur in association with age (as in Figures 5.4 and 5.5, pp. 159 and 162); (iii) variability data plotted in correspondence with mean scores (as in Figure 5.6, p.163).

5.3.3 Uncertainty in the model

Cross-sectional aging studies on which the bulk of the review was based, have different end-points in terms of the oldest age group assessed ranging from the 70's through to the late 80's or early 90's. None of the studies reviewed focus on the extremely old age periods of the ninth and tenth decades for which there is minimal research to date. Hence the end points of the model for all possible influences which cause the bulge to shift on the age axis, are uncertain at this stage. Furthermore it can be seen that unless the increase in variability has moved to the extreme right, bimodal comparisons between variability in the younger age groups and in the older old age groups may erroneously give rise to the impression that there has been no significant increase in variability in association with aging.

What does appear certain is that an initial bulge in variability occurs in association with the onset of cognitive aging (differentiation effect), and that this is frequently followed by a subsequent reduction in the variability bulge (dedifferentiation effect). What remains uncertain is whether a significant dedifferentiation effect will *always* occur at some point in the very late older years irrespective of the educational/IQ level or gender of the sample, or nature of the cognitive task. From the review it was noted that for some cognitive tasks only a progressive increase in variability was in evidence without a subsequent decrease. One possibility in such instances, which is implied in the shuttle model as outlined above, is that given the ability to document performance on progressively higher age groups, a decrease in variability would eventually become apparent in association with further declining mean scores, thus revealing the full inverted-U variability curve. However, it is also possible that there would be instances in which a progressive increase in variability, *without a subsequent decrease*, would occur into the extremely late adult years. In other words it may be that given certain circumstances a dedifferentiation effect would *never* occur however old the population tapped. Clearly such a *prolonged* occurrence of high variability would be most likely for those influences, or combination of influences, that contribute significant protective factors in terms of BRC and cause the bulge to shift to the right (Figure 5.5, p.162, lower graph; and Figure 5.6, p.163, lower graph).

5.3.4 Hypotheses for the present study

Building on the BRC theory hypotheses that were posed in Chapter 3 (reiterated at the beginning of this chapter), together with the more specific indications that have arisen from the literature review on variability in Chapter 4, it is now possible to elaborate on these with more specificity in terms of the 'shuttle' model of variability which has been developed above. Thus the following hypotheses are posed for testing in the present study:

1. There is likely to be increased variability of scores on cognitive tasks in association with the initial stages of the cognitive aging process as identified on central trends analysis, followed by a subsequent decrease in variability as the aging process progresses, producing an overall *non-linear pattern* of an initial increase and subsequent decrease in variability in association with aging;

2. The presentation of the non-linear pattern of an initial increase and subsequent decrease in inter-individual variability in cognitive test performance in association with aging, will be differentially affected by the cohort effects of lower versus higher levels of *education/IQ* and *male versus female gender*, and the task type effects of lower versus higher levels of *task challenge*, and lower versus higher levels of *age-sensitivity of tasks*. Specifically lower levels of education/IQ, male gender, higher levels of task challenge, and greater age-sensitivity of tasks, will cause the variability curve to shuttle to the left, that is the onset of increased variability and subsequent decrease in variability will occur *earlier* in the adult age range;

conversely higher levels of education/IQ, female gender, lower levels of task challenge, and lesser age-sensitivity of tasks will cause the variability curve to shuttle to the right, that is the onset of increased variability and subsequent decrease in variability will occur *later* across the adult age range. (Although it was not possible on the basis of Satz's review to indicate definitive directions for gender effects, this refinement was made possible on the basis of the subsequent review on the origins of inter-individual variability - see section 4.2.2.3, and the arguments posed above under cohort-specific effects, section 5.2.1.1).

3. In some instances higher levels of education/IQ, and/or female gender, and/or lower levels of task challenge and lesser age-sensitivity of tasks may reveal a *progressive increase* in variability across the adult age range without any subsequent decrease being in evidence. This may imply that a subsequent decrease in variability would become apparent to reveal the full non-linear variability pattern of an initial increase followed by a subsequent decrease, were progressively older age groups assessed. On the other hand it may imply that a prolonged retention of increased variability would become apparent, without a subsequent decrease in variability, were much older age groups assessed.

4. In terms of hypotheses two and three above, it is hypothesised that when making comparisons of two variability effects (for the differential influences of education, IQ, task-type and gender), variability effects may occur in any one of the following combination of forms:

(i) Two full non-linear variability patterns of an initial increase and subsequent decrease may be in evidence across the adult age axis with the one for lower levels of education/IQ, male gender, higher levels of task challenge, and higher age-sensitivity of tasks shifted comparatively more to the left (that is the initial increase and subsequent decrease in variability will occur earlier), than the one for higher levels of education/IQ, female gender, lower levels of task challenge and lesser age-sensitivity of tasks which will be shifted comparatively more to the right (that is the initial increase and subsequent decrease in variability will occur later);

(ii) One full non-linear variability pattern of an initial increase and subsequent decrease may be in evidence for lower levels of education/IQ, male gender, higher levels of task challenge, or higher age-sensitivity of tasks which will be shifted to the left (that is occur earlier), together with only a progressive increase in variability in evidence for higher levels of education/IQ, female gender, lower levels of task challenge, or lesser age-sensitivity of tasks respectively, which will be shifted comparatively more to the right (that is the initial increase in variability will occur later and there will be no subsequent decrease in evidence);

(iii) No full non-linear pattern of an inital increase and subsequent decrease in variability may be in evidence at all across the adult age axis, but only two progressive increases in variability with the one for lower levels of education/IQ, higher levels of task challenge, higher age-sensitivity of tasks, or male gender shifted comparatively more to the left (that is the progressive increase in variability will occur earlier), and the one for higher levels of education/IQ, lower levels of task challenge, lesser age-sensitivity of tasks, or female gender, respectively, which will be shifted comparatively more to the right (that is the progressive increase in variability will occur later).

5. The above presentations of variability in association with the progression of adult age will occur in the presence of an *overall steady decline in average cognitive test performance* as reflected in mean scores.

6. Since differential levels of education, IQ and gender are likely to be intricately related to differential variability effects in association with aging depending upon the extent to which they serve a protective function, it is hypothesized that the effects of education, IQ and gender on the individual distribution of scores will not be uniform across age groups, and that this will be reflected in *differential characterization of the disproportionately low and high scorers (that is the outliers)* as follows:

(i) The effect of education, IQ and gender will be more pronounced for the age groups with significantly increased variability than for those age groups which do not show increased variability;

(ii) Where education, IQ and gender are having an effect higher education/IQ and female gender will characterize the disproportionately high scorers, and lower education/IQ and male gender will characterize the disproportionately low scorers.

5.3.5 Shuttle model of inter-individual variability: conclusions

The threshold 'shuttle' model of inter-individual variability has been developed on the basis of an analytical synthesis of the research available in the literature within the broader theoretical context of the Satz BRC threshold concepts. In the process of building the model a reciprocal process occurred as follows. Trends that appeared to be emerging out of an apparently disparate and patternless assortment of data were identified, and then through the eyeglasses of an emergent model, research findings were analyzed. Such a lens has the advantage of allowing the viewer to identify patterns that might otherwise not be apparent, and facilitates the ability to see the wood and not get lost in a disparate plethora of trees. On the other hand, the danger of model or paradigm building lies in the fact that this will encourage an attempt to explain observational data in a manner which confirms the model, and an erroneous model might blind the researcher to the true nature of events. However, as Salthouse (1991) has emphasized, the need for theory building is overdue in cognitive aging research. Further Kuhn (1970) has suggested that science advances on the basis of the refutation of theories. Thus this model and its associated set of hypotheses is presented as a vehicle for advancing knowledge on the nature of inter-individual variability in cognitive aging, via further testing, possible verification, modification and conceivably even refutation of its basic tenets.

5.4 CHAPTER SUMMARY

This chapter served to identify a missing link in the hierarchy of theoretical discourse with respect to the scientific investigation of cognitive variability in association with aging, in that there was no specific model by which to describe and predict its particular characteristics. Thus a model of the causal connections for adult age-related variability patterns was presented, which was derived on an empirical basis from the research review, and located within the broader rubric of a neuropsychological framework and the threshold postulates of the Satz BRC theory. From this model, a number of hypotheses for further testing in the present study were elaborated.

VOLUME TWO

AN EMPIRICAL INVESTIGATION INTO VARIABILITY AND EVALUATION OF THE MODEL

6

CHAPTER 6: METHOD

In this chapter the research problem and specific research aims will be delineated as they have arisen out of the overall literature review. Following on this the procedural aspects of the empirical investigation will be described.

6.1 DELINEATION OF THE STUDY

In the broadest terms the present research set out to examine the extent and nature of heterogeneity in cognitive aging. On theoretical and research grounds the literature review provided support for the presence of enhanced inter-individual variability in association with aging. Furthermore a number of specific trends were identified on the basis of which a model of variability was developed and a number of specific hypotheses were posed for further investigation as follows (see section 5.3.4):

1. There is likely to be increased variability of scores on cognitive tasks in association with the initial stages of the cognitive aging process as identified on central trends analysis, followed by a subsequent decrease in variability as the aging process progresses, producing an overall *non-linear pattern* of an initial increase and subsequent decrease in variability in association with aging;

2. The presentation of the non-linear pattern of an initial increase and subsequent decrease in inter-individual variability in cognitive test performance in association with aging, will be differentially affected by the cohort effects of lower versus higher levels of *education/IQ* and *male versus female gender*, and the task type effects of lower versus higher levels of *task challenge,* and lower versus higher levels of *age-sensitivity of tasks*. Specifically lower levels of education/IQ, male gender, higher levels of task challenge, and greater age-sensitivity of tasks, will cause the variability curve to shuttle to the left, that is the onset of increased variability and subsequent decrease in variability will occur *earlier* in the adult age range; conversely higher levels of education/IQ, female gender, lower levels of task challenge, and lesser age-sensitivity of tasks will cause the variability curve to shuttle to the right, that is the onset of increased variability and subsequent decrease in variability will occur *later* across the adult age range. (Although it was not possible on the basis of Satz's review to indicate definitive directions for gender effects, this refinement was made possible on the basis of the subsequent review on the origins of inter-individual variability - see section 4.2.2.3, and the arguments posed above under cohort-specific effects, section 5.2.1.1).

3. In some instances higher levels of education/IQ, and/or female gender, and/or lower levels of task challenge and lesser age-sensitivity of tasks may reveal a *progressive increase* in variability across the adult age range without any subsequent decrease being in evidence. This may imply that a subsequent decrease in variability would become apparent to reveal the full non-linear variability pattern of an initial increase followed by a subsequent decrease, were progressively older age groups assessed. On the other hand it may imply that a prolonged retention of increased variability would become apparent, without a subsequent decrease in variability, were much older age groups assessed.

4. In terms of hypotheses two and three above, it is hypothesised that when making comparisons of two variability effects (for the differential influences of education, IQ, task-type and gender), variability effects may occur in any one of the following combination of forms:

(i) Two full non-linear variability patterns of an initial increase and subsequent decrease may be in evidence across the adult age axis with the one for lower levels of education/IQ, male gender, higher levels of task challenge, and higher age-sensitivity of tasks shifted comparatively more to the left (that is the initial increase and subsequent decrease in variability will occur earlier), than the one for higher levels of education/IQ, female gender, lower levels of task challenge and lesser age-sensitivity of tasks which will be shifted comparatively more to the right (that is the initial increase and subsequent decrease in variability will occur later);

(ii) One full non-linear variability pattern of an initial increase and subsequent decrease may be in evidence for lower levels of education/IQ, male gender, higher levels of task challenge, or higher age-sensitivity of tasks which will be shifted to the left (that is occur earlier), together with only a progressive increase in variability in evidence for higher levels of education/IQ, female gender, lower levels of task challenge, or lesser age-sensitivity of tasks respectively, which will be shifted comparatively more to the right (that is the initial increase in variability will occur later and there will be no subsequent decrease in evidence);

(iii) No full non-linear pattern of an initial increase and subsequent decrease in variability may be in evidence at all across the adult age axis, but only two progressive increases in variability with the one for lower levels of education/IQ, higher levels of task challenge, higher age-sensitivity of tasks, or male gender shifted comparatively more to the left (that is the progressive increase in variability will occur earlier), and the one for higher levels of education/IQ, lower levels of task challenge, lesser age-sensitivity of tasks, or female gender, respectively, which will be shifted comparatively more to the right (that is the progressive increase in variability will occur later).

5. The above presentations of variability in association with the progression of adult age will occur in the presence of an *overall steady decline in average cognitive test performance* as reflected in mean scores.

6. Since differential levels of education, IQ and gender are likely to be intricately related to differential variability effects in association with aging depending upon the extent to which they serve a protective function, it is hypothesized that the effects of education, IQ and gender on the individual distribution of scores will not be uniform across age groups, and that this will be reflected in *differential characterization of the disproportionately low and high scorers (that is the outliers)* as follows:

(i) The effect of education, IQ and gender will be more pronounced for the age groups with significantly increased variability than for those age groups which do not show increased variability;

(ii) Where education, IQ and gender are having an effect higher education/IQ and female gender will characterize the disproportionately high scorers, and lower education/IQ and male gender will characterize the disproportionately low scorers.

The first five of these hypotheses refer to *inter-individual variability trends as they occur between age groups,* and have been derived on the basis of tendencies identified from an examination of empirical data presented in the literature. However, whilst the review was comprehensive, the available research on variability was limited in that frequently observations were made without formal statistical analysis, and research investigations have tended to be unsystematic and consequently there are significant gaps in the findings across functional modalities. Furthermore there has been no methodical investigation of significant changes in mean scores and their complementary association with changes in variability within studies. The only two systematic and relatively comprehensive investigations into group variability changes appear to be those Morse (1993) and Christensen et al. (1994), in which the variability across a number of cognitive modalities have been examined by means of statistical analysis. However in both instances these authors findings are limited due to the use of severely restricted age ranges: Morse's large meta-analysis was concerned with a comparison of fluid versus crystallized functions across the bimodal young versus old dimension only; Christensen et al.'s study, whilst investigating a broad range of functions, was restricted to the 70 to 90 year old age range. Both of these studies investigate 'fluid' and 'crystallised' abilities, the composite nature of which may serve to obfuscate effects. Thus it is apparent from the literature review that there is a need for a study which systematically investigates inter-individual variability effects, in conjunction with central trends effects, across the total adult age range (in order to avoid confounding study window effects), for a representative set of functional domains in relatively non-composite form (in order to avoid confounding modality effects). This is a research endeavour which has not been achieved up to now.

The sixth and final hypothesis delineated in Chapter 5 (section 5.3.4) relates to *inter-individual characteristics as they occur within different age group distributions,* and has been derived purely speculatively in terms of the Satz (1993) BRC theory, and on the basis of tendencies identified from an examination of empirical data presented in the literature for between-group analyses. Rabbitt (1993c), appears to be the only researcher who has studied the characteristics of outliers on cognitive tasks *within* older age groups in order to gain clarity on changes across different age groups, but he has focused on the dissociation between memory impairment and IQ, and has not specifically examined the characteristics of distribution patterns in terms of education, IQ or gender within-groups. Thus there is a need for a study which systematically investigates the nature of the distribution of scores within age groups with a focus on unravelling the constitution of outliers in the manner which has been conducted by Rabbitt (1993c), except with the specific purpose of identifying individual distribution patterns for different age groups in terms of education, IQ and gender in a task-specific manner.

Thus it has emerged that are two very clear parameters of investigation needed to conduct an adequate enquiry into variability in cognitive aging. These are (i) a between-groups analysis of variability patterns, and (ii) a within-groups analysis of individual distribution patterns, and to date neither form of investigation has been adequately conducted. Discussion by Wohlwill (1973) reinforces the notion that there are these two distinct parameters required in order to study the differential aspect of development. As noted in Chapter 4 (see section 4.1.2), Wohlwill (1973) delineates these as follows: (i) the need to describe the *characteristics and course* of inter-individual differences (which is only possible via between-groups analyses), and (ii) the need to address the *nature and origin* of developmental individual differences (which calls for within-groups analyses). Thus, in accordance with these directions for research into variability it was decided to conduct a two-pronged study into inter-individual variability in cognitive aging in order to test the hypotheses which were posed in Chapter 5 and outlined again above. It was considered that such a dual focused study would provide the scope both to describe the course of inter-individual variability (by means of between-groups analysis), and at the same time address its origins more fully (via within-groups analysis).

The two phases of the study were as follows:

Phase 1. A wide-ranging analysis of between-groups trends in inter-individual variability across the adult years including data sets for a comprehensive spectrum of functional modalities and a broad spectrum of adult age groups ranging from 18 through to 90 years (*hereafter referred to as Phase 1 of the study*). This phase was implemented in order to test the first five hypotheses delineated above.

Phase 2. A more narrowly focused between-groups analysis of variability trends for data sets on a few selected neuropsychological tests, followed by a within-groups analysis of the individual distribution of raw scores for these same data sets, with a particular focus on the characteristics of good and poor outliers in terms of education, IQ and gender (*hereafter* *referred to as Phase 2 of the study*). This phase was implemented in order to test the first five hypotheses delineated above (via the between-groups analysis), and in particular the sixth and final hypothesis via the within-groups analysis.

6.2 RESEARCH PROCEDURE

6.2.1 Overall research procedure: A meta-analysis of variability data

The overall research procedure adopted comprised a form of meta-analysis in that both phases of the research examined multiple data sets that are available in the literature. For the purposes of this thesis a 'data set' is defined as a single *pre-existing* collection of normative data across a series of adult age groups for a particular neuropsychological test, which may include separate data collections for submodalities within a particular neuropsychological test. Thus in one aging *study* (for example, van Gorp et al., 1990) there may be several '*data sets*', including a data set for the Rey-Osterrieth Complex Figure Recall (with data for both the copy and recall versions of the test), a data set for the Trail Making Test (including data for Parts A and B), and a data set for the Wechsler Memory Scale-Logical Memory (including data for immediate and delayed versions of the test). Further, there may be several different data sets for a single test gleaned from disparate research studies and/or seminal normative data collations. Thus, for example, for the Wechsler Memory Scale-Logical Memory immediate, three data sets are used, one each as compiled in Lezak (1983) and Spreen & Strauss (1991), and one from van Gorp et al. (1990).

6.2.1.1 Definition of the meta-analytic technique

Craik and Jennings (1992), have emphasized the importance of moving to a meta-level of analysis in order to study the patterns that emerge across many sets of data, stating that "the search for order and consistency must move to a level higher than the individual experiment" (p. 98). A particular type of meta-analytic technique is exemplified in the collaborative work of Mortimer, van Duijn, Chandra, Fratiglioni, Graves, Heyman et al., (1991) with respect to head injury as a risk factor for Alzheimer's disease. In this much cited work statistical analyses were conducted on *pooled* data from a collection of studies of Alzheimer's disease in which head trauma exposure were available in order to demonstrate the effects of this occurrence more robustly. Specifically with respect to research into inter-individual variability in aging, Morse (1993) has produced a pioneering example of a meta-analytic investigation in this area via her examination of all studies in the issues of the Journal of

Gerontology and Psychology and Aging (from 1986 to 1990) in which there was a comparison of at least one group of healthy older adult subjects with at least one group of younger adult subjects. In Morse's study statistical comparisons were conducted for the *pooled* variance measures of studies grouped in terms of reaction time measures, memory measures, fluid and crystallized intelligence test measures.

Whilst the meta-analytic method of pooling data from a number of studies is useful for the examination of phenomena which occur rarely, Mortimer et al. (1991) has emphasized that one of the difficulties of this pooling technique lies in the frequent lack of comparability between samples which frequently occurs due to demographic heterogeneity. In particular, the pooling of data across disparate studies for which the demographic characteristics differ would be a worthless technique if, as was the case of the present investigation, a *focal purpose* of the study was precisely to examine the influence of such demographic differences. It might of course be useful to isolate a single task and search for a series of studies of equivalent demographic criteria, and then pool all such data in order to confirm the presence of a highly specific variability effect more robustly. However, the purpose of the present investigation was to get a broad comparative view of variability effects across a spectrum of tasks and demographic influences, as a first necessary research step. The verification of highly specific effects more robustly would be a route for future research.

Hence *the type of meta-analysis adopted for this study did not incorporate any pooling technique*. The procedure adopted is described as a meta-analysis in that it undertakes a statistical reanalysis of a broad spectrum of existing data from the literature, but differs from the method used by Mortimer et al. (1991) and Morse (1993) in that *it conducts the statistical analyses on each set of test data presented in the literature separately, and does not make use of pooled data*. On this basis it was considered that broad patterns and trends across multiple studies could be identified in the same manner as was achieved in the literature review on variability patterns in Chapter 4. Moreover it was considered that the mode of separate analyses of data sets would have an important advantage over the method of pooling data. Whereas differential demographic characteristics in pooled data collections would be likely to *confound* the analysis, the ability to take account of distinct demographic details between data sets by keeping them separate could be used to *enhance* the analysis.

6.2.1.2 Data Sources

As noted in the Introduction (see section 1.3), the neuropsychological angle has been identified as a fruitful direction for future research in cognitive aging (Craik & Jennings, 1992; Rabbitt, 1993b). Taking up this challenge, the source of data for analysis in this study is derived from this perspective. In the field of clinical neuropsychology there has been a surge of normative data collected for a broad spectrum of commonly employed single neuropsychological tests, on the basis of cross-sectional studies across a wide range of adult age groups (for example, Cornfield & Shuttleworth-Jordan, 1996; Shuttleworth-Jordan & Bode, 1996; Shuttleworth-Jordan & Bode, 1995b; Van Gorp et al., 1990; and as collated and cited in Lezak, 1983 and 1995; Spreen & Strauss, 1991)[15]. Since the goal of these normative data collections has been to provide indications about average age trends for pragmatic clinical purposes, as a body of data they have received marginal attention from the perspective of conceptual implications for aging processes, and in particular minimal attention has been paid to differences in inter-individual variability across the age groups in these studies. However, it was considered that the wide range of age groups typically presented in these collections - for example most of the data in Spreen and Strauss are presented in decades from the 20's through to the 80's - would allow for a progressive and differentiated analysis of variability effects, and would not be subject to the study-window artifact of a narrow age range which has been the major limitation of previous investigations in this area (see discussion in Chapter 4 Figure 4.7, p.112).

Typically these normative data studies are cross-sectional in nature, and most are conducted on volunteer, largely non-institutionalized, community-based populations, comprising individuals who have been carefully screened to exclude medical or psychiatric ill-health. Their purpose was to identify normal differences between groups in association with the independent variable of chronological age. Thus in terms of the methodological arguments presented in Chapter 2 (see section 2.2.2), for a study of inter-individual variability in normal aging, these studies are considered to be highly suited for use as the data base for the present investigation in that: (i) cross-sectional studies are required in order to provide the necessary *current* perspective on age-related cognitive patterns upon which policy and clinical decisions should be based; further, cross-sectional studies are not subject to the powerful

[15] Once again attention is drawn to the fact that the most recent collation of neuropsychological data is that of D'Elia, Boone & Mitrushina (in press) which is referred to in Lezak (1995). However this book was not in print at the time of the present investigation, and thus was not available as a data source.

confounding effect of attrition of the non-elite with the progression of age, and thus, provided samples are well-matched between age groups on influential demographic variables such as education and IQ, may be less likely than longitudinal studies to obscure subtle patterns of inter-individual variability; (ii) the optimal target population in normal aging research is considered to be that of generally representative healthy individuals living in community-based or relatively independent residential situations; and (iii) chronological age serves as the ideal pragmatically accessible and robust independent variable against which to study the psychosocial patterns of variability which it throws into relief.

6.2.2 Study selection

6.2.2.1 Phase 1: Wide-ranging between-groups analysis of group data

As a framework within which to conduct the literature review of psychometric findings on average and variability age-related effects in Chapter 3 and Chapter 4, respectively, the following specific functional modalities were delineated (see section 3.2.3): *attention, memory, language skills, visual skills, problem solving, and motor activity*. The delineation of these modalities was based on the divisions adopted by Lezak (1983; 1995) and Spreen and Strauss (1991) in their seminal texts on neuropsychological tests and assessment, and further narrowed down to include those functions which have received most attention with regard to age-related changes as are incorporated in the overviews of Lovelace (1990) and Nolan and Blass (1992). The same framework was thus adopted for the purposes of Phase 1 of the present study. As an additional refinement, the modality of memory was separated into *verbal memory* and *visual memory*.

The selection criteria for the choice of neuropsychological normative data sets, were based on an attempt to avoid the artifactual differences in variability which can result from: (i) the use of an overly limited set of age groups for the comparisons; (ii) the analysis of an overly limited sector of the adult age range; (iii) the inclusion of comparative groups with vastly disproportionate numbers of subjects particularly in conjunction with exceedingly low numbers of subjects in any of the groups (see discussion on key artifacts to be avoided in variability studies, section 4.2.1.3). On this basis a number of inclusion criteria were delineated. Further, the choice of data sets to be included in the study set out to avoid bias in favour of the shuttle variability model and its hypotheses by including as a baseline *all* the normative data sets presented in three major normative data collations, with the exception of those which were excluded because they did not meet the delineated inclusion criteria.

Specifically the selection of data sets for the Phase 1 study proceeded as follows:

1. *General Inclusion Criteria.* In order for a set of normative data to be *included* in the meta-analysis it had to meet the following criteria:

(i) A data set had to be with reference to a cognitive task which fell into at least one of the functional domains of *attention, verbal memory, visual memory, language skills, visual skills, problem solving, and motor activity* as defined in Chapter 3 (section 3.2.3). Composite measures which might obscure specific task effects were specifically excluded such as general tests of intellectual function, as well as the combined form of the Digit Span test. Tests for which the range of questions and level of difficulty was exceptionally low thus ensuring a ceiling effect when used with a normal population were also excluded. On this basis the Personal and Current Information, Orientation and Mental control subtests for the Wechsler Memory Scale were excluded;

(ii) A data set had to be based on a sample comprised of medically and psychiatrically healthy, community-based or relatively independent-living adults. Since specifications about residential characteristics of the normative samples tended to be sparse, a data set was included as long as there was no positive specification that a data set was based entirely on hospitalized or institutionalized adults;

(iii) A data set had to consist of at least three adult age groups, and cover an overall adult age range of at least three decades. A data set was not included unless at least one age group was with reference to individuals of 60+ years.

(iv) A data set needed to include the delineation of the number of subjects in the age groups, and there had to be at least 10 subjects per age group;

(v) Normative data had to be available in the form of group means and standard deviations.

2. *Baseline Data Base.* As the starting point, normative data sets for every neuropsychological test that appears in the collations of Lezak (1983), Spreen and Strauss (1991), and van Gorp et al. (1990) which met the inclusion criteria (i) to (v) above, were incorporated for analysis. In most instances all the normative data documented for each test in these two texts which met the inclusion criteria were included. However, to keep the data within manageable limits, for the Paced Auditory Serial-Addition Test (PASAT) which comprises data for several gradations of presentation speed, and for the memory tests which comprise multiple learning trials and sub-modalities (viz., the Rey Auditory-Verbal Learning Test - RAVLT, the Benton Visual Retention Test-Revised - BVRT-R, and the Buschke Selective reminding Test -BSRT), only representative portions of the normative data were included. These included the highest and lowest levels of presentation, and any data for which there were contrasting trends between levels, such that the general indications of all aspects of the tests were adequately revealed.

3. *Additional Data.* Next a number of additional data sets were included. This was implemented to provide a basis for the comparison of trends across different data sets for tests already included in the baseline selection from the Spreen and Strauss (1991), Lezak (1983) and Van Gorp et al. (1990) collations. The additional extra sets were extracted from the following normal aging studies: Axelrod & Henry (1992); Beatty (1993); Mitrushina et al.(1991); Roman et al.(1991); Stuss et al.(1988). In order to expand the data base for the underrepresented functional modality of 'attention' additional data were incorporated from Shuttleworth-Jordan (1992) for the Digit Span Test (Forward and Backward), a test that was not part of the original baseline selection. All of these additional data sets met the inclusion criteria listed in (i) to (v) above with respect to sample type, number and age range of groups, and the manner in which the data were presented.

All the studies included in the Phase 1 analysis are listed in Tables A and B (Appendix pp.1-17, and 18-21, respectively). Tables A and B include (i) the origins of the various data sets, together with a summary of the data including numbers of subjects, means and standard deviations per age group (Table A); and (ii) the available information on the educational and IQ levels of the samples (Table B). All the data included in the study are with respect to neuropsychological tests which are generally well-known and commonly employed in clinical settings. Thus intricacies about the derivation, administration and scoring of the many tests focused on in Phase 1 of this study, which are comprehensively described within the sources from which the data sets were extracted, were specifically not reiterated in the thesis.

Detailed specifications about the residential status of the samples from which the data sets used in the present study were derived, were unavailable in several instances. However, it does appear that most studies were based on healthy non-hospitalized volunteers living in the community or in relatively independent retirement conditions. Without exception all of the studies specify that their samples were comprised of healthy individuals. Several researchers state in addition that their studies were based largely on samples of individuals who were living in the community or in independent retirement villages (Axelrod & Henry, 1992; Roman et al., 1991; Shuttleworth-Jordan & Bode, 1992; Stuss et al., 1987; van Gorp et al., 1990). All of the unpublished data of Strauss and Spreen (or Spreen and Strauss) reported in Spreen and Strauss (1991) were apparently derived on a relatively superior, well-educated population, which, although not specifically reported as such, was presumably a community-based population, and/or a population comprised of relatively independent-living individuals.

The only data sets included in the present study, which appear to have been sampled in part from a purely institutional setting, are those extracted from Lezak (1983) for the Wechsler Memory Scale Associate Learning, Logical Memory and Visual Reproduction subtests. The normative data Lezak presents for these subtests are a compilation of three different studies (Hulick, 1966; Klonoff and Kennedy, 1966; Wechsler, 1945) without any details provided about the constitution of the samples in the collection of studies. However, from Spreen and Strauss (1991) it appears that the data sets reported in Lezak (1983) were only in part derived from a community-based set of samples: The Wechsler (1945) norms for the 20-29 year-old age group, and the Klonoff and Kennedy (1966) norms for the 80's age group were based on non-hospitalized and community-dwelling subjects respectively, but the Hulick (1966) norms for the 30-39 and 60-79 year-old subjects were based on a nonneurological population of hospitalized veterans. Nevertheless, the full age-spectrum of Lezak's data sets for each of these three Wechsler Memory Scale subtests (including those subsets derived from nonneurological hospitalized subjects) were incorporated for analysis in the present study. This was justified on the grounds that the data sets (which were in *some* part based on community-dwelling samples), were representative of a relatively less well-educated population and thus could serve as an important comparative base for the other data sets which were derived on relatively well-educated populations.

6.2.2.2 Phase 2: Narrowly focused between-groups and within-groups analyses

For this phase of the analysis a *limited* number of data sets from a selection of the author and coworkers' normative studies for which raw scores were available (Cornfield & Shuttleworth-Jordan, 1996; Shuttleworth-Jordan & Bode 1995b; 1996), were included. Data from these studies incorporated the following commonly employed neuropsychological tests: The Trail Making Test Parts A and B (Reitan, 1979), the SAWAIS Digit Symbol subtest (South African Wechsler Adult Intelligence Scale Manual, 1969), Digit Symbol Incidental Recall Test (as described in Shuttleworth-Jordan & Bode, 1995b), and the Finger Tapping Test Preferred and Non-Preferred Hands (Denckla, 1973). As indicated in Chapter 4 (Section 4.2.1.2), the variability between age groups had been analyzed previously for the Trail Making Test and Digit Symbol subtest (Bode & Shuttleworth-Jordan, 1993; Cornfield & Shuttleworth-Jordan 1996), but in each case was only briefly commented on as a side issue in these articles which focused on the presentation of older adult average tendencies. Thus there was scope to subject this data (alongside the data for the Incidental Recall and Finger Tapping Tests) to an analysis with a specific focus on variability issues. Phase 2 of the analysis was dependent on the availability of raw score data bases, and was designed as a *preliminary* investigation into detailed between-groups and within-in groups effects on a spectrum of data which did not aim to be comprehensive in terms of the functional modalities or age ranges covered. These data sets all met the same criteria for inclusion as delineated for Phase 1 of the analysis as follows:

(i) Each data set was with reference to a cognitive task which fell into at least one of the functional domains of *attention, verbal memory, visual memory, language skills, visual skills, problem solving, and motor activity* as defined in Chapter 3 (see section 3.2.3). Specifically visual skills (Trail Making Test Parts A and B, SAWAIS Digit Symbol test), visual memory (Digit Symbol Incidental Recall), and hand motor activity (Finger Tapping Test Preferred and Non-Preferred Hands) were represented.

(ii) Each data set was based on a sample comprised of non-hospitalized, healthy, largely community-dwelling or relatively independent-living adults;

(iii) Each data set consisted of at least three adult age groups, covered an overall adult age range of at least three decades, and included at least one age group that was with reference to individuals of 60+ years.

(iv) Each data set had at least 10 subjects per age group.

In effect each data set had a minimum of 20 subjects per age group, and met further favourable criteria in respect of a study on variability (which were not possible to adhere to strictly in the more broad-ranging first phase of the study) of well-balanced numbers in each group, and matched educational levels across groups. The studies included in the Phase 2 analysis are listed in Tables C and D (see Appendix pp.22-23, and 24, respectively). The tables include the origins of the various data sets, together with a summary of the descriptive statistical data including numbers of subjects, means and standard deviations per age group (Table C); and the available information on the educational and IQ levels of the samples (Table D). Statistical comparisons of the educational and IQ levels between age groups appear in the results (see Chapter 8, Tables 8.1, 8.2, and 8.3, pp. 323-325). Without exception all of the data sets were originally derived from healthy, relatively well-educated samples of individuals living in the community or living independently in retirement villages.

6.2.2.3 Limitations of the data sources

Whilst there are distinct advantages in conducting a retrospective meta-level of analysis in order to study general tendencies that emerge across many sets of data which are available in the literature - a much larger data bank than could ever be collected by any single group of researchers - there are also a number of disadvantages. The data available in this manner for the purposes of the present study, were not derived with the specific *intent* of investigating variability effects, and hence do not always appear in ideal form for this particular research purpose. By means of the inclusion criteria laid down for both Phase 1 and Phase 2 of the study, an attempt was made to control for some of the major potential confounding influences in data sets (such as the exclusion of groups with exceedingly low numbers, and the inclusion criterion of a relatively wide age range for analysis). However, in both phases of the study, due to the retrospective manner of accumulation of the data, there were necessarily a number of limitations which could not be ruled out as follows.

For *Phase 1* of the study, the major limitations of the included data were: (i) disproportionate numbers of subjects across groups in many of the data sets, although the confounding influence of this was reduced by the elimination of data sets with age groups containing very low subject numbers; (ii) the demographic characteristics of the samples were not always clearly reported, and in most instances age groups were not specifically matched for level of education and/or IQ. However, there was a reasonable amount of demographic information available, especially for the total sample of each data set, and in several instances for each age group within the data set, and these demographic details have been carefully reported (see Table B, Appendix pp.18-21). Hence, since the data sets were not pooled but examined separately, it was possible to take account of available demographic characteristics in the analysis; (iii) not all the functional modalities were equally well-represented in that normative data sets were quite prolific within certain modalities for the three normative collections (see Lezak, 1983; Spreen & Strauss, 1991; and van Gorp et al. 1990) which formed the basis of the analysis, and not for others such as attention and motor activity. However it was possible to supplement both of the under-represented areas with two data sets from the author and co-workers' normative studies: Digit Span was included for the modality of attention in Phase 1 of the study, since the data set did not meet the more stringent inclusion criteria for Phase 2 of the analysis (see below) of relatively equal numbers of subjects between age groups, and at least 20 subjects per age group; Finger Tapping was included for the modality of motor activity since it met the Phase 2 inclusion criteria.

For *Phase 2* of the study, unlike for Phase 1, data sets were available across adult age groups that were carefully matched in terms of numbers of subjects and levels of education and IQ. However, there were a number of limitations as follows: (i) the data covered a small number of functional modalities. However, this was deemed adequate in terms of the preliminary stage of the investigation, and the necessary confines of the study considering the intricate set of analyses conducted on these data; (ii) the spread of age ranges within the Shuttleworth-Jordan and Bode (1995b) data were uneven in that the number of years in the 20-39 and 40-59 groups were larger than those in the 60-69, 70-79 and 80-89 year-old age groups. Since this could be a source of increased variability in the younger age groups, thus confounding the effect noticeable due to age, it was taken into account in the data analyses; (iii) the data for the Trail Making Test (Cornfield & Shuttleworth-Jordan, 1996) did not cover the whole adult age range and were restricted to the 60's, 70's and 80's age groups only. The later age ranges from 60 years onwards, however, are the years most likely to reveal marked aging effects and thus these data were considered worthy of inclusion in the study. Further the groups in the Trail Making Test study were well-matched for NART IQ as well as for education, thus enabling the examination of outliers in terms of both IQ and education effects; (iv) the data sets were all with reference to samples which were restricted to a relatively high level of education (no subjects were included with less than 10 years of education, and the majority had at least some University education), which would tend to reduce the extent of variability effects. However, as pointed out by Rabbitt (1993c), such elitist samples offer a measure of stringency to variability studies in that they may serve to reduce rather than exaggerate effects. Thus should effects be identified it is probable that these would be even more pronounced on a sample which was more generally representative.

6.2.3 Statistical procedures

6.2.3.1. Phase 1

The statistical procedures for the Phase 1 analysis were as follows:

1. Comparisons of standard deviations across all groups. For each data set of neuropsychological test scores included in the Phase 1 analysis (as listed in Table A, Appendix pp.1-17), Bartlett's test, was used to compare the standard deviations across all age groups. The Bartlett's test statistic can be shown to follow an approximate chi-square distribution when sampling is done from populations with a normal distribution (Armitage, 1971), and has the capacity to examine the homogeneity of the variances between groups on the basis of group data where only the number of subjects, mean scores and standard deviations are available for each group. The test was chosen because raw data were not available for the analyses. Due to the absence of raw data for the Phase 1 data sets, the assumption of normality of the distributions could not be empirically verified for all the tests included in the analysis. Typically, however, comparative neuropsychological test interpretation is based on the assumption that tests of intellectual function assume a normal distribution (see Lezak, 1983, p. 139). Accordingly when the raw scores available for all the tests included in Phase 2 of the study were subjected to the Kolmogorov-Smirnov test of normality procedure, each analysis resulted in a p-value > 0.05 indicating that the assumption of normality of the distributions of the scores for those neuropsychological tests was valid. Thus it was considered legitimate to assume normality of the distributions in the Phase 1 data sets, and to use the Bartlett's test for the variability analyses.

2. *Multiple pairwise comparisons of standard deviations.* For each data set which revealed a significant difference between standard deviations across age groups at the 5% level ($p < 0.05$), the following multiple comparisons procedure for the delineation of the variability patterns was adopted. By means of the F ratio test, pairwise comparisons of standard deviations between age groups were conducted between the oldest age group and each

younger age group (*backwards comparisons*), in order to identify where the differences lay. In the majority of instances the nature of the difference was adequately revealed in this way. However, if the significant difference as detected by the Bartlett's test on the overall analysis was not identified via this series of backwards comparisons, a second series was conducted which compared the youngest age group with each older age group (*forwards comparisons*). If the difference was still not revealed, a set of comparisons was conducted between each advancing pair of age groups (*progressive comparisons*). On occasion when the increases in standard deviations were very large between each age group and did not show a subsequent decrease in the older age groups, progressive forwards comparisons were also conducted in order to establish the point at which the increases in standard deviations were initiated. On the isolated occasion when the significant difference as identified by the Bartlett's statistic was not revealed via any of these three modes of serial comparisons, an additional spot comparison was made for the pair which appeared to be the source of the difference.

Each of the three mechanisms for sequential comparisons contained the capacity to obscure the presence of differences depending upon the unit of measurement and particular pattern of variability effects. The *backwards* comparison mode was isolated as the first option because it would be able to establish the presence of the late dedifferentiation effect of reduced variability between the later and middle years most proficiently; this effect could be obscured in a forwards comparison mechanism if the reduced variability in the later years was not as low as the measure of variance in the earliest age group. The *forwards* comparison mechanism would establish most effectively the exact stage of onset of progressive increases in variability if the measurement increments were gradual (but would fail to do this if the increments were large); in contrast the *progressive comparisons* would detect the stages of advancing progressive increases in variability most effectively if the increments were large (but would fail to detect small incremental differences across multiple age groups). Thus, in view of the fact that the overall data base consisted of multiple data sets from different tests with contrasting measurement indices over widely differing age ranges and numbers of groups, it was considered appropriate to implement this laddered system of comparison modes for the exploration of significant variability patterns in the data. Where series of forwards or progressive comparisons were required in addition to the backwards comparisons, in order to find where the difference lay for a submodality within a test, comparisons of the same type were completed for *all* submodalities of the test. Adjustments for alpha (see below) were calculated on the basis of the number of comparisons required for *a single set* of sequential comparisons. All sets of comparisons were conducted strictly in the order delineated until the significant patterns were satisfactorily identified, and reported in this order in the tables.

Finally, for any data sets which revealed a p-value of $>.05$ and $<.10$, a set of backwards pairwise comparisons was conducted in order to examine strong trends, and to locate any significant pairwise differences which were not evident on the overall analysis. This was justified considering the very small numbers in several studies and/or very disproportionate numbers of subjects between groups in some studies, which may have obscured differences. However, if no significant differences were identified by means of the backwards comparisons, no further comparisons were conducted. Since this occurred in most instances it provided a measure of confirmation that the overall level of significance which was set for the alpha adjustments was appropriately stringent.

3. Single pairwise comparisons of means. Due to the absence of raw data it was not possible to test for the equality of the mean scores across all age groups using the regular ANOVA, and thus all groups comparisons were not conducted for means. However, for all data sets on which multiple pairwise comparisons of standard deviations were conducted because of significant variability effects (or strong trends) noted on the basis of the Bartlett's test statistic, the mean scores between the oldest and youngest age groups were compared using the Student's t statistic. Since the focus of the study was on the pattern of changes in standard deviations across the older age group which included the likelihood of both linear and non-linear effects, it was essential to conduct multiple comparisons on the standard deviations in order to track the differences exactly. However, since a linear progression of declining cognitive performance in association with age is well-established on central tendencies analysis, it was considered adequate to examine the means in a single comparison between the youngest and oldest age group only. This single means analysis per data set would be able to identify whether there was an overall decline in cognitive performance in association with significant changes in variability across the adult age groups under investigation; it would not have the capacity to establish the exact point of onset of diminished average performance. However, it was considered sufficient for this broad first phase of the analysis, and was a necessary limitation in order to keep the overall number of pairwise comparisons within manageable proportions for the purposes of the thesis.

4. Alpha adjustments. The sequential Bonferroni method of multiple tests of individual significance of means and variances was used for all the pairwise comparisons to ensure that the overall level of significance was not larger than 0.10 (Miller, 1981). The 10% level of significance was chosen to guard against a high probability of committing a Type II error due to the large alpha adjustment necessary on a data base which involved many pairwise comparisons due to multiple age groups in the data sets. The adjustments in the level of significance were set according to the number of comparisons made between groups as follows (see summary of adjustments in the levels of significance in Table E, Appendix p.25): For each of the three modes of serial comparison described above (*backwards, forwards, and progressive forwards),* the number of comparisons amounted to one less than the number of groups in the data set (i.e. N - 1), and the alpha adjustment for the level of significance was made accordingly for each set of serial comparisons conducted. In the interests of consistency, and as a measure of stringency, when single spot pairwise comparisons were made between standard deviations, as well as all the single pairwise comparison of mean scores, these were subjected to the same adjustment in significance level as would be demanded by a full set of backwards or forwards serial comparisons for the particular data set (i.e. N - 1 comparisons).

5. Presentation of results. Separately, for each of the functional modalities of *attention, verbal memory, visual memory, language skills, visual skills, problem solving, and motor activity,* the results were summarized for presentation in tables, and illustrated graphically for emphasis of significant findings.

6.2.3.2 Phase 2

The statistical procedures for the Phase 2 analysis progressed as follows:

1. Analysis of subject characteristics. For each data set included in the Phase 2 analysis (as listed in Table C, Appendix pp.22-3), a one-way Analysis of Variance and Levene's F test, respectively, were used to compare the means and standard deviations across all age groups for the subject characteristics of age and education (all tests), and IQ (Trail Making Test only). For the Digit Symbol, Digit Symbol Incidental Recall and Finger Tapping Preferred and Non-Preferred Hands Tests, a significant effect for variability was noted for age across the five age groups, due to the wider age ranges in the two younger groups. Thus a second set of comparisons was implemented for these data sets to investigate the homogeneity of variance for age and education between the three older age groups only.

2. *Between-groups analyses.*

(i) Comparisons of means and standard deviations across all groups. For each data set of neuropsychological test scores included in the Phase 2 analysis, a one-way Analysis of Variance, and Levene's F test, respectively, were used to compare the mean scores and standard deviations across all age groups. In addition, for the Digit Symbol, Digit Symbol Incidental Recall and Finger Tapping (Preferred and Non-Preferred Hands) Tests, the mean scores and standard deviations were compared across the three oldest age groups only. Levene's F test was used for the comparisons of standard deviations across all groups in the Phase 2 analyses, instead of the Barlett's test which was used in the Phase 1 analyses, because raw data were available. The test procedure is based on an ANOVA using the absolute deviation of each case from its cell mean (Brown & Forsythe, 1974). Compared with Bartlett's test, Levene's F test is more conservative and is more stringent with regard to the possibility of a Type I error. As noted above, all the raw scores available for the tests included in Phase 2 of the study were subjected to the Kolmogorov-Smirnov test of normality procedure. Each analysis resulted in a p-value >0.05 indicating that the assumption of normality of the distributions of the scores for these tests was valid.

(ii) Multiple pairwise comparisons of means and standard deviations. For each data set of neuropsychological test scores the Student's t and the F ratio tests were used, respectively, to conduct pairwise comparisons of means and standard deviations between the oldest age group and each younger age group (*backwards comparisons*). These backwards comparisons were sufficient to identify significant differences noted on the overall group analyses and no further sets of serial pairwise comparisons were required. The general absence of significant differences for the backwards comparisons for those data sets which did not reveal an overall significant difference across all groups, provided a measure of confirmation that the overall level of significance which was set for the alpha adjustments was appropriately stringent.

(iii) Alpha adjustments. As for the Phase 1 analyses, the sequential Bonferroni method of multiple tests of individual significance of means and variances was used for all the pairwise comparisons to ensure that the overall level of significance was not larger than 0.10 (Miller, 1981). The adjustments in the level of significance were set according to the number of comparisons made between groups, which for each set of *backwards* comparisons amounted to one less than the number of groups in the data set (N - 1 comparisons). For specifications of the alpha adjustments used see Table E (Appendix p.25).

3. Within-groups analyses.

(i) Comparisons of test scores for levels of education, IQ and gender. Within each age group for each of the data sets, the Student's t and the F ratio tests were used, respectively, to compare the means and standard deviations of the neuropsychological test scores for levels of education (10-13 versus 14^+; and 10-15 versus 16^+) and for gender (female versus male). In addition two levels of NART IQ (< 130 versus 130^+) were compared within each age group for the Trail Making Test only.

(ii) Calculation of percentages of outliers in terms of levels of education, IQ and gender. Within each age group for each of the data sets, the percentages of subjects that fell beyond $\pm$ 1 Standard Deviation and beyond $\pm$ 2 Standard Deviations, for two levels of education (10-13 years versus 14+ years) and two levels of gender (female versus male), were calculated. In addition within each age group for the Trail Making Test only, the percentage of subjects that fell beyond $\pm$ 1 Standard Deviation and beyond $\pm$ 2 Standard Deviations, for two levels of NART IQ (< 130 versus 130+), were calculated. The percentages of subjects calculated were tabulated for comparative purposes and were not statistically compared.

(iii) Calculation of correlation coefficients for all variables. Correlation coefficients were computed to investigate the association between the Trail Making Part A, Trail Making Part B, education and NART IQ for each of the 60-69, 70-79, and 80-89 year-old age groups. Correlation coefficients were calculated to investigate the relationship between Digit Symbol, Digit Symbol Incidental Recall, Finger Tapping Preferred, Finger Tapping Non-preferred test scores, and education for each of the 20-39, 40-59, 60-69, 70-79, and 80-89 year-old age groups.

(iv) Presentation of results. The results for each of the data sets in the Phase 2 analysis were grouped for presentation in terms of the three levels of the statistical procedure: *Analysis of subject characteristics, Between-groups analyses,* and *Within-groups analyses.* The data were tabulated and illustrated graphically for emphasis of significant findings.

6.3 STATISTICAL HYPOTHESES FOR THE STUDY

Based on the experimental hypotheses proposed in Chapter 5 and outlined above, the following formal statistical hypotheses were posed:

6.3.1 Phase 1 and Phase 2: Statistical hypotheses for the between-groups analyses

1. It was hypothesized that there would be an overall significant difference in the standard deviations across adult age groups for the majority of neuropsychological tests included in both phases of the analysis, and that if this did not occur it would be for those tasks which have a relatively high crystallized verbal component and/or present relatively low task challenge, and/or have been shown on central trends analysis not to be particularly age-sensitive.

2. It was hypothesized that the significant differences in variability data would be due to one of two distinctive patterns across the middle to late adult years when plotted in association with a decline in average cognitive performance: either there would be a non-linear inverted-U effect with an initial increase in variability followed by a subsequent decrease in variability in the later adult years; or there would be a linear effect of progressively increasing variability across the middle to late adult years. More specifically it was hypothesized that this would occur as follows:

(i) Should the Bartlett's test or Levene's F test reveal an overall significant difference in the standard deviations between all age groups for a data set with a non-linear inverted-U trend, backwards multiple pairwise comparisons of standard deviations from the oldest group with each younger age group, would reveal significantly lower standard deviations for the oldest age group (or groups) compared with the middle age group (or groups), and no significant difference compared with the youngest age group (or groups); forwards comparisons from the youngest age group with each older age group would reveal significantly lower standard deviations for the youngest age group (or groups) compared with the middle age group (or groups), and no significant difference compared with the oldest group (or groups). Simultaneously the pairwise comparison of the mean cognitive score between the oldest and youngest age group would reveal a significantly lower mean score (where a low score represents poor functioning) when compared with the older group, or a significantly higher mean score when compared with the older group (where a high score represents good functioning).

(ii) Should the Bartlett's statistic or Levene's F statistic reveal an overall significant difference in the standard deviations between age groups for a data set with a linear trend, backwards multiple pairwise comparisons of standard deviations from the oldest age group with each younger age group, would reveal significantly higher standard deviations for the oldest group (or groups) compared with the middle age group (or groups), and the youngest age group (or groups); forwards comparisons from the youngest age group with each older age group, would reveal significantly lower standard deviations for the youngest age group (or groups) compared with the middle age group (or groups) and/or the oldest age group (or groups). Simultaneously the pairwise comparison of the mean cognitive score between the oldest and youngest age group would reveal a significantly lower mean score (where a low score represents poor functioning) when compared with the older group, or a significantly higher mean score when compared to the older group (where a high score represents poor functioning).

3. It was hypothesized that the non-linear inverted-U variability trend (rather than the linear trend) would be more likely to occur for samples with a low educational/IQ level rather than a high educational/IQ, for samples of males rather than females; and for those tasks which have a high fluid/perceptual speed component, and/or present high challenge, and/or are particularly age-sensitive, rather than for those tasks which have a high crystallized/verbal component, and/or present low challenge and/or are not particularly age-sensitive.

4. It was hypothesized that for the same neuropsychological task, a significant increase in standard deviations would occur at a relatively earlier age stage, and a subsequent significant decrease in standard deviations would occur at a relatively earlier age stage for lower levels of education, for males, for tasks which present greater challenge, and for tasks which are more age-sensitive. Conversely it was hypothesized that a significant increase in standard deviations would occur at a relatively later age stage for higher levels of education, females, lower levels of task challenge, and tasks which are less age-sensitive, such that a subsequent

significant decrease may be apparent at a later age stage, or may not appear at all even for the latest age stage.

6.3.2 Phase 2: Statistical hypotheses for the within-groups analyses

1. It was hypothesized that there would be an effect for education and IQ (as revealed in the comparisons of standard deviations for different levels of education and IQ within age groups, and in the positive correlation coefficients for education/IQ and cognitive test performance across age groups), that would be more pronounced for the age groups with significantly increased variability across the adult age range than for those age groups which did not show increased variability, that would be more pronounced for those tasks which showed a prolonged progressive increase in variability in later old age compared with those which showed an initial increase followed by a subsequent decrease in variability in later old age, and that would be more pronounced for those tasks which call upon educationally-based/crystallized/reasoning skills to a greater extent than speeded perceptual/fluid type skills.

2. It was hypothesized that where levels of education/IQ have a significant effect, the effect could be revealed in any one of the following forms for a particular age group:

(i) The standard deviation for low education/IQ may be significantly greater than that for high education/IQ due to a high proportion of individuals with low education who achieve disproportionately low scores which account for the significant increase in the overall distribution, and this effect would be reflected in a significant positive correlation between education/IQ and performance for that age group;

(ii) The standard deviation for high education/IQ may be significantly greater than that for low education/IQ due to a relatively large proportion of individuals with high education who achieve disproportionately high scores which account for the significant increase in the distribution, and this effect will be reflected in a significant positive correlation between education/IQ and performance for that age group;

(iii) There may be no difference between the standard deviations for low and high education because a relatively large number of disproportionately low scorers with low education/IQ, together with a relatively large number of disproportionately high scorers with high education/IQ, account for the significant increase in the distribution, but the effect will be revealed in a significant positive correlation between education/IQ and performance.

3. It was hypothesized that there might be differences between the standard deviations within age-groups for gender, and associated gender differences in the characterization of disproportionately good and poor scorers that account for significant increases in variability in the late adult years, and that the effect could be revealed in any one of the following forms for a particular age group:

(i) The standard deviation for males may be significantly greater than that for females due to a high proportion of males who achieve disproportionately low scores which account for the significant increase in the overall distribution;

(ii) The standard deviation for females may be significantly greater than that for males due to a relatively large proportion of females who achieve disproportionately high scores which account for the significant increase in the distribution;

(iii) There may be no difference between the standard deviations between males and females because a relatively large proportion of disproportionately low scoring males, together with a relatively large proportion of disproportionately high scoring females, contribute to the significant increase in the distribution.

7

CHAPTER 7: RESULTS PHASE 1

In this section the results for Phase 1 of the analysis are presented in a separate subsection for each of the functional modalities of *attention, verbal memory, visual memory, language skills, visual skills, problem solving, and motor activity.* The results are summarized in tables for each functional modality and are presented together at the end of each subsection. Figures are presented at the next available opportunity after they are first mentioned in the text. Thus page references will *not* be given for tables and figures which, when alluded to, follow on obviously within each subsection in the chapter. However, as for the rest of the thesis, page numbers *will* be given when cross-reference is made to tables (or figures) which are contained outside of the chapter, and/or are not easily accessible within the chapter.

For all tests within each modality there is a single table for the Bartlett's test comparisons of standard deviations across all age groups, and a second table for the multiple comparisons of means and standard deviations between age groups. This grouping was done in the interests of clarity in that, due to adjustments in the levels of significance for the multiple comparisons, the significance levels of the all-groups analyses are denoted differently from the multiple pairwise comparisons. In addition, this clustering facilitates the ability to scan the general trends across the all groups analyses for an entire modality. For further ease of reference, whenever comparisons are made between tests and/or data sets, these are illustrated graphically, discussed and referenced in the order in which they appear in the subsection tables. This is also the order in which the full spectrum of the Phase 1 group data and demographic details are collated in the Appendix Tables A and B (pp.1-21). The order is based on the step-wise procedure by which the data were accessed into the meta-analysis. Thus, within each modality, all data from the collation of Spreen and Strauss (1991) appear first, followed by any data from van Gorp et al. (1990) and Lezak (1983), and followed finally by all the additional data sets. Where a data set is made up from a number of sources, only the author of the collation is referenced in the tables and used for graphical purposes, and will appear as 'In Lezak' or 'In Spreen and Strauss'; otherwise the actual data source will be referenced.

Due to the large amount of data in the analysis, except for particular emphasis, only significant p-values are reported in the text. For the purpose of cross-referencing p-values are reported in the text as they appear in the tables. For the all-groups comparisons, these values require no adjustment in terms of the levels of significance which they denote; for the multiple pairwise comparisons, the levels of significance as determined by Bonferroni's adjustment for simultaneous testing are indicated in the text (as in the results tables) by means of asterisks that denote the following levels of significance: $*p<0.10$; $**p<0.05$; and $***p<0.01$ (see Appendix Table E, p.25). As noted in the method (section 6.2.3), in accordance with Miller (1981), the overall level of significance for the multiple comparisons was set at $p=0.10$ in order to guard against the probability of a Type II error (failure to identify significant differences detected on the analyses across all groups) in view of the large numbers of comparison groups for most of the data sets.

7.1 ATTENTION

For the modality of attention the data sets submitted to analysis are collated in Table A (Appendix pp.1-2), and a summary of the educational and/or IQ levels of the age groups appear in Table B (Appendix p.18). The following tests and accompanying sources of normative data sets were included in the analysis (ordered for discussion as they appear in the tables):

(i) Paced Auditory Serial Addition Test (PASAT): Three data sets (Stuss, Stethem, & Pelchat, 1988, in Spreen & Strauss, 1991; Stuss, Stethem, & Poirier, 1987; and Roman, Edwall, Buchanan & Patton, 1991).

(ii) Digit Span Forward and Backward: One data set (Shuttleworth-Jordan, 1992).

7.1.1 Comparisons of standard deviations across age groups

For the *PASAT* (see Figure 7.1, based on data contained in Tables 7.1 and 7.2), the patterns of standard deviations across groups between the three data sets were as follows. For the first PASAT study *(Stuss et al., 1988 in Spreen & Strauss, 1991)* there was a highly consistent linear trend of increased variability in association with older age for both the 2.4 and 1.2 second presentations (Figure 7.1, top graph). The comparisons of standard deviations across all age groups revealed no significant effects for either presentation although the effect for the 1.2 second presentation approached significance ($p=0.0915$). Backwards pairwise

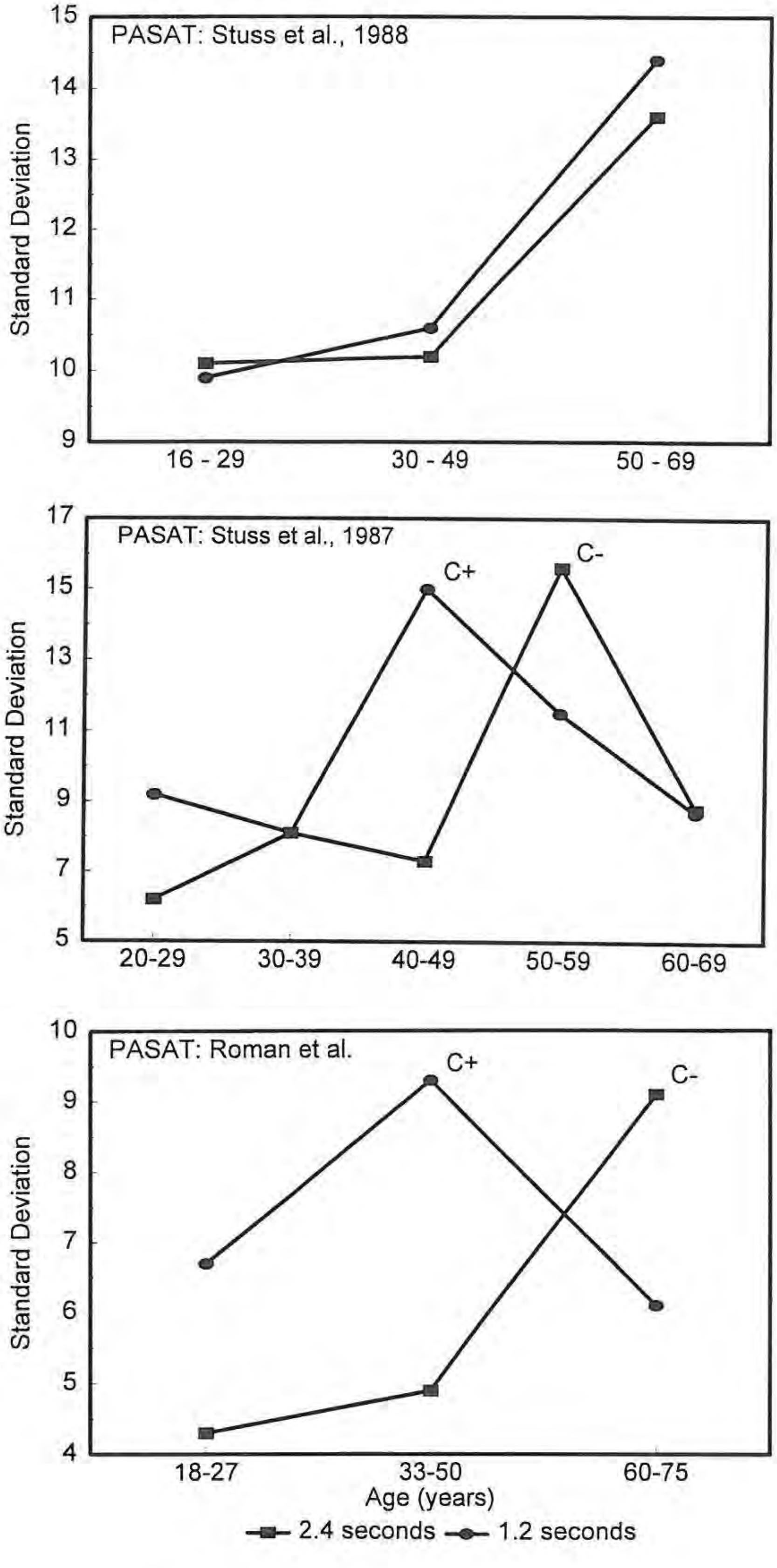


Figure 7.1. PASAT (2.4 and 1.2 second presentations): Standard deviations across adult age groups for three data sets (Stuss et al., 1988, in Spreen & Strauss, 1991; Stuss et al., 1987; Roman et al., 1991). The points at which variability peaks for tasks of relatively high versus relatively low challenge are designated c+ (left shuttle effect) and c- (right shuttle effect), respectively.

comparisons between age groups for the 1.2 second presentation revealed that the standard deviation for the oldest 50-69 year-old age group was significantly higher than that of each of the two younger 30-49 and 16-29 year-old age groups ($p=0.0284^{*}$ and $p=0.0240^{**}$, respectively) which provided support for the presence of a significant linear pattern of increased variability across age groups in this subset of PASAT data, although the effect had only approached significance on the analysis across all groups.

For the second PASAT study *(Stuss et al., 1987)*, there was a highly consistent inverted-U pattern of an initial increase and subsequent decrease in variability for both the 2.4 and 1.2 second levels of presentation (Figure 7.1, middle graph). On the analyses of standard deviations across all age groups there was a significant effect for the 2.4 second presentation ($p<0.01$), but the effect for the 1.2 second presentation did not reach significance. Forwards comparisons between age groups revealed that for the 2.4 second presentation the standard deviation of the 20-29 year-old age group was significantly lower than that of the 50-59 year-old group ($p=0.0056^{**}$) but that there was no significant difference between 20-29 year-old age group and the 30-39, 40-49 and 60-69 year-old age groups, denoting that for this data set there was a significant peak in variability for the 50-59 year old age group followed by a subsequent significant decline in variability for the 60-69 year old age group. For the 1.2 second presentation (the level of greater challenge) there was evidence of an identical inverted-U trend except that it occurred earlier, with a clearly discernable trend of an initial increase in the 40-49 year-old age group, followed by a subsequent decrease in the 50-59 and 60-69 year-old age groups, although the apparent peak in the 40-49 year-old age group did not reach significance.

For the third PASAT study *(Roman et al., 1991)*, there was the combination of a linear trend of increased variability, and an inverted-U trend of an initial increase and subsequent decrease in variability, respectively, for the 2.4 and 1.2 second presentations (Figure 7.1, bottom graph). On the analyses across all groups, there were significant effects for both levels of presentation ($p<0.01$ and $p=0.0146$, respectively). The backwards comparisons for the 2.4 second presentation revealed a significantly higher standard deviation for the oldest 60-75 year-old age group compared with those of each of the 18-29 and 33-50 year-old younger age groups ($p<0.001^{***}$; $p<0.001^{***}$), which denotes a linear pattern of significantly increased variability across the younger to older adult age groups with no subsequent decrease in variability. In contrast the backwards comparisons for the 1.2 second presentation (the level of greater challenge) revealed a significantly lower standard deviation for the oldest 60-75 year-old age group compared with the middle 33-50 year-old age group ($p<0.0046^{**}$), but no difference between the standard deviation of the 60-75 year-old group compared with that of the youngest 18-27 year-old group which denotes a significant non-linear pattern of variability across the age groups with a significant peak in variability at the 33-50 year interval, and a subsequent significant decrease in variability at the 60-75 year interval. In this last PASAT data set (Roman et al.), for which the overall effects across age groups were the most clearly significant across both the 2.4 and 1.2 levels of presentation, it is notable that the groups comprised larger numbers of subjects per age group than the other two PASAT studies. In the Roman et al. study there were three groups of 62, 40 and 41 subjects each, whereas the other studies comprised 30 subjects per age group (Stuss et al. 1988, in Spreen & Strauss 1991) and as low as 10 subjects per age group (Stuss et al., 1977). With increased subject numbers strong trends in the latter two data sets might also have reached significance.

The results for the PASAT data sets of Stuss et al. (1987) and Roman et al. (1991) which are illustrated graphically in the middle and lower graphs of Figure 7.1 lay emphasis on a consistent shuttle pattern that is evident for these two studies. In both cases the initial increase in variability occurs earlier for the 1.2 second presentation (the task of greater challenge designated C+ on the graphs) when compared with the initial increase in variability for the 2.4 second presentation (the task of less challenge designated C- on the graphs). For the Stuss et al. (1987) data set two full non-linear inverted-U patterns were revealed with the total curve for the more challenging task shifted to the left. For the Roman et al. study a full non-linear pattern of an increase and subsequent decrease in variability was apparent for the more challenging 1.2 second presentation (left shuttle effect), in comparison with the less challenging 2.4 second presentation where only a later increase in variability was revealed with no subsequent decrease in evidence (right shuttle effect).

The reason for the apparent absence of a dedifferentiation effect for the first Stuss et al. (1988, in Spreen & Strauss, 1991) data set, compared with the second Stuss et al. (1987) data set, may be due to the manner in which the data have been clustered by Spreen and Strauss in the former collation (across two age decades rather than only a single age decade), such that an inherent dedifferentiation effect in the 60's age decade may have been obscured by a large differentiation effect in the 50's age decade. This particular combination of data may also have obscured a left versus right shuttle effect between the two levels of presentation. It is also possible that were a 70-89 year-old age group available for the Stuss et al. (1988) data set, that a subsequent dedifferentiation effect would have become apparent. Similarly it is possible that were a 76-89 year-old age group available for the Stuss et al. (1987) 2.4 second presentation that a subsequent dedifferentiation effect would have been evident. Hence the three sets of PASAT data are considered to show a high level of consistency particularly between the second two studies, and not to be fundamentally incompatible when including the first study.

For the *Digit Span test* (Shuttleworth-Jordan, 1991), (see Figure 7.2, based on the data contained in Tables 7.1 and 7.2), there was a highly consistent non-linear inverted-U pattern across age groups for both levels of presentation. The comparisons of standard deviations across all groups revealed an effect which approached significance for Digits Forward ($p=0.09$), and an effect which was highly significant for Digits Backward ($p<=0.01$). Backwards pairwise comparisons revealed that in the case of both Digits Forward and Digits Backward there was a significantly lower standard deviation for the oldest 80-89 year-old age group compared with the 60-69 year-old age group ($p<0.0119^{**}$ and $p<0.0101^{**}$, respectively) and no significant difference between the 80-89 year-old age group age group and the 70-79, 40-49 and 18-25 year-old age groups.

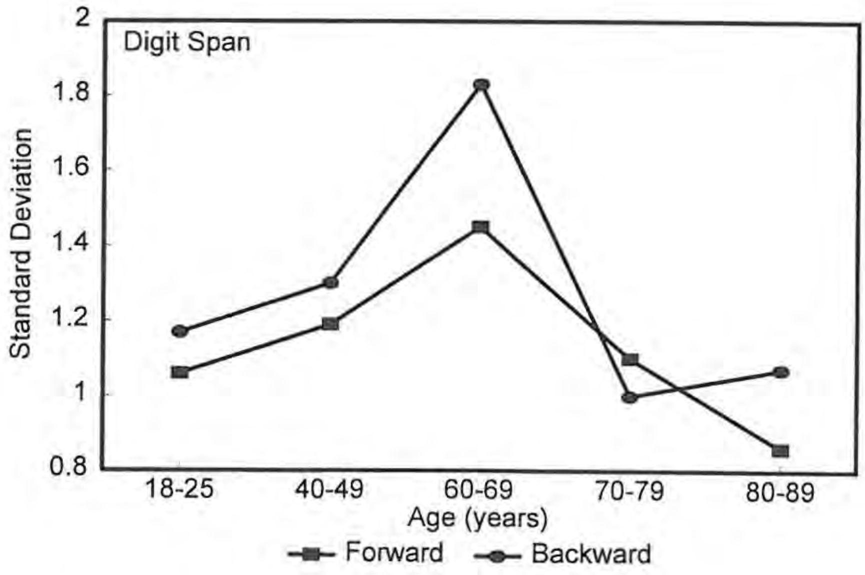


Figure 7.2. Digit Span (Forward and Backward): Standard deviations across adult age groups (data from Shuttleworth-Jordan, 1992) illustrating non-linear inverted-U patterns of variability in association with age.

This denotes that for both submodalities of this test there was a significant non-linear pattern of variability across the age groups with a significant peak in the 60's and subsequent significant decrease in the 70's and 80's age groups. From the graph in Figure 7.2 it is evident that an identical pattern occurred across age groups for Digits Forwards and Backwards, with neither shifted to the left or right of each other.

7.1.2 Comparisons of means

For the modality of attention (see Table 7.1 for a summary of mean scores across all groups, and Table 7.2 for pairwise statistical comparisons between the oldest and youngest age groups) in all instances with the exception of the Stuss et al., (1987) PASAT data, there was a clear trend of declining mean scores across older age groups. The fluctuating mean score trends on the Stuss et al. (1987) study are not consistent between the 2.4 and 1.2 second presentations of the data set. This contrasts with a coherent pattern of standard deviations in this study which appears consistently between the two levels of presentation, and across two PASAT studies, suggesting that the variability effects for this Stuss et al. (1987) data set have been more robust than possible sampling inequities that have influenced the central tendencies.

In most instances of data sets where the overall comparisons of standard deviations across age groups revealed a significant effect, there was also a significantly lower mean score for the oldest age group compared with the youngest age group. Conversely, when the overall effect for standard deviations across all age groups failed to reach significance, the differences between the mean scores of the youngest and oldest age groups also failed to reach significance. This indicates that with advancing age there was a significant decline in average performance which occurred in parallel with either a significant linear or significant non-linear inverted-U pattern of variability (see Figures 7.3 and 7.4). In particular where the multiple pairwise comparisons of standard deviations between-groups revealed highly significant effects for all levels of presentation within a data set, the significant decline in mean scores was clear-cut. Thus for the PASAT data set from Roman et al. (1991), the pairwise comparisons of mean scores for the 2.4 and 1.2 seconds revealed highly significant differences between the youngest and the oldest age groups indicating diminished average performance in the older age groups for both the 2.4 and 1.2 second presentations ($p<0.001^{***}$; $p<0.001^{***}$, respectively), as did the pairwise comparisons of mean scores between the youngest and oldest age groups for both Digits Forward and Digits Backward ($p < 0.0177^{*}$; $p < 0.001^{***}$, respectively).

Of note from Figures 7.3 and 7.4) is that the significant increases (and subsequent decreases) in variability in association with aging occurred broadly in parallel with declining average performance. However the *onset of significantly increased variability* either occurred *simultaneously* with the onset of markedly (and consistently) declining average performance (Figure 7.3 upper graph), or *preceded* it by one to two decades (Figure 7.3 lower graph, and Figure 7.4, upper and lower graphs).

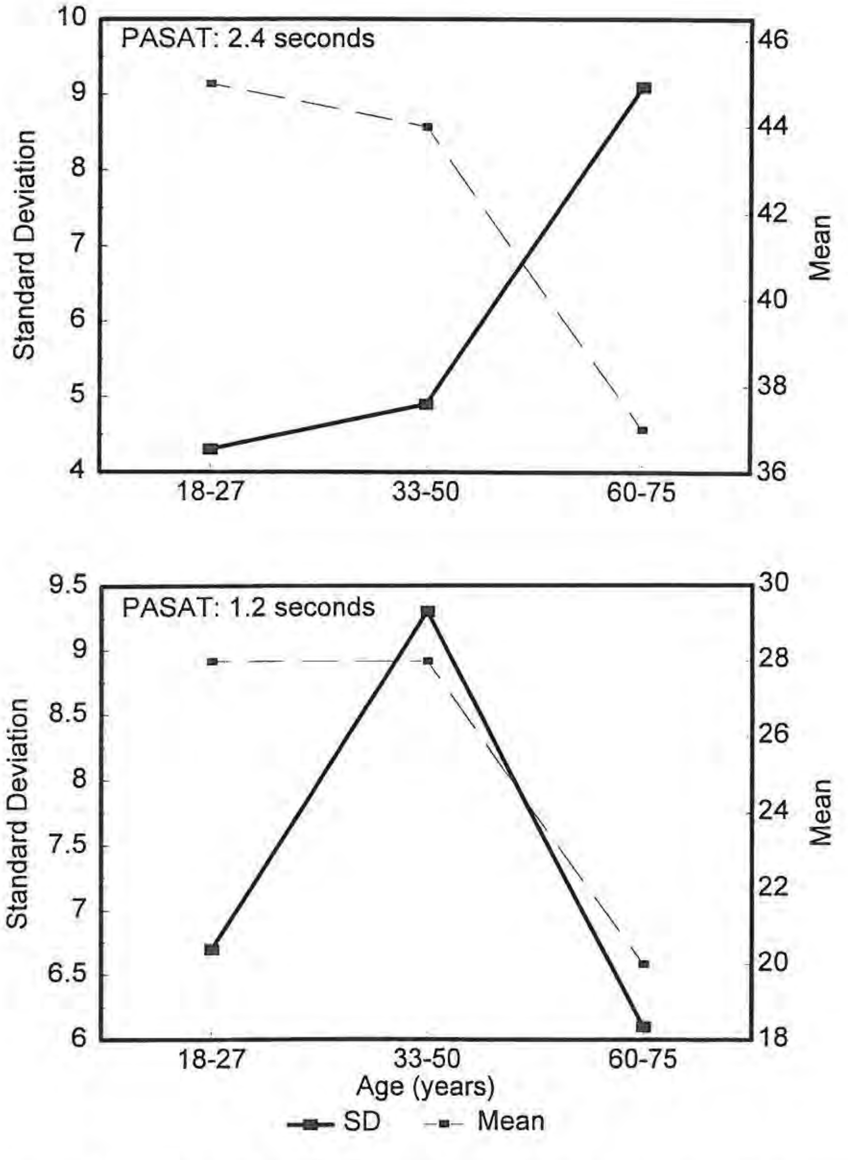


Figure 7.3. PASAT (2.4 and 1.2 seconds): Standard deviations and mean scores across adult age groups (data from Roman et al., 1991) illustrating linear (top graph) and non-linear inverted-U (lower graph) patterns of variability in association with declining mean scores.

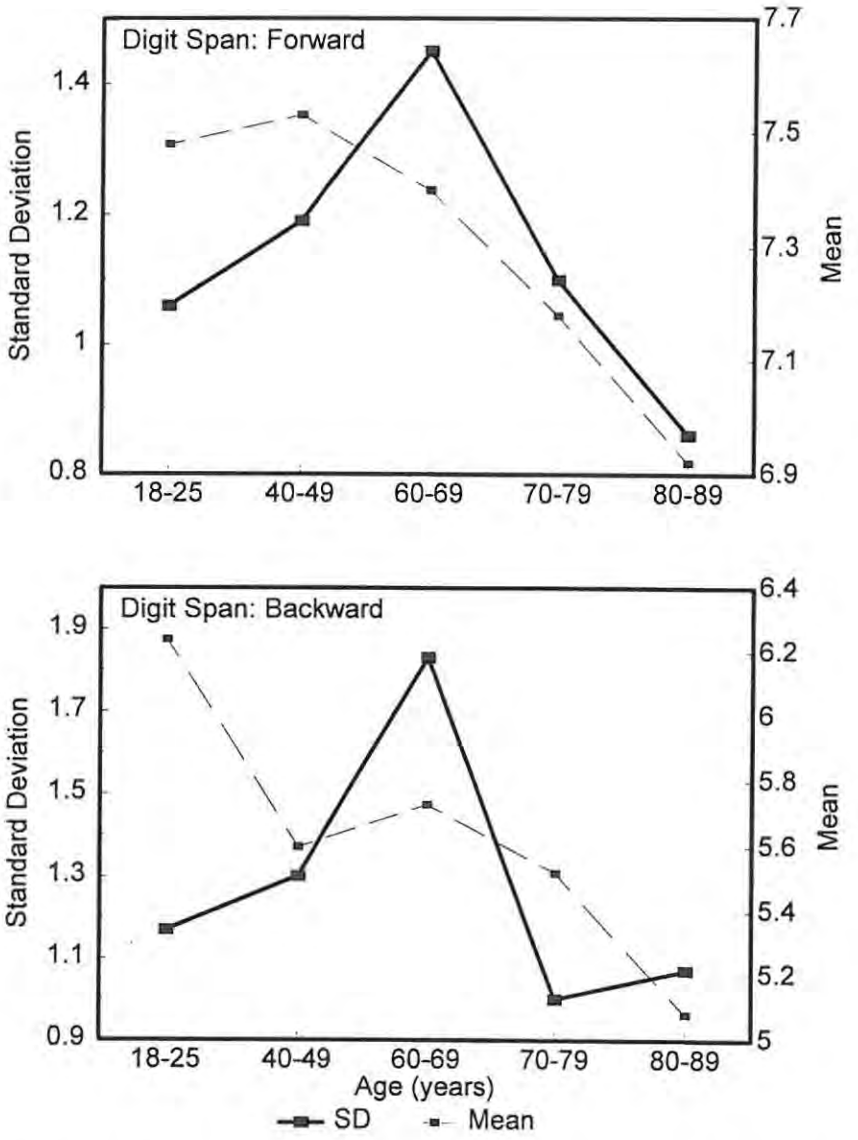


Figure 7.4. Digit Span (Forward and Backward): Standard deviations and mean scores across adult age groups (data from Shuttleworth-Jordan, 1992) illustrating non-linear inverted-U patterns of variability in association with declining mean scores.

7.1.3 Attention: Summary of effects

Generally for attention, as tapped by the PASAT and Digit Span Tests, there was support for a significant increase in variability with advancing age by the 60's age decade. In addition it appeared that for an attention task of particularly high challenge (such as the PASAT 1.2 second presentation), this significant increase in variability occurred two decades earlier, around the 40's age stage. In addition there was support for an overall significant non-linear inverted-U pattern of variability for attention (that is the full shuttle bulge effect of initial differentiation followed by subsequent dedifferentiation), in that for both the PASAT and Digit Span tests an initial increase in variability was most often followed by a subsequent decrease around one decade later. Where this did not occur either it was in an instance where an old enough age group had not been examined in isolation (Stuss et al., 1988, in Spreen & Strauss, 1991) or in the instance of the less challenging submodality of a task (Roman et al., 1991).

For two PASAT studies, a clear shift pattern was detected between the 1.2 second attention task of higher challenge (for which there was a consistent left shuttle effect across two studies) and 2.4 second task of lesser challenge (for which there was a consistent right shuttle effect across two studies), such that in one study there were two full non-linear inverted-U patterns in evidence with the one of higher challenge shifted more to the left, and the one of lower challenge shifted more to the right; in another PASAT study there was one full non-linear inverted-U pattern in evidence for the task of higher challenge shifted more to the left, in conjunction with only a linear increase in evidence for the task of lesser challenge which was shifted more to the right.

All of the above variability effects for attention occurred broadly in correspondence to a significant decline in average performance due to aging. It appeared, however, that the precise *onset* of significant variability effects might differ in relation to changes in the central tendency, in that increases in variability either occurred closely in parallel with the onset of consistently diminishing mean performances, or in a number of instances *preceded* clear-cut signs of steadily declining average performance by one or two decades.

Table 7.1. **ATTENTION:** Mean Scores and Comparisons of Standard Deviations across all Age Groups.

***PASAT* (Stuss et al. 1988, in Spreen & Strauss 1991)**

AGE GROUP	16-29	30-49	50-69	df	B[1]	p
n	30	30	30	-	-	-
PASAT-2.4 sec						
Mean	47.4	43.4	43.5	-	-	-
SD	10.1	10.2	13.6	2	3.44	0.1792
PASAT-1.2 sec						
Mean	27.4	24.6	21.2	-	-	-
SD	9.9	10.6	14.4	2	4.78	0.0915⁺

***PASAT* (Stuss et al. 1987)**

AGE GROUP	20-29	30-39	40-49	50-59	60-69	df	B	p
n	10	10	10	10	10	-	-	-
PASAT-2.4 sec								
Mean	51.3	44.8	49.1	35.3	47.7	-	-	-
SD	6.2	8.1	7.3	15.6	8.8	4	9.91	<0.01**
PASAT-1.2 sec								
Mean	33.6	27.6	23.0	8.2	30.7	-	-	-
SD	9.2	8.1	15.0	11.5	8.7	4	4.70	0.1171

***PASAT* (Roman et al. 1991)**

AGE GROUP	18-27	33-50	60-75	df	B	p
n	62	40	41	-	-	-
PASAT-2.4 sec						
Mean	45.0	44.0	37.0	-	-	-
SD	4.3	4.9	9.1	2	31.85	<0.01**
PASAT-1.2 sec						
Mean	28.0	28.0	20.0	-	-	-
SD	6.7	9.3	6.1	2	8.45	0.0146*

[1] Bartlett's test statistic

Table 7.1 *continued*

Table 7.1. ATTENTION: Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

Digit Span (Shuttleworth-Jordan 1992) AGE GROUP	18-25	40-49	60-69	70-79	80-89	df	B	p
n	33	15	15	33	25	-	-	-
Digit Span-Forward								
Mean	7.48	7.53	7.40	7.18	6.92	-	-	-
SD	1.06	1.19	1.45	1.10	0.86	4	5.24	0.09⁺
Digit Span-Backward								
Mean	6.24	5.60	5.73	5.52	5.08	-	-	-
SD	1.17	1.30	1.83	1.00	1.07	4	9.17	0.01*

Significant difference (*$p<0.05$; **$p<0.01$); Approaching significance ($^{+}$ $p<0.10$).

Table 7.2. ATTENTION: Pairwise Comparisons of Means between the oldest and youngest Age Groups and of Standard Deviations between the oldest Age Group and each younger Age Group unless otherwise specified.

PASAT **(Stuss et al. 1988, in Spreen & Strauss 1991)**						
AGE GROUP	16-29	30-49	50-69	df	t/F	p
n	30	30	30	-	-	-
PASAT-2.4 sec						
Mean	47.4	-	43.5	58	1.26	0.1062
SD	-	10.2	13.6	29,29	0.56	0.0636
SD	10.1	-	13.6	29,29	0.55	0.0574
PASAT-1.2 sec						
Mean	27.4	-	21.2	58	1.94	0.0284*
SD	-	10.6	14.4	29,29	0.54	0.0523
SD	9.9	-	14.4	29,29	0.47	0.0240**

PASAT **(Stuss et al. 1987)**								
AGE GROUP	20-29	30-39	40-49	50-59	60-69	df	t/F	p
n	10	10	10	10	10			
PASAT-2.4 sec								
Mean	51.3	-	-	-	47.7	18	1.06	0.1521
SD	-	-	-	15.6	8.8	9,9	3.14	0.0516
SD	-	-	7.3	-	8.8	9,9	0.69	0.2933
SD	-	8.1	-	-	8.8	9,9	0.85	0.4045
SD	6.2	-	-	-	8.8	9,9	0.50	0.1558
PASAT-1.2 sec								
Mean	33.6	-	-	-	30.7	18	0.72	0.2391
SD	-	-	-	11.5	8.7	9,9	1.75	0.2092
SD	-	-	15.0	-	8.7	9,9	2.97	0.0601
SD	-	8.1	-	-	8.7	9,9	0.87	0.4174
SD	9.2	-	-	-	8.7	9,9	1.12	0.5467

Table 7.2 *continued*

Table 7.2. ATTENTION: Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

PASAT **(Stuss et al. 1987)** AGE GROUP	20-29	30-39	40-49	50-59	60-69	df	t/F	p
Forwards Comparisons								
PASAT-2.4 sec								
SD	6.2	8.1	-	-	-	9,9	0.59	0.2190
SD	6.2	-	7.3	-	-	9,9	0.72	0.3172
SD	6.2	-	-	15.6	-	9,9	0.16	0.0056**
SD	6.2	-	-	-	8.8	9,9	0.50	0.1558
PASAT-1.2 sec								
SD	9.2	8.1	-	-	-	9,9	1.29	0.6447
SD	9.2	-	15.0	-	-	9,9	0.38	0.0807
SD	9.2	-	-	11.5	-	9,9	0.64	0.2583
SD	9.2	-	-	-	8.7	9,9	1.12	0.5647

PASAT **(Roman et al. 1991)** AGE GROUP	18-27	33-50	60-75	df	t/F	p
n	62	40	41	-	-	-
PASAT-2.4 sec						
Mean	45.0	-	37.0	101	5.99	<0.001***
SD	-	4.9	9.1	39,4	0.29	<0.001***
SD	4.3	-	9.1	61,4	0.22	<0.001***
PASAT-1.2 sec						
Mean	28.0	-	20.0	101	6.14	<0.001***
SD	-	9.3	6.1	39,4	2.32	0.0046***
SD	6.7	-	6.1	61,4	1.21	0.2662

Table 7.2 *continued*

Table 7.2. ATTENTION: Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

Digit Span (Shuttleworth-Jordan 1992) AGE GROUP	18-25	40-49	60-69	70-79	80-89	df	t/F	p
n	33	15	15	33	25	-	-	-
Digit Span-Forward								
Mean	7.48	-	-	-	6.92	56	2.16	0.0177*
SD	-	-	-	1.10	0.86	32,24	1.64	0.1078
SD	-	-	1.45	-	0.86	14,24	2.84	0.0119**
SD	-	1.19	-	-	0.86	14,24	1.92	0.0783
SD	1.06	-	-	-	0.86	32,24	1.52	0.1465
Digit Span-Backward								
Mean	6.24	-	-	-	5.08	56	3.88	<0.001***
SD	-	-	-	1.00	1.07	32,24	0.87	0.3555
SD	-	-	1.83	-	1.07	14,24	2.93	0.0101**
SD	-	1.30	-	-	1.07	14,24	1.48	0.1945
SD	1.17	-	-	-	1.07	32,24	1.20	0.3288

Significant difference (*$p<0.10$; **$p<0.05$; *** $p<0.01$) as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

7.2 VERBAL MEMORY

For the modality of verbal memory the data sets submitted to analysis are collated in Table A (Appendix pp.2-6), and a summary of the educational and/or IQ levels of the age groups appear in Table B (Appendix p.18). The following tests and accompanying sources of normative data sets were included in the analysis (ordered for discussion as they appear in the tables):

(i) Wechsler Memory Scale (WMS) Associate Learning: Two data sets (Strauss & Spreen, 1989, in Spreen & Strauss, 1991; data collated from Wechsler, 1945, Hulicka, 1966, and Klonoff & Kennedy, 1966, in Lezak, 1983).

(ii) Wechsler Memory Scale (WMS) Logical Memory: Three data sets (Abikoff, Alvir, Hong, Sukoff, Orazio, Soloman, et al. 1987, in Spreen & Strauss, 1991; Van Gorp, Satz., & Mitrushina, 1990; data collated from Wechsler, 1945, Hulicka, 1966, and Klonoff & Kennedy (80-92), 1966, in Lezak, 1983).

(iii) Rey Auditory-Verbal Learning Test (RAVLT): Two data sets including Trial 1, Trial 5, Total, Delayed, Recognition (Geffen, Hoar, & O'Hanlon 1990, in Spreen & Strauss, 1991); Trial 1, Trial 5, Recognition (Mitrushina, Satz, Chervinsky & D'Elia, 1991).

(iv) Buschke Selective Reminding Test (BSRT): One data set (Larrabee, Trahan, Curtiss, & Levin, 1988, in Spreen & Strauss).

7.2.1 Comparisons of standard deviations across all groups

For the *WMS Associate Learning test* (see Figure 7.5a, based on the data contained in Tables 7.3 and 7.4), the comparisons of standard deviations across all age groups revealed a highly significant effect for both data sets included in the study (Strauss & Spreen 1989, in Spreen & Strauss, 1991, and the collated data in Lezak, 1983) ($p < 0.01$ in each case). Backwards pairwise comparisons for the *Strauss and Spreen (1989)* data set (Figure 7.5a, upper graph) revealed a significant difference only between the oldest 70+ age group and the youngest

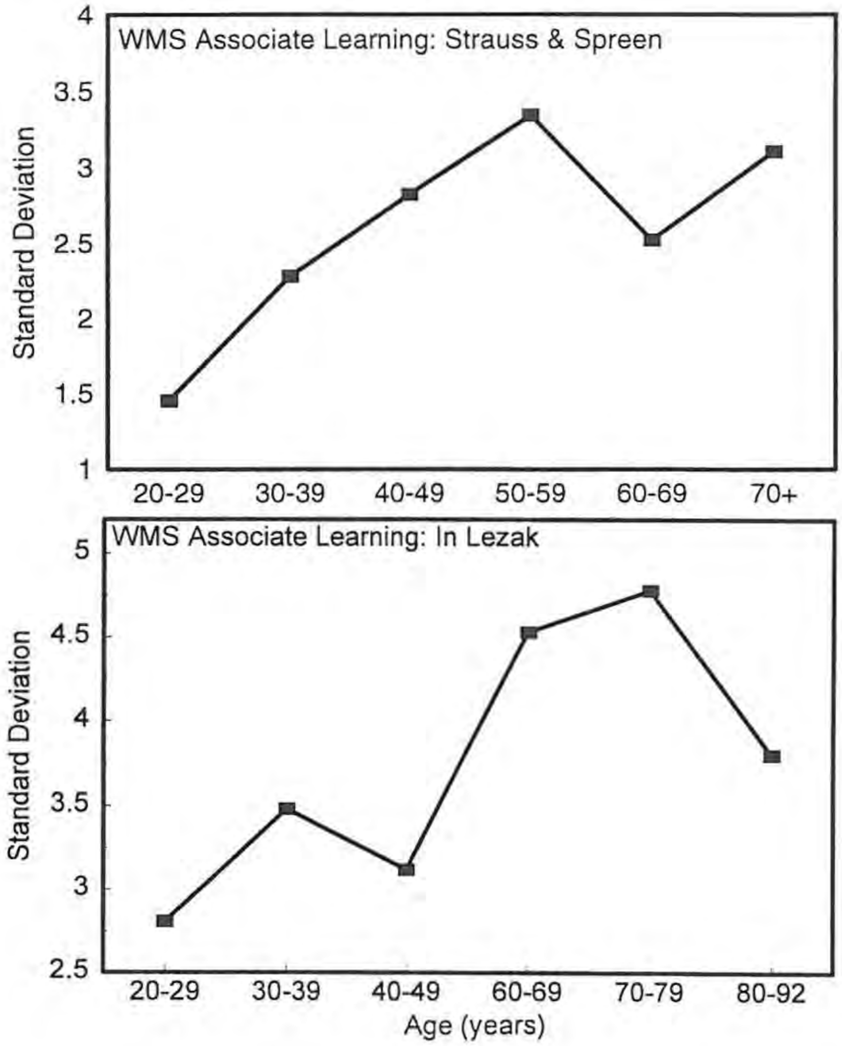


Figure 7.5a. Wechsler Memory Scale (WMS) Associate Learning: Standard deviations across adult age groups for two data sets (Strauss & Spreen, 1989, in Spreen & Strauss, 1991); collated data in Lezak, 1983).

20-29 year old age group ($p=0.0009^{**}$) and no difference between this oldest 70+ year-old age group and any of the other age groups. Similarly for the data set in *Lezak (1983)* (Figure 7.5a, lower graph) backwards comparisons revealed a significant difference only between the oldest 80-92 year-old age group and the youngest 20-29 year old age group ($p=0.0009^{**}$) and no difference between this oldest 80-92 age group and any of the intermediary age groups ($p=0.0094^{**}$). This implies that an initial increase in variability occurred as early as the 30's age decade for both data sets, and as is evident from Figure 7.5a, there was a continued tendency for variability in both data sets to increase up to the 70's age stage.

The inconsistent apparent dips in the standard deviation in the 60-69 year-old age group relative to the 70+ age group for the Strauss and Spreen data set, and in the 40-49 year-old age group relative to the 80-92 year-old age group for the Lezak data set, were *not* *significant*, and thus do not fundamentally detract from the overall pattern of increasing variability up to the 70's age group for this test across the two studies. On the other hand the apparent older-age dip in the standard deviation of the 80-92 year old group compared with the 70-79 age group for the Lezak data set only just missed significance ($p < 0.0269$), and may thus be indicative of the beginnings of a very late subsequent reduction in variability for the WMS Associate Learning test. The fact that this possible older-age dedifferentiation effect was not evident in the Strauss and Spreen data set may be due to the fact that unlike Lezak, Strauss and Spreen did not include an 80's older old age group (see direct comparison of the two studies in Figure 7.5b,), and had a data set for the older age group been available it might also have revealed the tendency for a dedifferentiation effect. (Again it is emphasized that the dip in the standard deviation of 60-69 year-old age group relative to the 70+ age group in the Strauss and Spreen data set was not significant: $p=0.2327$).

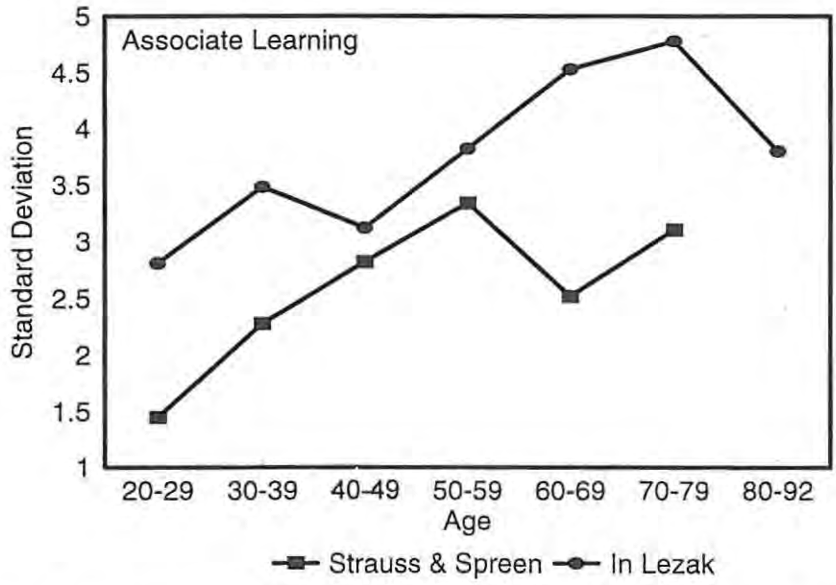


Figure 7.5b. Wechsler Memory Scale (WMS) Associate Learning: Standard deviations across adult age groups compared between two data sets (Strauss & Spreen, 1989, in Spreen & Strauss, 1991; collated data in Lezak, 1983). For graphical purposes, an artificial midpoint was created for the Lezak set at the 50-59 year-old stage, for which no actual data were available.

For the *WMS Logical Memory test* (see data contained in Tables 7.3 and 7.4) there was no clearly discernable pattern of trends for the *Abikoff et al. (1987)* study, and there were no significant effects on any of the statistical analyses. The comparisons of standard deviations across all groups revealed an effect which approached significance for immediate recall ($p=0.0757$), but follow-up backwards pairwise comparisons for the data set did not identify any significant differences in standard deviations between the age groups. In contrast for the

WMS Logical Memory test (see Figure 7.6), there was a highly comparable non-linear inverted-U pattern of trends for the other two data sets included in the analysis (van Gorp et al., 1990; collated data in Lezak, 1983), with variability peaking around the 70's and diminishing around the later 80's age stages in both studies. For the *van Gorp et al. (1990)* data set the comparisons of standard deviations across all groups for immediate recall revealed an effect which approached significance ($p < 0.0728$), and a significant effect for Delayed recall ($p < 0.0187$). Similarly for the data collated in *Lezak (1983)* the comparisons of standard deviations across all groups for immediate recall revealed a significant effect

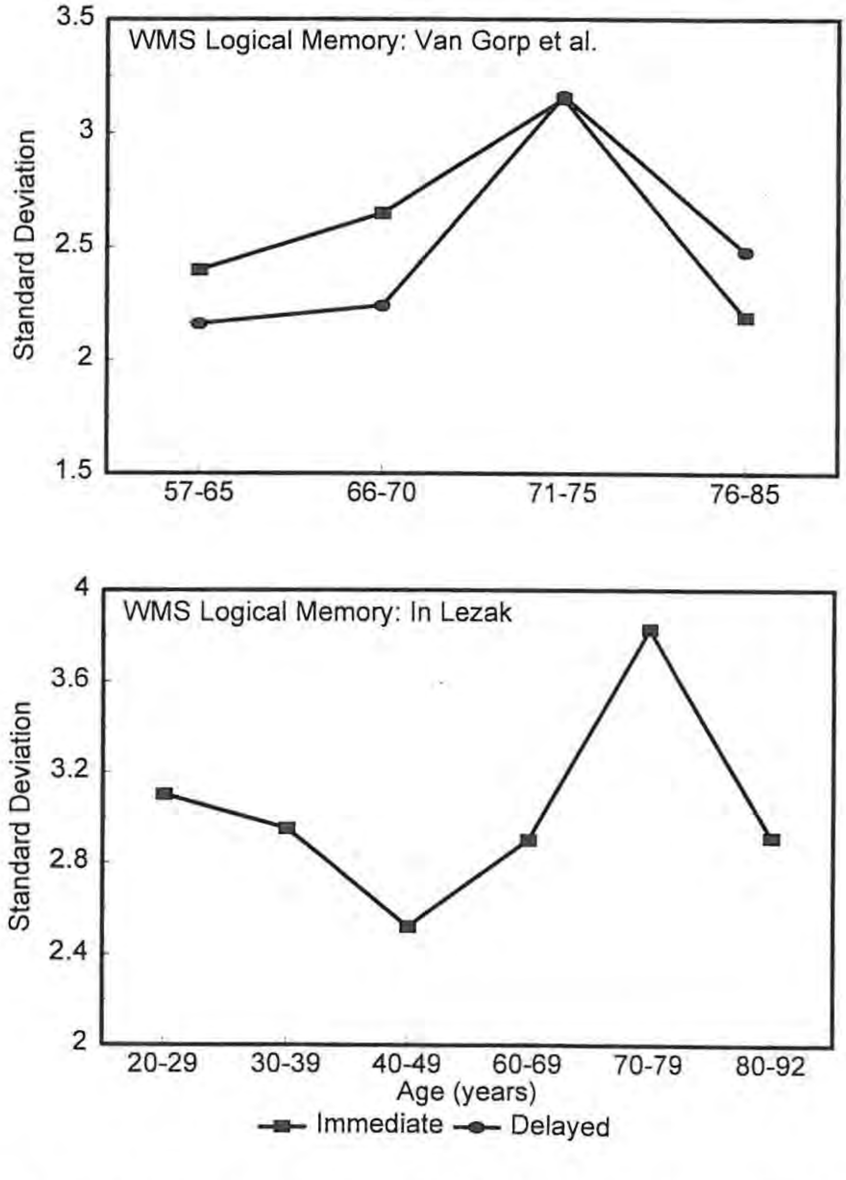


Figure 7.6. Wechsler Memory Scale (WMS) Logical Memory: Standard deviations across adult age groups for two data sets (Van Gorp et al., 1990; collated data in Lezak, 1983).

($p < 0.0141$). (No data for a Delayed recall version of this test were available from Lezak, 1983). For the van Gorp et al. study, backwards comparisons for immediate recall revealed a significantly lower standard deviation for the oldest 76-85 year-old age group compared with the next oldest 71-75 year-old age group ($p = 0.0245^*$) and no differences between the oldest 76-85 year-old age group and the 66-75 and 57-65 year-old age groups; forwards comparisons for Delayed recall revealed a significantly lower standard deviation for the youngest 56-65 year-old group compared with the 71-75 year-old age group ($p = 0.0167^*$), and no significant differences between the 57-65 year-old group and the 66-70 and the 76-85 year-old groups. Thus there was a significant peak in the 71-75 year-old age group compared with the older and younger age groups on the van Gorp et al. study for both immediate and Delayed recall. Entirely commensurate with this pattern, backwards comparisons for the Lezak data set revealed a significantly lower standard deviation for the 80-92 year-old group compared with the 70-79 year-old group ($p = 0.0104^*$), and no significant differences between the 80-92 year-old group and any of the other younger age groups. Thus for both studies there was a significant non-linear trend in variability with a significant peak in variability by the 70's age stage, and a subsequent significant decrease in the oldest age groups incorporating 80 year-old subjects. (The apparent dip in variability at the 40's age stage for the Lezak data does not contradict a relatively stable pattern of variability effects up to the 70's age decade, in that all the comparisons between the standard deviation of the 80-92 year-old age group and those of each of the 60's, 40's, 30's and 20's age groups were *not significant*).

The reason for the lack of consistency for Logical Memory between the Abikoff et al. data set for which there were no significant effects, and the data sets of van Gorp et al. and Lezak for which there were matching effects, is unclear. Since the two totally different studies of van Gorp et al. and Lezak revealed highly comparable significant variability effects it would appear likely that this is a robust finding. However, educational levels were not controlled for between age groups for any of the three data sets, and thus it is not possible to know the extent to which this or other confounding sample factors may have obscured an effect in the Abikoff et al. study, or produced a coinciding artifact of a non-linear inverted-U pattern in the latter two studies.

For the *Rey Auditory-Verbal Learning Test (RAVLT)* (see Figure 7.7a, based on data in Tables 7.3 and 7.4), the comparisons between standard deviations across all groups for the

Geffen et al., (1990, in Spreen and Strauss, 1991) data set for males revealed a significant effect for Trial 1 and the Recognition trials ($p<0.0093$ and $p<0.01$, respectively), and similarly for females there was an effect which only just missed significance for Trial 1 and a significant effect for Recognition ($p<0.0543$ and $p<0.0021$, respectively). Entirely analogously for both males and females there were no significant effects for the RAVLT Trial 5, Total and Delayed recall, although in several instances, especially for females, the trends for these submodalities were similar to the overall trends across age groups for Trial 1 (see Figure 7.7a, top graph) which *did* reach significance as follows. For males, backwards

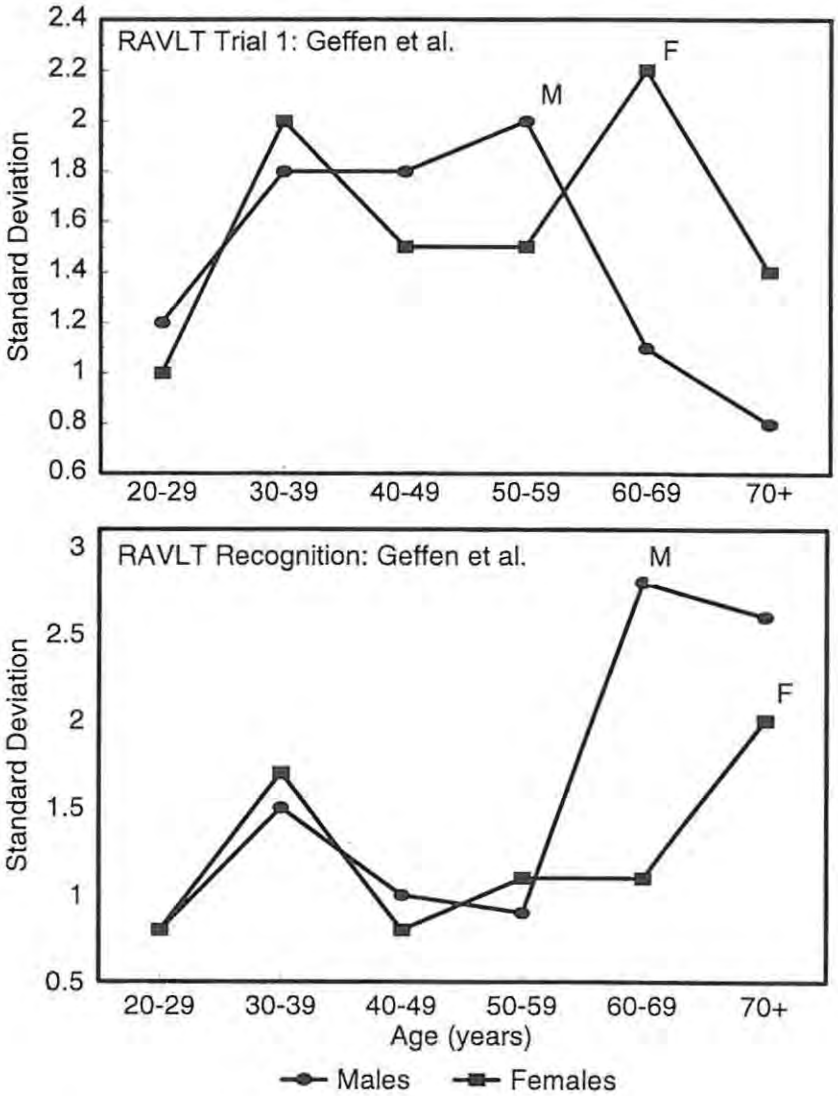


Figure 7.7a. Rey Auditory-Verbal Learning Test (RAVLT) Trial 1 and Recognition: Standard deviations for males (M) and females F) across adult age groups (data from Geffen et al., 1990, in Spreen & Strauss, 1991). The points at which variability peaks for males and females are denoted by M and F, respectively.

comparisons revealed that the standard deviation of the oldest (70+) age group was no different from the 60's age group (p=0.1783), was significantly lower than that of the 50's, 40's and 30's age groups (p=0.0055**, p=0.0114*, 0.0120*, respectively), and was no different from the 20's age group, which indicates a significant non-linear pattern of variability with an initial increase in variability in the 30's which was sustained through the 40's and 50's, and a subsequent decrease in the 60's which was sustained in the 70+ age group; for females forwards comparisons revealed no significant difference between the 20's age group and the 30's, 40's, 50's and 70+ age groups, but a significantly lower standard deviation in the 20's compared with the 60's age group (p=0.0134*). Overall this indicates that for the RAVLT Trial 1 there was a highly consistent non-linear inverted-U pattern of variability for both males and females, that occurred later for females than for males.

Thus for the RAVLT Trial 1 a left versus right shuttle pattern in variability was evident for gender in association with aging in that there was an earlier peak in the 50's for males followed by a sustained decline from the 60's onwards (left shuttle effect), whereas there was a later variability peak in the 60's for females followed by a decline only in the 70+ age group (right shuttle effect). This effect is illustrated in Figure 7.7a (upper graph) in which the significant peaks in variability immediately prior to significant declines in variability are denoted as M and F for males and females, respectively. The trends on the RAVLT data for Total recall provided additional support for an earlier increase in variability for males over females (see Table 7.3), although the effect for Total recall did not reach significance on the all-groups analyses. (The apparent peak in variability for females in the 30's age group relative to the 20's and 30's age groups was *not significant*, and thus does not contradict the overall pattern of a late increase in variability for females).

On the *RAVLT-Recognition* task (see Figure 7.7a lower graph) backwards comparisons for males revealed no significant difference between the 70+ age group and the 60's and 30's age groups, but a significantly higher standard deviation between the 70+ age group and the 50's, 40's and 20's age groups (p=0.0013***, p=0.0031** and p=0.0008***, respectively). For females backwards comparisons revealed that there was no significant difference between the 70+ age group and the 60's and 30's age group, but a significantly higher standard deviation between the 70+ age group and the 40's and 20's age group (p=0.0042**, and 0.0059**, respectively). These results indicate that for both males and females there was a significant peak in variability in older age which occurred in the 60's age decade for males and only in the 70+ age stage for females. Thus as with Trial 1 of the RAVLT, for Recognition there was support for a left versus right shuttle pattern of variability in association with aging in that the onset of increased variability occurred earlier for males (left shuttle effect), compared with a later onset of increased variability for females (right shuttle effect). This effect is illustrated in Figure 7.7a (lower graph) in which the points at which there are significant peaks in variability are denoted as M and F for males and females, respectively. In contrast to Trial 1 for which there was a non-linear pattern of an initial increase and subsequent decrease in variability in association with aging for both males and females, for Recognition only a pattern of significantly increased variability was evident, and there was no significant dedifferentiation effect noted for either males or females. An isolated additional indication from the pairwise comparisons for the RAVLT Recognition task (as reported on above) was that for both males and females there was the indication of a *significant* peak in variability in the 30's age group.

The second RAVLT data set (*Mitrushina et al., 1991*) which tapped the older age groups from the 50's through to the mid-80's, reflected identical overall trends across the older age ranges for Trial 1 and the Recognition trial as were noted for females in the over 50's age groups in the Geffen et al., (1990) data set (see Figure 7.7b upper and lower graphs based on data in Tables 7.3 and 7.4), although only the effect for the Recognition trial reached significance ($p<0.01$). As for the Geffen et al. data, the pattern of variability for the Mitrushina et al. data for Trial 1 (although the statistical comparisons across all groups were not significant) was a non-linear one of an initial increase and subsequent decrease in variability. For the Recognition trial, backwards comparisons for the Mitrushina et al. data revealed significantly higher standard deviations between the oldest and two youngest age groups ($p=0.0007^{***}$, and 0.0282^{**}) denoting, as was the case also for the Recognition trial of the data set in the Geffen et al. data set, a significant pattern of progressively increasing variability up to the oldest age group, with no evidence of a subsequent dedifferentiation effect.

For the *Buschke Selective Reminding Test - BSRT* (Larrabee, 1988 et al., in Spreen & Strauss, 1991) (see Figure 7.8, based on the data in Tables 7.3 and 7.4) the comparisons of standard deviations across age groups revealed a highly significant effect for all levels of presentation of the test ($p<0.01$ in each case) including Total recall, Delayed recall, RLTR (recall twice in a row) and CLTR (continuous recall). For BSRT Total and Delayed recall

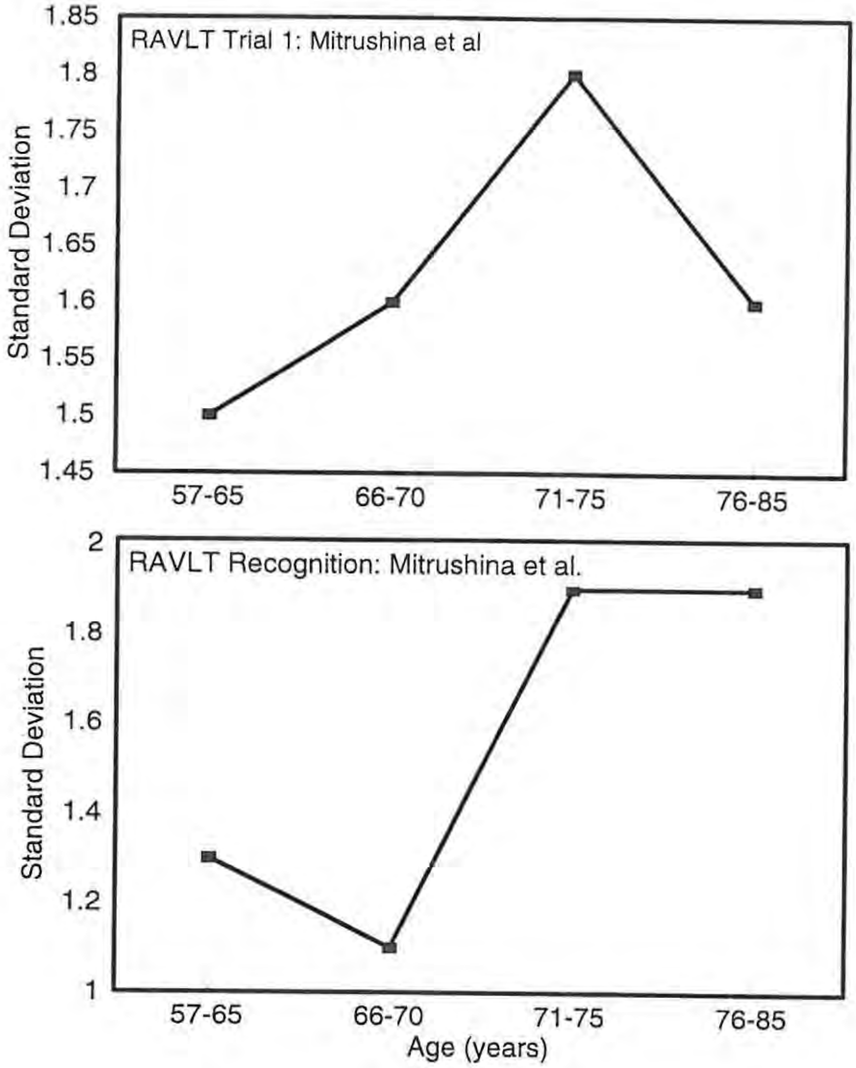


Figure 7.7b. Rey Auditory-Verbal Learning Test (RAVLT) Trial 1 and Recognition: Standard deviations across adult age groups (data from Mitrushina et al., 1991).

backwards comparisons revealed no difference between the standard deviations of the 80's and each of the 70's, 60's and 30's age groups, but significantly higher standard deviations in the 80's age groups than in each of the 50's and 20's age groups ($p=0.0079^{**}$, 0.001^{***}; and $p=0.0032^{**}$, $p=0.001^{***}$, respectively for Total and Delayed recall); for Delayed recall there was also a significantly higher standard deviation in the 80's than in the 40's age group ($p=0.0018^{**}$). This indicates (see Figure 7.8, upper graph), that for both Total and Delayed recall there was a highly consistent pattern of significantly increased variability from the 60's age group compared with the youngest age group and no subsequent decrease in later old age. In both cases (as was evident for the Geffen et al. RAVLT Recognition data set) there was also a significant increase at the 30's age group and subsequent significant decrease in variability at the 50's age group.

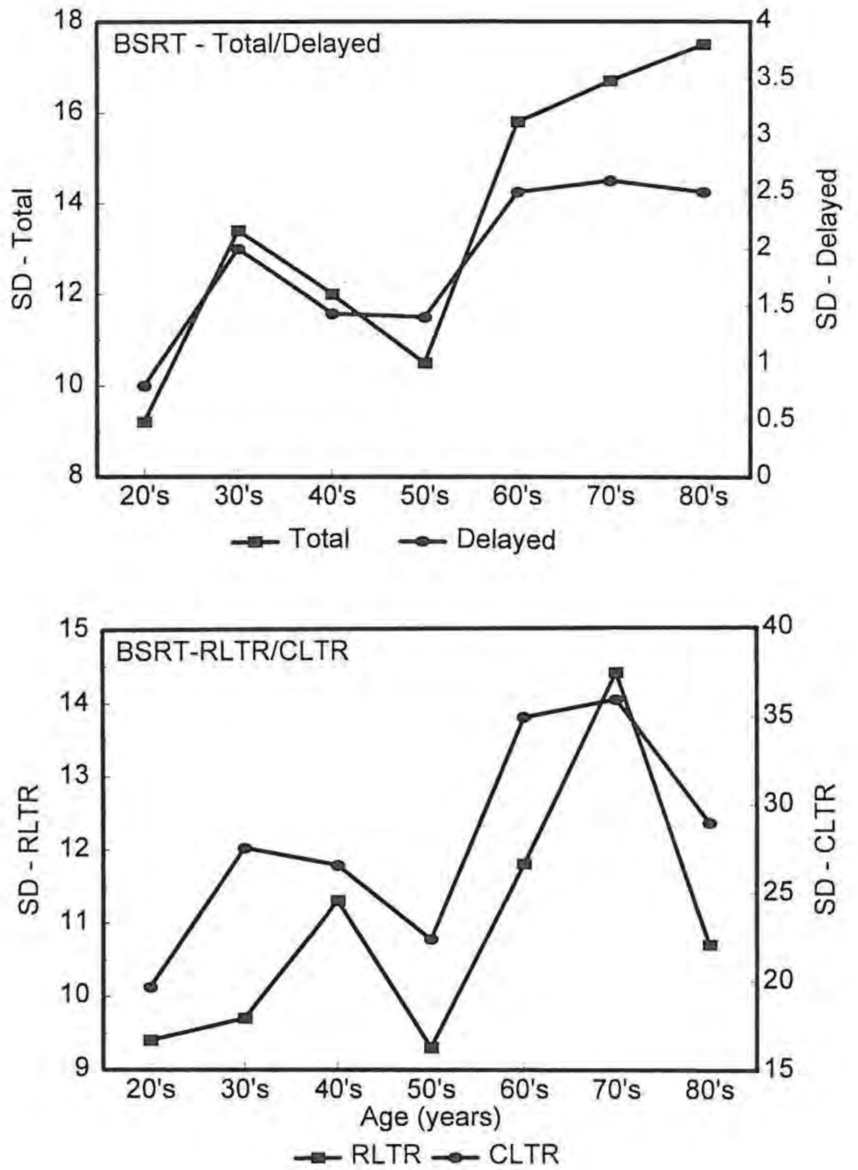


Figure 7.8. Buschke Selective Reminding Test (BSRT) Total, Delayed, RLTR and CLTR: Standard deviations across adult age groups (data from Larrabee et al., 1988, in Spreen & Strauss, 1991).

For BSRT-RLTR forwards comparisons revealed a significantly higher standard deviation for only the 70's age group in comparison with the 20's age group ($p=0.0012^{***}$), and no other significant differences between age groups; for BSRT-CLTR backwards comparisons revealed a significantly higher standard deviation for the 80's age group compared with the 20's age group ($p=0.0094^{*}$), and no other significant differences between age groups. This indicates that for BSRT-RLTR and CLTR (see Figure 7.8, lower graph) there was a highly significant increase in standard deviations by the 60's and 70's age groups compared with the 20's age group consistent with BSRT Total and Delayed recall. However, in contrast with

BSRT Total and Delayed recall which showed a pattern of progressively increasing standard deviations in the older age groups without a later old age decrease in variability, there was a highly consistent non-linear pattern with a peak in standard deviations at the 70's age group followed by a subsequent decrease in standard deviations at the 80's age group for both RLTR and CLTR. This pattern was highly significant for the RLTR trial ($p=0.0012^{***}$), although the tendency did not reach significance for the CLRT trial. The indication of an earlier increase in standard deviations at the 30's age group for RLTR, and at the 40's age group for CLTR, followed by a dip in the variability at the 50's age group, was *not significant* for either RLTR or CLTR.

Overall these findings were strongly indicative of a task-specific left versus right shuttle effect in the patterns of variability for the BSRT task. It is likely that the cumulative learning capacity which is called upon in the RLTR and CLTR tasks is more age sensitive than haphazard memory span as required in the Total and Delayed modalities. Thus a dedifferentiation effect occurred by the 80's decade for the more age-sensitive BSRT RLTR and CLTR tasks (left shuttle effect), and was not in evidence by the 80's decade for the BSRT Total and Delayed tasks (right shuttle effect).

7.2.2 Comparisons of means

For the modality of verbal memory (see Table 7.3 for a summary of mean scores across all groups, and Table 7.4 for pairwise statistical comparisons between the oldest and youngest age groups) there was a highly consistent overall trend across all data sets for average test performance to decline between the youngest and oldest age groups. In all instances for which statistical analyses were conducted, with the exception of the RAVLT Recognition trial for two data sets (females in Geffen et al. 1990, in Spreen & Strauss, 1991; Mitrushina et al., 1991), there were highly significant differences between the mean scores of the youngest and oldest age groups in the direction which indicated deteriorated performance in the oldest relative to the youngest age groups. For verbal memory, the recurring significant variability patterns that were identified in association with an overall significant decline in average performance are illustrated selectively for WMS Associate Learning and Logical Memory, the RAVLT-Trial 1 and Recognition, and the BSRT Delayed recall and CLTR (see Figures 7.9, 7.10, and 7.11, respectively).

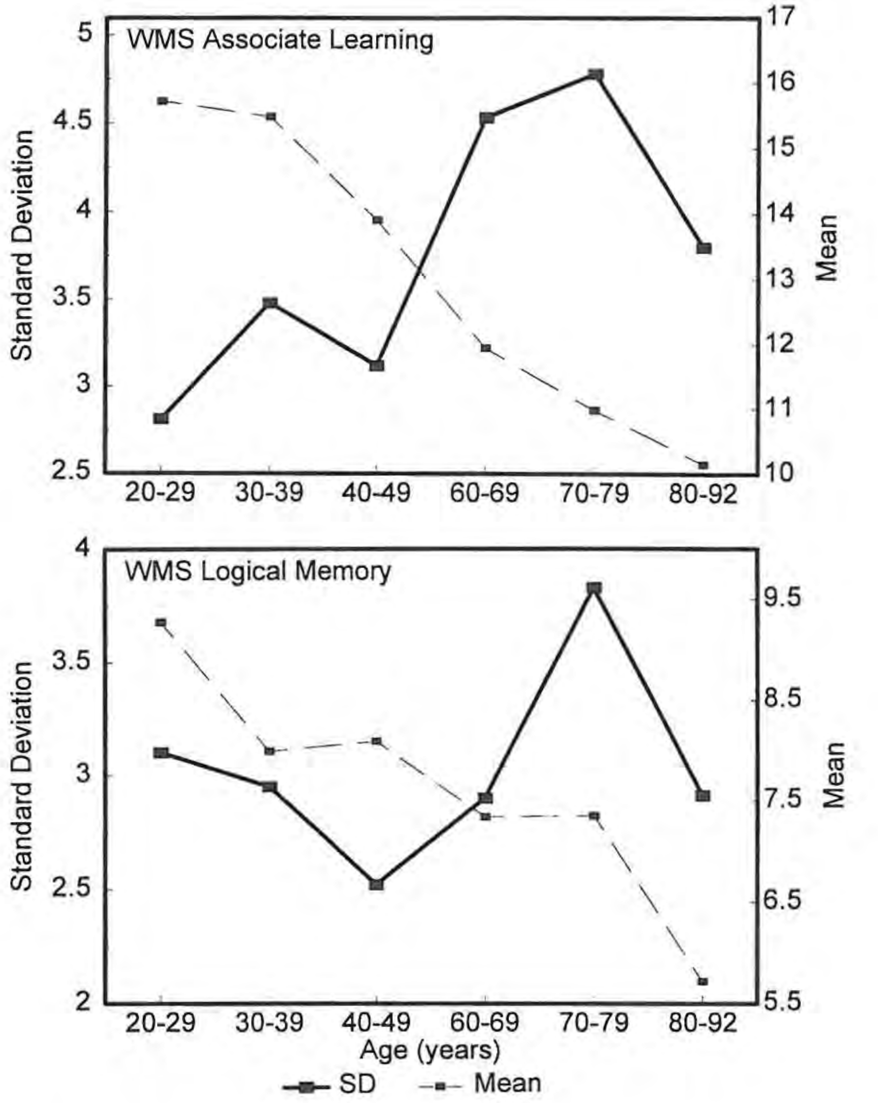


Figure 7.9. Wechsler Memory Scale (WMS) Associate Learning and Logical Memory: Standard deviations and mean scores across adult age groups (collated data in Lezak, 1983) illustrating non-linear inverted-U patterns of variability in association with declining mean scores.

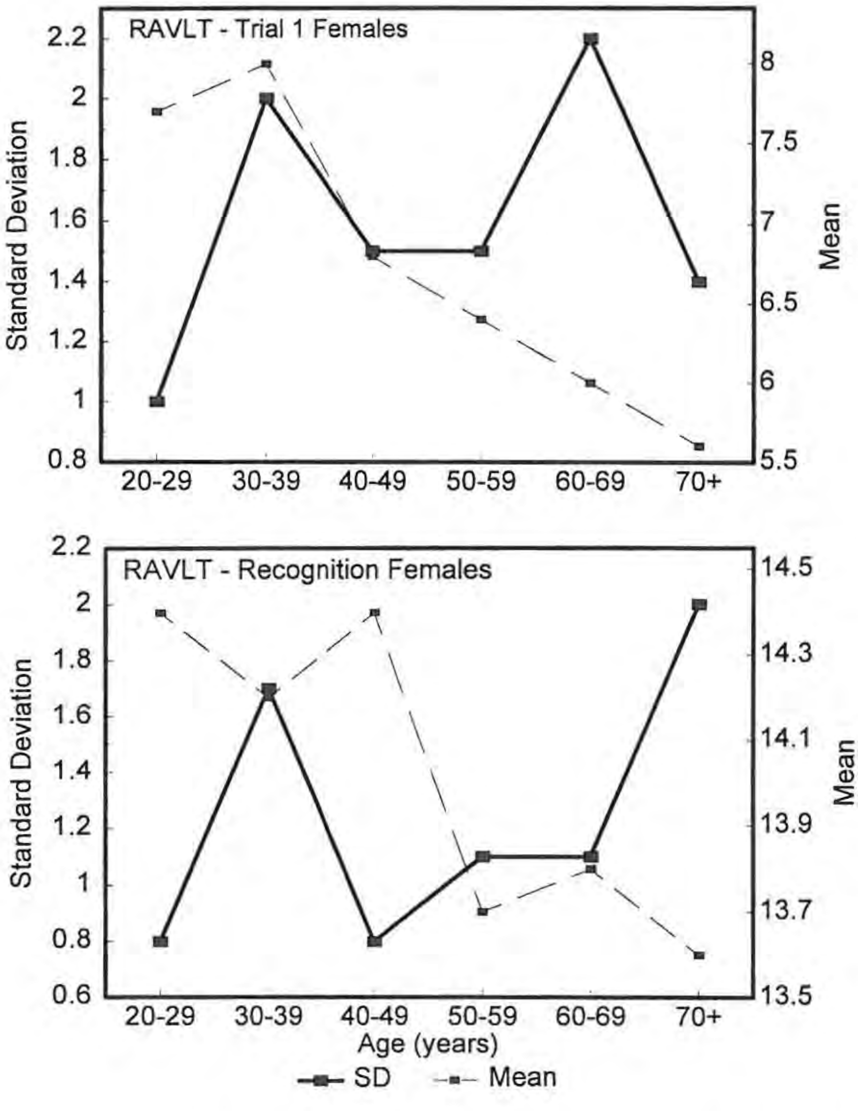


Figure 7.10. Rey Auditory-Verbal Learning Test (RAVLT) Trial 1 and Recognition: Standard deviations and mean scores across adult age groups (data from Geffen et al., 1990, in Spreen & Strauss, 1991) illustrating non-linear inverted-U (top graph) and linear (lower graph) patterns of variability in the later adult years in association with declining mean scores.

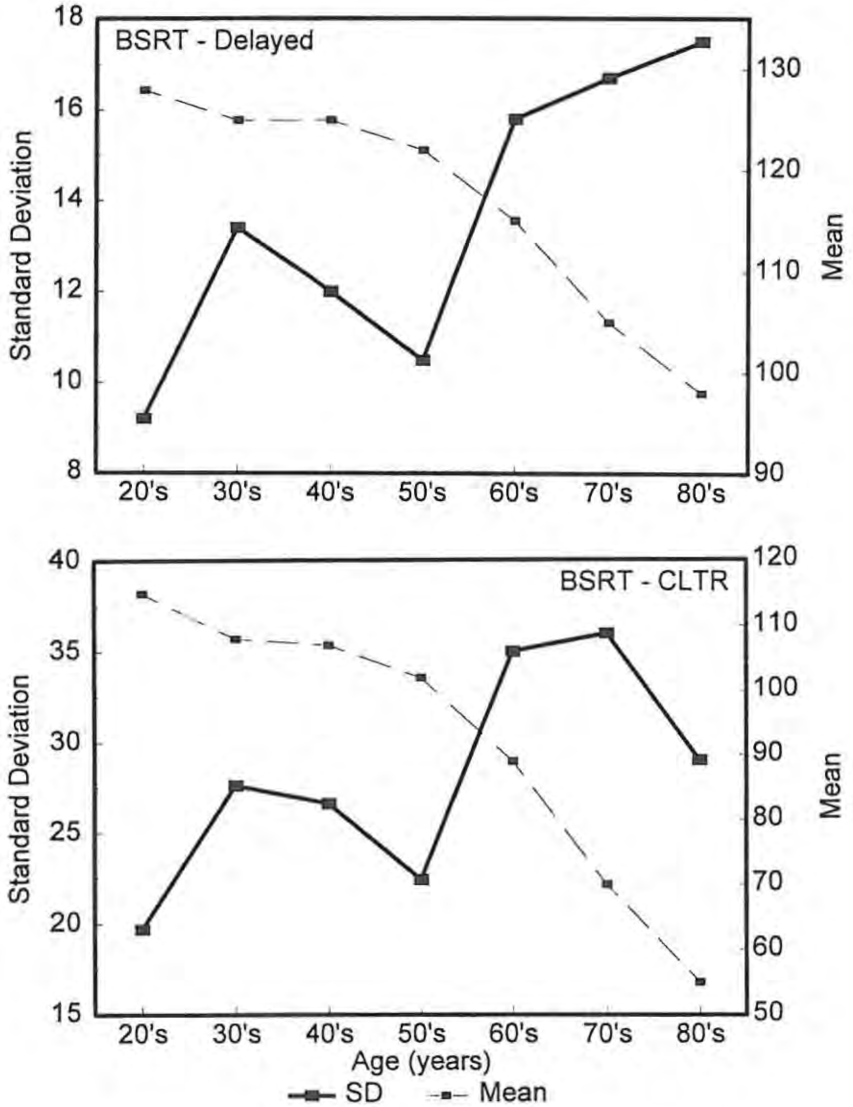


Figure 7.11. Buschke Selective Reminding Test (BSRT) Delayed, and CLTR: Standard deviations and mean scores across adult age groups (data from Larrabee al., 1988, in Spreen & Strauss, 1991) illustrating linear (top graph) and non-linear inverted-U (lower graph) patterns of variability in the later adult years in association with declining mean scores.

Of note when making reference to Figure 7.9 (upper graph), Figure 7.10 (upper and lower graphs) and Figure 7.11 (upper and lower graphs), is that for verbal memory when the decline in average performance was clearly apparent from the 50's age groups onwards, this occurred in association with the robust finding of either a highly significant progressive increase in variability, or a significant increase followed by a significant decrease in variability. However, the apparent tendency for an initial bulge in variability to occur in the 30's age group for several of the verbal memory tasks, and which was a significant effect for the RAVLT Recognition and BSRT Delayed and Total submodalities, did *not* occur in association with a clear decline in average performance. Rather (see Figures 7.9, 7.10 and 7.11) it occurred in association with relatively stable, fluctuating or even increasing tendencies in the progression of mean scores. Since this tendency occurred across a number of different studies for different verbal memory tasks it is not likely to be the consequence of a sample artifact. Rather, considered in conjunction with the stable means this tendency indicates that there was a ceiling effect for the youngest 20's age groups, and that this was 'released' in the early 30's when a number of individuals started to fall away from the ceiling in sufficient number to increase the standard deviation, but not in sufficient number to effect the mean. The artifact of the proposed ceiling effect disappeared once the mean scores showed clearly pronounced decline. As is evident from mean scores generally, recognition tasks in particular are clearly prone to ceiling effects in high functioning individuals in younger age groups. Similarly this would be the case for a test with a relatively low information load such as the WMS Associate Learning test. By contrast a ceiling effect is not likely for the WMS Logical Memory test for which individuals seldom perform close to the ceiling. Accordingly the proposed release of a ceiling effect in the 30's age group resulting in a sporadic increase in variability at this stage, was *not* apparent for Logical Memory (see Figure 7.9), and should not therefore be mistaken for a pervasive variability effect for verbal memory tasks due to aging.

7.2.3 Verbal memory: Summary of effects

Generally for verbal memory, as tapped by the WMS Associate Learning and WMS Logical Memory tests, the RAVLT and the BSRT, there was support for a significant increase in variability with advancing age by the 60's and 70's age decades that is the initial differentiation stage of the shuttle bulge effect. Further there was evidence of a subsequent significant decrease in variability by the 80's age stage (that is the full shuttle bulge effect of initial differentiation and subsequent dedifferentiation) in the case of WMS Logical Memory, RAVLT Trial 1, and the BSRT-RLTR, and probably also for WMS Associate Learning and BSRT-CLTR (for which strong trends in this direction were noted). A number of initial increases in variability in the 30's age decade are likely to have been an artifact of a released ceiling effect from the earlier 20's decade.

Over a number of verbal memory data sets two clear shuttle shift effects were detected in the variability data, one for gender and one for age-sensitivity of tasks. The data for the RAVLT Trial 1 and Recognition supported a shuttle shift effect for gender, in that for males variability effects on both tasks occurred earlier than for females (left versus right shuttle effect, respectively). Both the RAVLT and the BSRT data supported a shuttle shift effect for age-sensitivity of tasks in that a non-linear pattern of an initial increase and subsequent decrease in variability was apparent for the relatively more age-sensitive submodality <u>within</u> the RAVLT of Trial 1 short-term verbal memory (left shuttle effect), and only a progressive increase in variability without a subsequent decrease was apparent for the relatively less-age sensitive submodality <u>within</u> the RAVLT of recognition (right shuttle effect); similarly a non-linear pattern of an initial increase and subsequent decrease in variability was apparent for the relatively more age-sensitive submodalities <u>within</u> the BSRT which tapped consistent recall (left shuttle effect), and only a progressive increase in variability without a subsequent decrease was apparent for the relatively less-age sensitive modalities <u>within</u> the BSRT of total and delayed recall (right shuttle effect).

All of the above variability effects for verbal memory occurred in association with a significant decline in average performance in association with older age, and with broadly equivalent onset. Proposed artifactual increases in variability due to the release of a ceiling effect in the early adult years for a number of tasks were accompanied by the clear absence of a decline in average performance.

Table 7.3. MEMORY (VERBAL): Mean Scores and Comparisons of Standard Deviations across all Age Groups.

WMS Associate Learning (Strauss & Spreen 1989, in Spreen & Strauss 1991)									
AGE GROUP	20-29	30-39	40-49	50-59	60-69	70+	df	B[1]	p
n	23	13	14	12	12	14	-	-	-
Mean	18.33	18.21	18.29	17.30	14.30	15.89	-	-	-
SD	1.45	2.28	2.82	3.34	2.52	3.10	5	13.20	<0.01**

WMS Associate Learning (Collated data in Lezak 1983)									
AGE GROUP	20-29	30-39	40-49	60-69	70-79	80-92	df	B	p
n	50	53	46	70	46	115	-	-	-
Mean	15.72	15.48	13.91	11.94	10.98	10.15	-	-	-
SD	2.81	3.48	3.12	4.53	4.78	3.80	5	21.29	<0.01**

[1]Bartlett's test statistic.

Table 7.3 *continued*

Table 7.3. MEMORY (VERBAL): Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

WMS Logical Memory (Abikoff et al. 1987, in Spreen & Strauss 1991)									
AGE GROUP	18-29	30-39	40-49	50-59	60-69	70+	df	B	p
n	74	67	41	54	56	46	-	-	-
WMS Log Mem-Immed									
Mean	22.99	24.57	23.44	23.63	20.48	19.11	-	-	-
SD	6.66	6.97	5.01	6.14	6.42	6.74	5	5.72	0.0757⁺
WMS Log Mem-Del									
Mean	19.84	22.16	21.07	20.13	17.34	15.33	-	-	-
SD	6.67	7.57	5.91	6.48	6.71	7.57	5	4.38	0.1436

WMS Logical Memory (Van Gorp et al. 1990)							
AGE GROUP	57-65	66-70	71-75	76-85	df	B	p
n	28	45	57	26	-	-	-
WMS Log Mem-Immed							
Mean	9.75	8.47	9.27	8.05	-	-	-
SD	2.40	2.65	3.15	2.19	3	5.47	0.0728⁺
WMS Log Mem-Del							
Mean	7.80	5.73	6.31	5.32	-	-	-
SD	2.16	2.24	3.16	2.48	3	8.23	0.0187**

WMS Logical Memory (Collated data in Lezak 1983)									
AGE GROUP	20-29	30-39	40-49	60-69	70-79	80-92	df	B	p
n	50	53	46	70	46	115	-	-	-
WMS Log Mem-Immed									
Mean	8.40	6.00	6.40	6.50	4.90	3.60	-	-	-
SD	3.10	2.95	2.52	2.90	3.83	2.91	-	9.20	0.0141*

Table 7.3 *continued*

Table 7.3. **MEMORY (VERBAL)**: Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

RAVLT - Males (Geffen et al. 1990, in Spreen & Strauss 1991)									
AGE GROUP	20-29	30-39	40-49	50-59	60-69	70+	df	B	p
n	10	10	11	11	10	10	-	-	-
RAVLT-Trial 1									
Mean	8.4	6.0	6.4	6.5	4.9	3.6	-	-	-
SD	1.2	1.8	1.8	2.0	1.1	0.8	5	10.06	0.0093*
RAVLT-Trial 5									
Mean	12.2	11.4	10.9	11.8	8.9	8.2	-	-	-
SD	2.2	2.6	2.0	2.6	2.0	2.5	5	1.39	0.5724
RAVLT-Total									
Mean	54.9	46.0	47.5	47.6	36.7	32.6	-	-	-
SD	7.0	10.9	8.3	8.5	8.4	8.3	5	1.84	0.4673
RAVLT-Del									
Mean	10.6	10.4	10.5	10.0	7.1	5.6	-	-	-
SD	2.4	2.3	2.7	2.6	3.8	2.6	5	3.16	0.2542
RAVLT-Recog									
Mean	14.2	13.5	14.2	13.9	12.4	11.5	-	-	-
SD	0.8	1.5	1.0	0.9	2.8	2.6	5	26.00	<0.01**

Table 7.3 *continued*

Table 7.3. MEMORY (VERBAL): Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

RAVLT - Females (Geffen et al. 1990, in Spreen & Strauss 1991)									
AGE GROUP	20-29	30-39	40-49	50-59	60-69	70+	df	B	p
n	10	10	11	11	10	10	-	-	-
RALVT-Trial 1									
Mean	7.7	8.0	6.8	6.4	6.0	5.6	-	-	-
SD	1.0	2.0	1.5	1.5	2.2	1.4	5	6.41	0.0543⁺
RAVLT-Trial 5									
Mean	12.9	12.7	12.8	11.6	11.9	10.1	-	-	-
SD	1.5	1.3	1.4	2.1	1.6	1.2	5	3.90	0.1841
RAVLT-Total									
Mean	55.3	55.9	52.1	47.6	49.0	41.6	-	-	-
SD	6.6	6.3	7.1	7.7	7.1	6.6	5	0.48	0.8436
RAVLT-Del									
Mean	11.0	12.2	11.1	10.2	10.3	8.3	-	-	-
SD	2.0	2.5	2.3	2.7	2.3	2.1	5	1.10	0.6505
RAVLT-Recog									
Mean	14.4	14.2	14.4	13.7	13.8	13.6	-	-	-
SD	0.8	1.7	0.8	1.1	1.1	2.0	5	13.14	0.0021*

Table 7.3 *continued*

Table 7.3. MEMORY (VERBAL): Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

RAVLT (Mitrushina et al. 1991)							
AGE GROUP	57-65	66-70	71-75	76-85	df	B	p
n	28	45	57	26	-	-	-
RAVLT-Trial 1							
Mean	6.4	5.9	5.1	5.1	-	-	-
SD	1.5	1.6	1.8	1.6	3	1.46	0.5117
RAVLT-Trial 5							
Mean	12.1	11.5	10.3	9.7	-	-	-
SD	2.4	3.0	2.9	2.8	3	1.68	0.4599
RAVLT-Recog							
Mean	13.2	13.0	12.7	12.6	-	-	-
SD	1.3	1.1	1.9	1.9	3	16.98	<0.01**

BSRT (Larrabee et al. 1988 in Spreen & Strauss 1991)										
AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	B	p
n	51	29	31	24	50	59	27	-	-	-
BSRT-Total										
Mean	128	125	125	122	115	105	98	-	-	-
SD	9.2	13.4	12.0	10.5	15.8	16.7	17.5	6	26.46	<0.01**
BSRT-Del										
Mean	11.5	10.7	11.0	10.6	9.6	9.1	6.4	-	-	-
SD	0.8	2.0	1.4	1.4	2.5	2.6	2.5	6	76.85	<0.01**
BSRT-RLTR										
Mean	8.1	10.1	11.2	10.8	14.7	20.7	22.2	-	-	-
SD	9.4	9.7	11.3	9.3	11.8	14.4	10.7	6	13.87	<0.01**
BSRT-CLTR										
Mean	115	108	107	102	89	70	55	-	-	-
SD	19.7	27.6	26.6	22.4	35	36	29	6	24.67	<0.01**

Significant difference (*$p<0.05$; **$p<0.01$); Approaching significance ($^{+}$ $p<0.10$).

Table 7.4. MEMORY (VERBAL): Pairwise Comparisons of Means between the oldest and youngest Age Groups, and of Standard Deviations between the oldest Age Group and each younger Age Group unless otherwise specified.

WMS Associate Learning (Strauss & Spreen 1989, in Spreen & Strauss 1991)									
AGE GROUP	20-29	30-39	40-49	50-59	60-69	70+	df	t/F	p
n	23	13	14	12	14	14	-	-	-
Mean	18.33	-	-	-	-	15.89	35	3.26	0.0013***
SD	-	-	-	-	2.52	3.10	13,13	0.66	0.2327
SD	-	-	-	3.34	-	3.10	11,13	1.17	0.3941
SD	-	-	2.82	-	-	3.10	13,13	1.25	0.3455
SD	-	2.28	-	-	-	3.10	12,13	0.54	0.1482
SD	1.45	-	-	-	-	3.10	22,13	0.22	0.0009***

WMS Associate Learning (Collated data in Lezak 1983)									
AGE GROUP	20-29	30-39	40-49	60-69	70-79	80-92	df	t/F	p
n	50	53	46	70	46	115	-	-	-
Mean	15.72	-	-	-	-	10.15	163	9.31	<0.001***
SD	-	-	-	-	4.78	3.80	45,114	1.58	0.0269
SD	-	-	-	4.53	-	3.80	69,114	1.42	0.0483
SD	-	-	3.12	-	-	3.80	45,114	0.67	0.0674
SD	-	3.48	-	-	-	3.80	52,114	0.84	0.2410
SD	2.81	-	-	-	-	3.80	49,114	0.55	0.0094**

Table 7.4 *continued*

Table 7.4. MEMORY (VERBAL): Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

WMS Logical Memory (Abikoff et al. 1987, in Spreen & Strauss 1991)									
AGE GROUP	18-29	30-39	40-49	50-59	60-69	70+	df	t/F	p
n	74	67	41	54	56	46	-	-	-
WMS Log Mem-Immed									
Mean	22.99	-	-	-	-	19.11	118	3.09	0.0012***
SD	-	-	-	-	6.42	6.74	55,45	0.91	0.3630
SD	-	-	-	6.14	-	6.74	53,45	0.83	0.2559
SD	-	-	5.01	-	-	6.74	40,45	0.55	0.0295
SD	-	6.97	-	-	-	6.74	66,45	1.07	0.4104
SD	6.66	-	-	-	-	6.74	73,45	0.98	0.4562
WMS Log Mem-Del									
Mean	19.84	-	-	-	-	15.33	118	3.42	<0.001***
SD	-	-	-	-	6.71	7.57	55.45	0.79	0.1960
SD	-	-	-	6.48	-	7.57	53.45	0.73	0.1380
SD	-	-	5.91	-	-	7.57	40.45	0.61	0.0571
SD	-	7.57	-	-	-	7.57	66.45	1.00	0.5069
SD	6.67	-	-	-	-	7.57	73.45	0.78	0.1661

Table 7.4 *continued*

Table 7.4. **MEMORY (VERBAL):** Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

WMS Logical Memory (van Gorp et al. 1990)							
AGE GROUP	57-65	66-70	71-75	76-85	df	t/F	p
n	28	45	57	26	-	-	-
WMS Log Mem-Immed							
Mean	9.75	-	-	8.05	52	2.71	0.0045**
SD	-	-	3.15	2.19	56,25	2.07	0.0245*
SD	-	2.65	-	2.19	44,25	1.46	0.1551
SD	2.40	-	-	2.19	27,25	1.20	0.3241
WMS Log Mem-Del							
Mean	7.80	-	-	5.32	52	3.93	<0.001***
SD	-	-	3.16	2.48	56,25	1.62	0.0927
SD	-	2.24	-	2.48	44,25	0.82	0.2716
SD	2.16	-	-	2.48	27,25	0.76	0.2412
Forwards Comparisons							
WMS Log Mem-Immed							
SD	2.40	2.65	-	-	27,44	0.82	0.2958
SD	2.40	-	3.15	-	27,56	0.58	0.0624
SD	2.40	-	-	2.19	27,25	1.20	0.3241
WMS Log Mem-Del							
SD	2.16	2.24	-	-	27,44	0.93	0.4283
SD	2.16	-	3.16	-	27,56	0.47	0.0167*
SD	2.16	-	-	2.48	27,25	0.76	0.2412

WMS Logical Memory (Collated date in Lezak 1983)									
AGE GROUP	20-39	30-39	40-49	60-69	70-79	80-92	df	t/F	p
n	50	53	46	70	46	115	-	-	-
WMS Log Mem-Immed									
Mean	9.28	-	-	-	-	5.72	163	7.08	<0.001***
SD	-	-	-	-	3.83	2.91	45,115	1.73	0.0104*
SD	-	-	-	2.90	-	2.91	69,115	0.99	0.4943
SD	-	-	2.52	-	-	2.91	45,115	0.75	0.1376
SD	-	2.95	-	-	-	2.91	52,115	1.03	0.4424
SD	3.10	-	-	-	-	2.91	49,115	1.13	0.2884

Table 7.4 *continued*

Table 7.4. MEMORY (VERBAL): Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

RAVLT - Males (Geffen et al. 1990, in Spreen & Strauss)									
AGE GROUP	20-29	30-39	40-49	50-59	60-69	70+	df	t/F	p
n	10	10	11	11	10	10	-	-	-
RAVLT-Trial 1									
Mean	8.4	-	-	-	-	3.6	18	10.52	<0.001***
SD	-	-	-	-	1.1	0.8	9,9	1.89	0.1783
SD	-	-	-	2.0	-	0.8	10,9	6.25	0.0055**
SD	-	-	1.8	-	-	0.8	10,9	5.06	0.0114*
SD	-	1.8	-	-	-	0.8	9,9	5.06	0.0120*
SD	1.2	-	-	-	-	0.8	9,9	2.25	0.1214
RAVLT-Recog									
Mean	14.2	-	-	-	-	11.5	18	3.14	0.0028**
SD	-	-	-	-	2.8	2.6	9,9	1.16	0.4075
SD	-	-	-	0.9	-	2.6	10,9	0.12	0.0013***
SD	-	-	1.0	-	-	2.6	10,9	0.15	0.0031**
SD	-	1.5	-	-	-	2.6	9,9	0.33	0.0564
SD	0.8	-	-	-	-	2.6	9,9	0.09	0.0008***

Table 7.4 *continued*

Table 7.4. MEMORY (VERBAL): Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

RAVLT - Females (Geffen et al. 1990, in Spreen & Strauss) AGE GROUP	20-29	30-39	40-49	50-59	60-69	70+	df	t/F	p
n	10	10	11	11	10	10			
RAVLT-Trial 1									
Mean	7.7	-	-	-	-	5.6	18	3.86	0.0057**
SD	-	-	-	-	2.2	1.4	9,9	2.47	0.0971
SD	-	-	-	1.5	-	1.4	10,9	1.15	0.4227
SD	-	-	1.5	-	-	1.4	10,9	1.15	0.4227
SD	-	2.0	-	-	-	1.4	9,9	2.04	0.1514
SD	1.0	-	-	-	-	1.4	9,9	0.51	0.1653
RAVLT-Recog									
Mean	14.4	-	-	-	-	13.6	18	1.17	0.1278
SD	-	-	-	-	1.1	2.0	9,9	0.30	0.0448
SD	-	-	-	1.1	-	2.0	10,9	0.30	0.0381
SD	-	-	0.8	-	-	2.0	10,9	0.16	0.0042**
SD	-	1.7	-	-	-	2.0	9,9	0.72	0.3180
SD	0.8	-	-	-	-	2.0	9,9	0.16	0.0059**
Forwards Comparisons									
RAVLT-Trial 1									
SD	1.0	2.0	-	-	-	-	9,9	0.25	0.0255
SD	1.0	-	1.5	-	-	-	9,10	0.44	0.1189
SD	1.0	-	-	1.5	-	-	9,10	0.44	0.1189
SD	1.0	-	-	-	2.2	-	9,9	0.21	0.0134*
SD	1.0	-	-	-	-	1.4	9,9	0.51	0.1653
RAVLT-Recog									
SD	0.8	1.7	-	-	-	-	9,9	0.22	0.0175*
SD	0.8	-	0.8	-	-	-	9,9	1.00	0.4954
SD	0.8	-	-	1.1	-	-	9,9	0.53	0.1761
SD	0.8	-	-	-	1.1	-	9,9	0.53	0.1783
SD	0.8	-	-	-	-	2.0	9,9	0.16	0.0059**

Table 7.4 *continued*

Table 7.4. MEMORY (VERBAL): Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

RAVLT (Mitrushina et al. 1991)							
AGE GROUP	57-65	66-70	71-75	76-85	df	t/F	p
n	28	45	57	26	-	-	-
RAVLT-Recog							
Mean	13.2	-	-	12.6	52	1.36	0.0894
SD	-	-	1.9	1.9	56,25	1.00	0.4826
SD	-	1.1	-	1.9	44,25	0.34	0.0007***
SD	1.3	-	-	1.9	27,25	0.47	0.0282*

BSRT (Larrabee 1988 et al. in Spreen & Strauss 1991)										
AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	t/F	p
n	51	29	31	24	50	59	27	-	-	-
BSRT-Total										
Mean	128	-	-	-	-	-	98	76	9.95	<0.001***
SD	-	-	-	-	-	16.7	17.50	58,26	0.91	0.3735
SD	-	-	-	-	15.8	-	17.50	49,26	0.81	0.2636
SD	-	-	-	10.5	-	-	17.50	23,26	0.36	0.0079**
SD	-	-	12.0	-	-	-	17.50	30,26	0.47	0.0240
SD	-	13.4	-	-	-	-	17.50	28,26	0.59	0.0846
SD	9.2	-	-	-	-	-	17.50	50,26	0.28	<0.001***
BSRT-Del										
Mean	11.5	-	-	-	-	-	8.40	76	8.14	<0.001***
SD	-	-	-	-	-	2.6	2.50	58,26	1.08	0.5752
SD	-	-	-	-	2.5	-	2.50	49,26	1.00	0.4860
SD	-	-	-	1.4	-	-	2.50	23,26	0.31	0.0032**
SD	-	-	1.4	-	-	-	2.50	30,26	0.33	0.0018**
SD	-	2.0	-	-	-	-	2.50	28,26	0.64	0.1248
SD	0.8	-	-	-	-	-	2.50	50,26	0.10	<0.001***

Table 7.4 *continued*

Table 7.4. MEMORY (VERBAL): Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

BSRT (Larrabee 1988 et al. in Spreen & Strauss 1991)										
AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	t/F	p
BSRT-RLTR										
Mean	8.1	-	-	-	-	-	22.2	76	-6.01	<0.001***
SD	-	-	-	-	-	14.4	10.7	58,26	1.81	0.0492
SD	-	-	-	-	11.8	-	10.7	49,26	1.22	0.2997
SD	-	-	-	9.3	-	-	10.7	23,26	0.76	0.2500
SD	-	-	11.3	-	-	-	10.7	30,26	1.12	0.3913
SD	-	9.7	-	-	-	-	10.7	28,26	0.82	0.3051
SD	9.4	-	-	-	-	-	10.7	50,26	0.77	0.2125
BSRT-CLTR										
Mean	115	-	-	-	-	-	55	76	10.82	<0.001***
SD	-	-	-	-	-	36	29	58,26	1.54	0.1138
SD	-	-	-	-	35	-	29	49,26	1.46	0.1515
SD	-	-	-	22.4	-	-	29	23,26	0.60	0.1072
SD	-	-	26.6	-	-	-	29	30,26	0.84	0.3222
SD	-	27.6	-	-	-	-	29	28,26	0.91	0.3977
SD	19.7	-	-	-	-	-	29	50,26	0.46	0.0094*

Table 7.4 *continued*

Table 7.4. MEMORY (VERBAL): Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

BSRT (Larrabee 1988 et al. in Spreen & Strauss 1991) AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	t/F	p
Forwards Comparisons										
BSRT-RLTR										
SD	9.4	9.7	-	-	-	-	-	50,28	0.94	0.4133
SD	9.4	-	11.3	-	-	-	-	50,30	0.69	0.1227
SD	9.4	-	-	9.3	-	-	-	50,23	1.02	0.4939
SD	9.4	-	-	-	11.8	-	-	50,49	0.64	0.0620
SD	9.4	-	-	-	-	14.4	-	50,58	0.43	0.0012***
SD	9.4	-	-	-	-	-	10.7	50,26	0.77	0.2125
BSRT-CLTR										
SD	19.7	27.6	-	-	-	-	-	50,28	0.51	0.0186
SD	19.7	-	26.6	-	-	-	-	50,30	0.55	0.2956
SD	19.7	-	-	22.4	-	-	-	50,23	0.77	0.2206
SD	19.7	-	-	-	35	-	-	50,49	0.32	<0.001***
SD	19.7	-	-	-	-	36	-	50,58	0.30	<0.001***
SD	19.7	-	-	-	-	-	29	50,26	0.46	0.0094*

Significant difference (*p<0.10; **p<0.05; ***p<0.01) as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

7.3 VISUAL MEMORY

For the modality of visual memory the data sets submitted to analysis are collated in Table A (Appendix pp.7-9), and a summary of the educational and/or IQ levels of the age groups appear in Table B (Appendix p.19). The following tests and accompanying sources of normative data sets were included in the analysis (ordered for discussion as they appear in the tables):

(i) **Wechsler Memory Scale (WMS) Visual Reproduction:** Three data sets (data collated from Trahan, Quintana, Willingham, & Goethe, 1988 and Haaland et al., in Spreen & Strauss, 1991; van Gorp, Satz, & Mitrushina, 1990; and data collated from Wechsler, 1945, Hulicka, 1966, and Klonoff & Kennedy, 1966, in Lezak, 1983).

(ii) **Rey Complex Figure Recall (RCF) Delayed:** Three data sets (data collated from Kolb & Whishaw, 1985, and Strauss & Spreen, unpublished study, in Spreen & Strauss, 1991; van Gorp, Satz, & Mitrushina, 1990; and Troyer, Graves, & Cullum, 1994).

(iii) **Benton Visual Retention Test-Revised (BVRT-R):** One data set for males and females with and without a college degree (Robertson-Tehabo & Arenberg, 1989, in Spreen & Strauss, 1991).

(iv) **Rey Visual Design Learning Test (RVDLT):** One data set including Trial 1, Trial 5, Total and Recognition (Strauss & Spreen, unpublished study, in Spreen & Strauss, 1991).

7.3.1 Comparisons of standard deviations across groups

For the *WMS Visual Reproduction* (see Figure 7.12, based on data in Tables 7.5 and 7.6) the comparisons of standard deviations across all groups revealed a highly significant effect for all three data sets included in the study (data in Spreen & Strauss, 1991; van Gorp et al., 1990, and data in Lezak, 1983), for both the Immediate and Delayed recall modalities ($p = < 0.01$ in all instances).

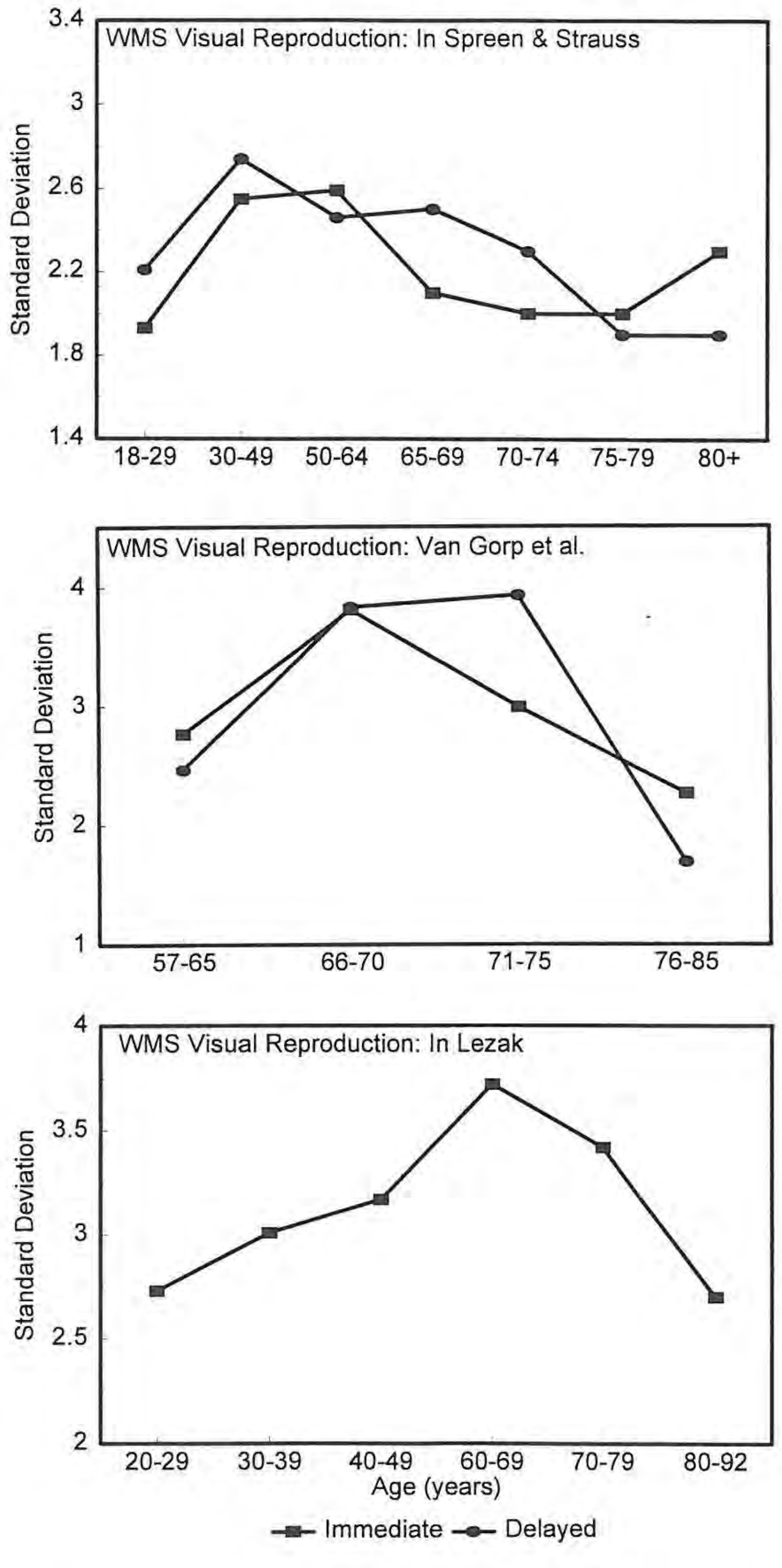


Figure 7.12. Wechsler Memory Scale (WMS) Visual Reproduction: Standard deviations across adult age groups for three data sets (collated data in Spreen & Strauss, 1991; van Gorp et al., 1990; collated data in Lezak, 1983).

For the data set *in Spreen and Strauss (1991)* (Figure 7.12, top graph), forwards comparisons revealed a significantly lower standard deviation for the 18-29 year-old age group in comparison with the 30-49 and 50-64 year-old age groups for Immediate recall ($p=0.0046^{**}$ and 0.0069^{**}), a significantly lower standard deviation for the 18-29 year-old group in comparison with the 30-49 year-old age group for Delayed recall ($p=0.0094^{*}$), and no significant differences between the youngest age group and any of the other age groups for either Immediate or Delayed recall, indicating a significant non-linear inverted-U pattern of variability across the age groups for which the initial significant increase in variability occurred at the 30-49 year-old age group for both levels of presentation. For the second data set (*van Gorp et al., 1990*) (Figure 7.12, middle graph), backwards comparisons revealed a significantly lower standard deviation between the oldest 76-85 year-old age group and the 66-70 year-old age group for Immediate recall ($p=0.0033^{**}$), a significantly lower standard deviation between the oldest 76-85 year-old age group compared with the 71-75 and 66-70 year-old age groups for Delayed recall ($p=<0.001^{***}$, in both instances), and no other significant differences between the oldest age group in each case for Immediate and Delayed recall, and any of the younger age groups. Consistent with the analyses for the Spreen and Strauss data set (Figure 7.12, top graph), the results for the van Gorp et al. analyses (Figure 7.12, middle graph) indicate a significant overall non-linear inverted-U pattern of variability across age groups for which the initial significant increase in variability occurred at the same age stage for Immediate and Delayed recall within each data set. For the *Lezak (1983)* data set (Figure 7.12, bottom graph), backwards comparisons confirmed a significantly lower standard deviation for the 80-92 year-old age group compared with the 60-69 year-old age group ($p=<0.001^{***}$), and no significant differences between the 80-92 year-old age group and any of the other age groups.

Thus for each of these three data sets for WMS Visual Reproduction there was a significant non-linear inverted-U pattern of variability effects. Although the onset of a *significant* increase in variability occurred at a relatively early age stage for the data collated in Spreen and Strauss (that is by the 30-49 year-old age interval), the significant increase was sustained through to the 50-64 year-old age group for Immediate recall, followed by a declining trend in variability from the late 60's and early 70's age stages. This pattern was not incommensurate therefore with the highly consistent pattern of effects which occurred for the other two data sets (van Gorp et al. and data in Lezak) of a significant increase in variability by the 60's and subsequent decrease in the 70's age groups. Moreover, whilst the effect did not reach significance, there was a clearly apparent trend on the Lezak data set for steadily increasing standard deviations from the 30's age decade prior to the peak in the 60's age decade, which was again broadly commensurate with the effects evident for the Spreen and Strauss data set.

For the *Rey Complex Figure Recall (RCF) Delayed* (see Figure 7.13, based on data in Tables 7.5 and 7.6) in the case of the first two studies included (data collated in Spreen & Strauss, 1991; and van Gorp et al., 1990), there was a highly consistent non-linear pattern of variability with the tendency (as was in evidence for two of the WMS Visual Reproduction data sets) for variability to increase around the late 60's and early 70's and to decrease again in the late 70's and 80's (Figure 7.13 top and middle graph), but which did not reach significance for either data set. However, the *van Gorp et al. (1990)* data set (Figure 7.13, middle graph) revealed an effect that approached significance ($p=0.0690$). Forwards pairwise comparisons of their data indicated that there was a significantly higher standard deviation for the 66-70 year-old age group compared with the 57-65 year-old age group ($p=0.0182^{*}$), and no difference between this youngest group and any of the other age groups. This lent support to the presence of a significant non-linear pattern of variability for this data set, with an initial increase in variability by the late 60's age stage, followed by a subsequent significant decrease in variability in the 70's and 80's age stages. The third RCF recall data set *(Troyer et al. 1994)* (Figure 7.13, bottom graph) revealed a broadly matching trend, in that there was a tendency for variability to be high in the 60's and 70's, and to diminish in later old age, although again this effect did not reach significance. It is notable that the only RCF data set that strongly approached significance on the overall analysis (van Gorp et al.) comprised groups with consistently higher numbers of subjects than the other two studies (28, 45, 57, and 26 subjects in each group), whereas the other two studies contained some groups with only 10, 13, and 14 subjects (data collated in Spreen & Strauss), and 15, 17, and 19 subjects (Troyer et al.). Thus with larger numbers in the samples it appears highly probable that the conceptually consistent trends which were present would also have reached significance.

For the *Benton Visual Retention Test-Revised (BVRT-R)* including males and females with and without a college degrees (see Figure 7.14, based on data in Tables 7.5 and 7.6) the comparisons of standard deviations across all groups revealed a highly significant effect ($p= <0.01$) in all instances. For *males (no degree)* the backwards comparisons revealed a

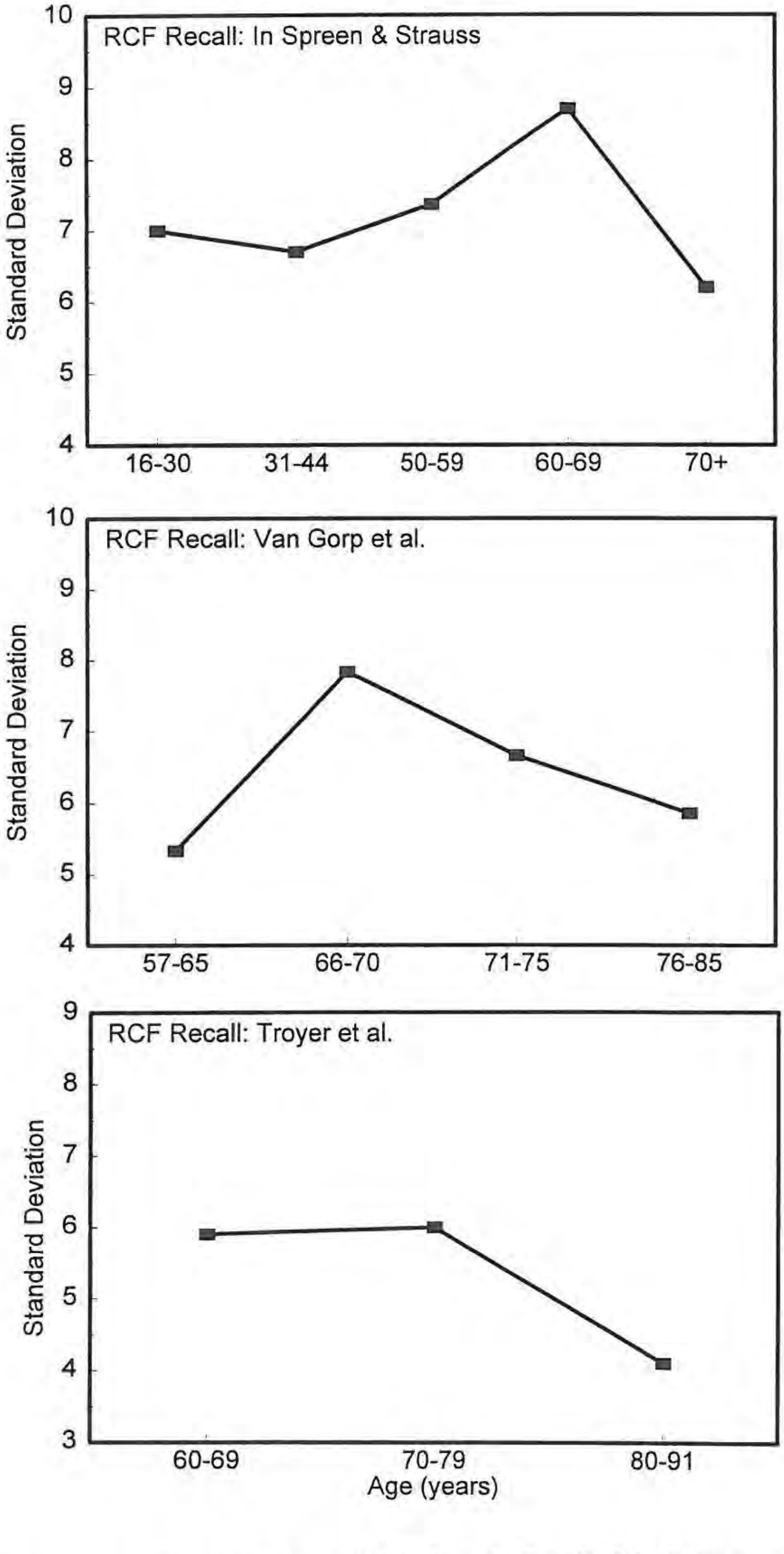


Figure 7.13. Rey Complex Figure (RCF) Recall: Standard deviations across adult age groups for three data sets (collated data in Spreen & Strauss, 1991; van Gorp et al., 1990; Troyer et al., 1994).

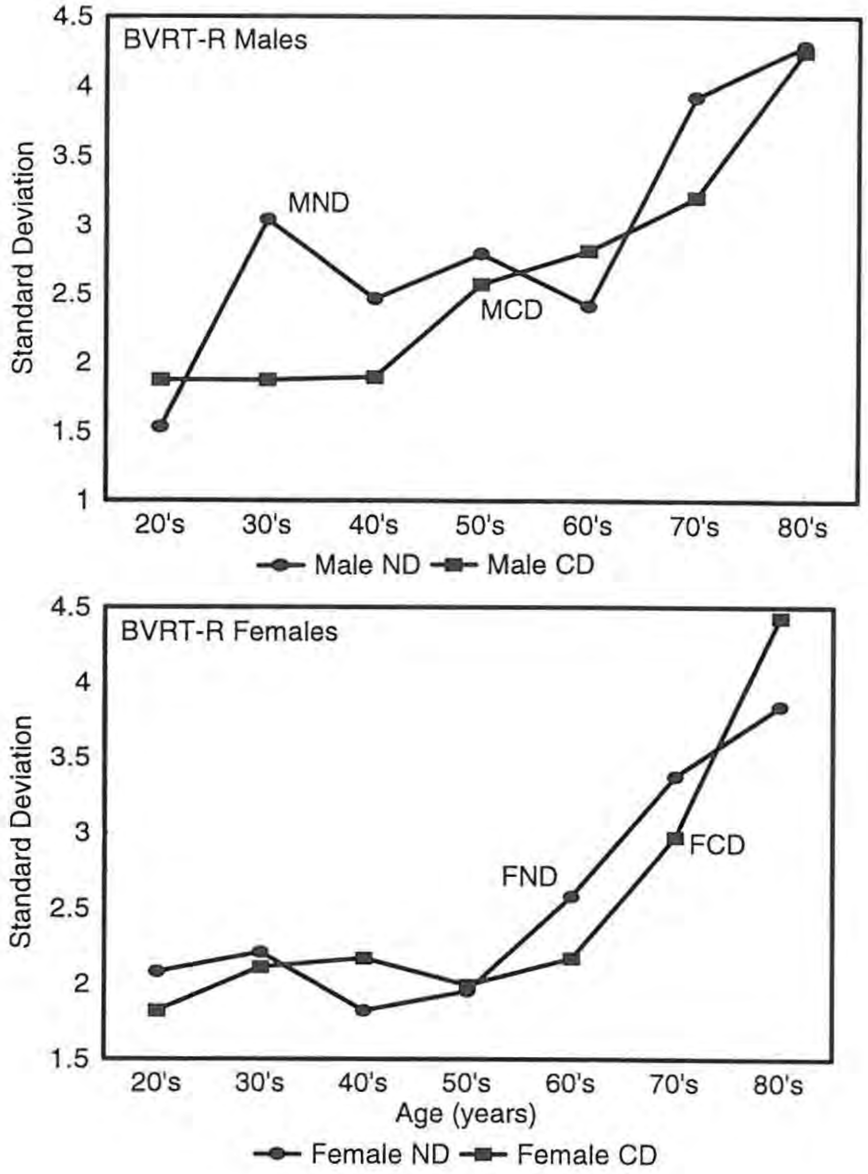


Figure 7.14. Benton Visual Retention Test-Revised (BVRT-R): Standard deviations across adult age groups for Males with no college degree (MND), Males with a college degree (MCD), Females with no degree (FND) and Females with a college degree (FCD) (data from Robertson-Tehabo & Arenberg, 1989, in Spreen & Strauss, 1991). The initial points at which a significant increase in variability occurs in association with age occur at progressively older ages for MND, MCD, FND, and FCD, and are denoted accordingly.

significantly higher standard deviation for the 80's age group compared with all other age groups apart from the 30's age group ($p=0.0049^{**}$, $p=0.0164^{*}$, $p=0.0019^{**}$, $p=<0.001^{***}$), although the difference in the 80's age group compared with the 30's age group also approached significance in the same direction ($p=0.0354$); progressive comparisons of advancing pairs revealed a highly significant increase in variability in the 30's age group compared with the 20's ($p=<0.001^{***}$), *no significant increases in variability between the*

30's and 40's, 40's and 50's, 50's and 60's age decades, a significant increase in variability in the 70's age group compared with the 60's age group (0.0042**), and no significant difference between the 70's and 80's age groups. Considering the indications from the two sets of serial comparisons, and in particular those from the progressive comparisons, there was clear support for an initial peak in standard deviations at the 30 year-old age decade for males (no degree) which remained somewhat inflated until there was a second peak in the 70 and 80 year-old age decades. For *males (college degree)* backwards comparisons revealed a significantly higher standard deviation for the 80's age group compared with all younger age groups ($p = < 0.001^{***}$, in all instances), and the progressive comparisons indicated that as was the case for males (no degree), there was a two-step increase in variability, although in this instance there was an initial increase in standard deviations in the 50's compared with the 40's age group ($p = < 0.001^{***}$) and a second increase in standard deviations in the 80's compared with the 70's age group (0.0090*). Thus overall the variability findings for males on the BVRT-R supported a shuttle shift effect for education in association with aging, in that a significant increase in variability occurs earlier for males with no degree - MND (a left shuttle effect), and later for males with a college degree - MCD (right shuttle effect). These shifting variability effects for males due to differential educational conditions are highlighted in Figure 14, upper graph).

For *females (no degree)* on the BVRT-R, backwards comparisons revealed that there was no significant difference in the standard deviation of the 80's age group compared with those of the 70's and 60's age groups, but that it was significantly higher than those of the 50's, 40's and 20's age groups (0.0043**, 0.0052**, 0.0155*, and 0.0146*, respectively). For *females (college degree)* backwards comparisons revealed that the standard deviation of the 80's age group was not significantly different from that of the 70's age group, but that it was significantly higher than those of the 60's, 50's, 40's and 20's age groups ($p=0.001^{***}$, in all instances). As was the case for males, these results clearly supported a shuttle shift effect for education in association with aging, in that a significant increase in variability occurred earlier for females with no degree - FND (a left shuttle effect), and later for females with a college degree - FCD (right shuttle effect). These shifting variability effects for females due to differential educational conditions are highlighted in Figure 14 (lower graph). They supported a pattern of progressively increasing variability for females across the older adult age groups with an initial increase for females no degree (FND) at the comparatively earlier 60's age decade, when compared with an initial increase for females college degree (FCD) at the later 70's age decade (see Figure 7.14, lower graph). Overall it was apparent from a comparison of both upper and lower graphs (Figure 7.14) that in addition to a shuttle shift effect for educational level, the BVRT-R results supported an additional shuttle shift effect for gender, in that for each instance of educational condition the initial increases in variability occurred earlier for males (that is demonstrated a left shuttle effect), and occurred later for females (that is demonstrated a right shuttle effect).

For the *Rey Visual Design Learning Test (RVDLT)* (see Figure 7.15, based on data in Tables 7.5 and 7.6), for which there are data from only one study (Strauss & Spreen, unpublished study, in Spreen & Strauss, 1991), there was an overall non-linear trend of increasing standard deviations at the 40's age decade and a subsequent decrease in standard deviations in the 50's age group for Trials 1, 5, and Total recall, which on the comparisons of standard deviations across all groups failed to reach significance for Trial 1, was significant for Trial 5 ($p=0.0408$), and only just missed significance for Total recall ($p=0.0550$). For *Trial 5* (see Figure 7.15 top graph) forwards pairwise comparisons revealed a significant increase in the standard deviation of the 40's age groups when compared with the 20's age group, and no other significant differences. For *Total recall* (see Figure 7.15 middle graph) backwards comparisons revealed no significant differences between the 70+ age group and any of the other age groups with the exception of the 40's age group which was significantly higher than the 70+ age group ($p=0.0200^{*}$). Thus generally the findings for the RVDLT Trial 5 and Total recall supported a significant non-linear pattern of an initial increase in variability in the 40's and subsequent decrease in the 50's see Figure 7.15 (upper two graphs). (The apparent increase at age 60 for Trial 5 was *not significant*, which suggests that this was not a robust tendency since it disappeared quite clearly in the cumulative Total score which incorporated the trends across all 5 trials).

Whilst the findings for the RVDLT test supported a non-linear pattern of variability effects as was also present for the WMS Visual Reproduction and Rey Complex Figure Recall, the effect clearly occurred *much earlier* for the RVDLT with a significant peak in variability by the 40's age decade, and a significant decline by the 50's age decade, whereas generally for the other two tasks there was sustained high variability through the 60's, with a subsequent decrease only in the 70's and 80's age stages. For the BSRT-R also there was a sustained increase in variability right up to the 80's age decade with no evidence of a dedifferentiation effect for any of the conditions of the test. It is not possible to rule out the possibility that

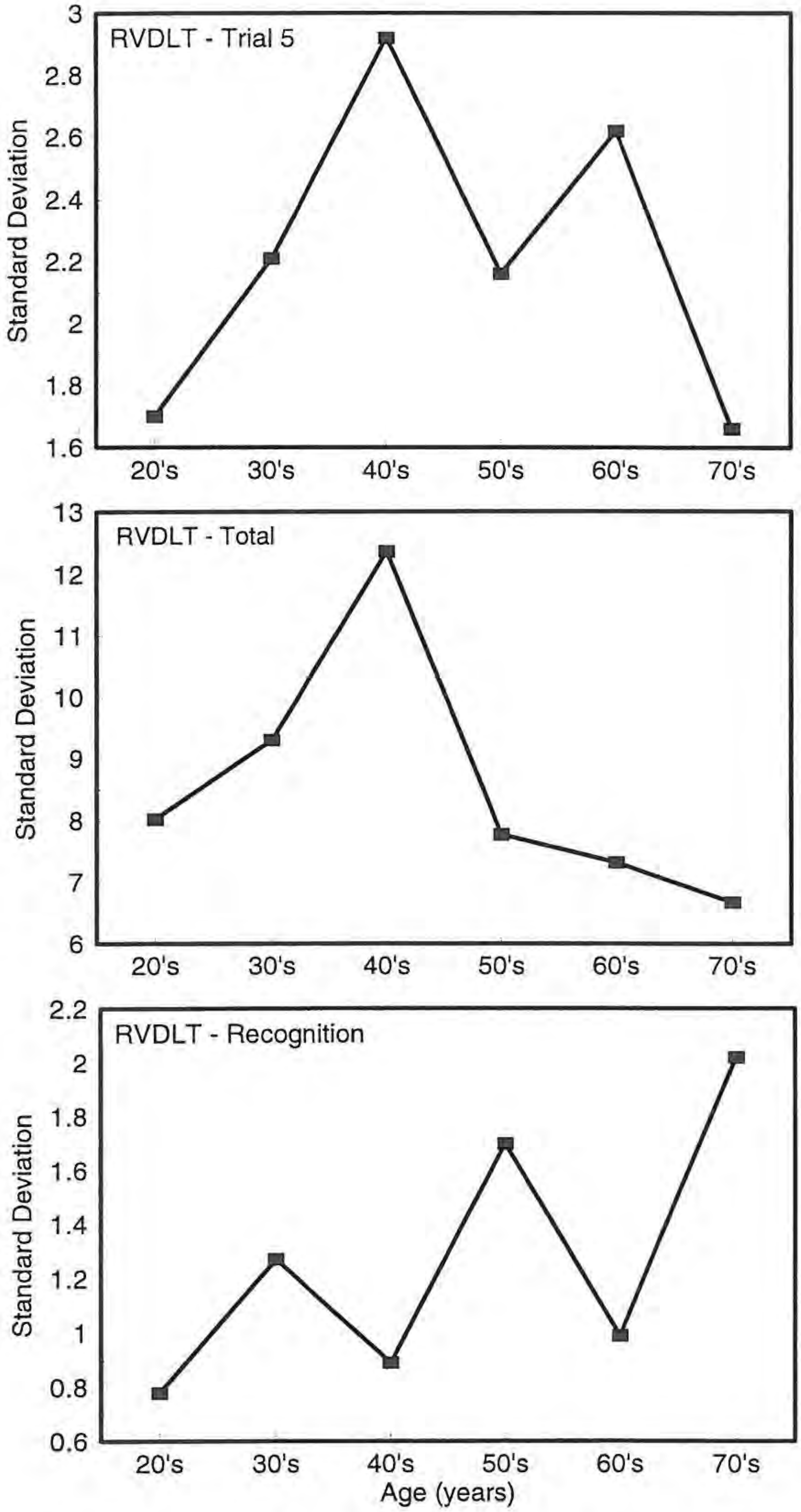


Figure 7.15. Rey Visual Design Learning Test (RVDLT): Standard deviations across adult age groups for Trial 5, Total recall and Recognition (data from Strauss & Spreen, unpublished study, in Spreen & Strauss, 1991).

the earlier onset of the variability bulge for the RVDLT task compared with the other visual memory tasks is a sample effect, particularly since educational levels for the RVDLT data set were not reported. However, it seems much more probable that this was a task-related effect similar to that noted for the BSRT-R, in which it was noted that where consistent verbal learning was tapped (the arguably more age-sensitive task) there was an earlier differentiation effect than where only haphazard verbal recall was tapped (the arguably less age-sensitive task). Similarly the RVDLT is a visual *learning* task compared with the one-off recall tasks of WMS Visual Reproduction, Rey Complex Figure Recall and BSRT-R tasks, and even more importantly it is a *speeded* task in the reproduction phase unlike any of the other visual memory tasks. Thus cumulatively it is clearly a much more age-sensitive task than the others. Accordingly an earlier differentiation occurred (that is a significant left shuttle effect was demonstrated) for the RVDLT relative to all the other arguably relatively less age-sensitive visual memory tasks for which a later differentiation effect occurred (that is a right shuttle effect).

For *RVDLT Recognition* there was a significant effect for variability on the analysis of standard deviations across all groups (<0.01). However the backwards comparisons revealed an erratic zigzag pattern of effects with the 70+ age group being significantly higher than every alternate age group being the 60's, 40's and 20's age groups ($p=0.0198^{*}$, $p=0.0031^{**}$, and 0.001^{***}, respectively), and no different from each of the 30's and 50's age groups (see Figure 7.15, bottom graph).

7.3.2 Comparisons of means

For the modality of visual memory (see Table 7.5 for a summary of mean scores across all groups, and Table 7.6 for pairwise statistical comparisons between the oldest and youngest age groups) in all instances with the single exception of the RVDLT Recognition data set there was absolutely clearcut evidence of steadily declining average performance as reflected in uniformly declining or increasing mean scores (where low scores denote poor performance and high scores denote poor performance, respectively). This was the case whether the variability pattern was non-linear which occurred in most instances (WMS Visual Reproduction, Rey Complex Recall, and the RVDLT trials 1,5, and Total recall), or whether there was a pattern of progressively increasing variability which occurred for the BVRT-R. For each of the data sets with the exception of RVDLT Recognition, the pairwise comparisons of the mean of the youngest age group with the mean of the oldest age group was highly significant ($p = < 0.001^{***}$, in all instances). This effect of a highly significant steady decline in average performance for visual memory tasks is illustrated in Figure 16 (upper two graphs), as it occurs in association with a significant non-linear inverted-U pattern in the standard deviations across the adult age range, for WMS Visual Reproduction where the variability bulge was shifted relatively to the right (Figure 7.16, upper graph), and for the arguably more age-sensitive speeded RVDLT Total recall task where the variability curve was shifted relatively more to the left (Figure 7.16, middle graph).

For RVDLT Recognition the mean of the youngest age group was also significantly lower than that of the oldest age group (($p=0.0099^{**}$). However, as is evident from Figure 7.16 (bottom graph), there was some fluctuation in the course of the overall declining mean when compared with the other data sets for visual memory. In addition the RVDLT Recognition mean covered a markedly small range of only 1.5 points compared with the other visual memory tasks (a range from 3 to 12 points for WMS visual Reproduction, and from 30 to 55 points for RVDLT-Total, as illustrated in Figure 16). This erratic mean tendency for RVDLT Recognition was accompanied by a markedly fluctuating pattern of standard deviations across age groups.

7.3.3 Visual memory: Summary of effects

Generally for visual memory, as tapped by the WMS Visual Reproduction, Rey Complex Figure Recall and BVRT-R tests, there was support for a significant increase in variability with advancing age by the 60's age decade (that is the initial differentiation stage of the shuttle bulge effect). Further there was evidence of a subsequent significant decrease in variability by the late 70's and 80's age stages for the WMS Visual Reproduction and Rey Complex Figure Recall tasks, (that is the full shuttle bulge effect of initial differentiation followed by subsequent dedifferentiation), although this was not present for the BSRT-R for which there was a late sustained differentiation effect. For the RVDLT, a visual learning task with a speeded reproduction component, a significant variability bulge in association with aging occurred much earlier than for the other unspeeded visual reproduction tasks, peaking in the 40's age decade, and declining by the 50's age decade.

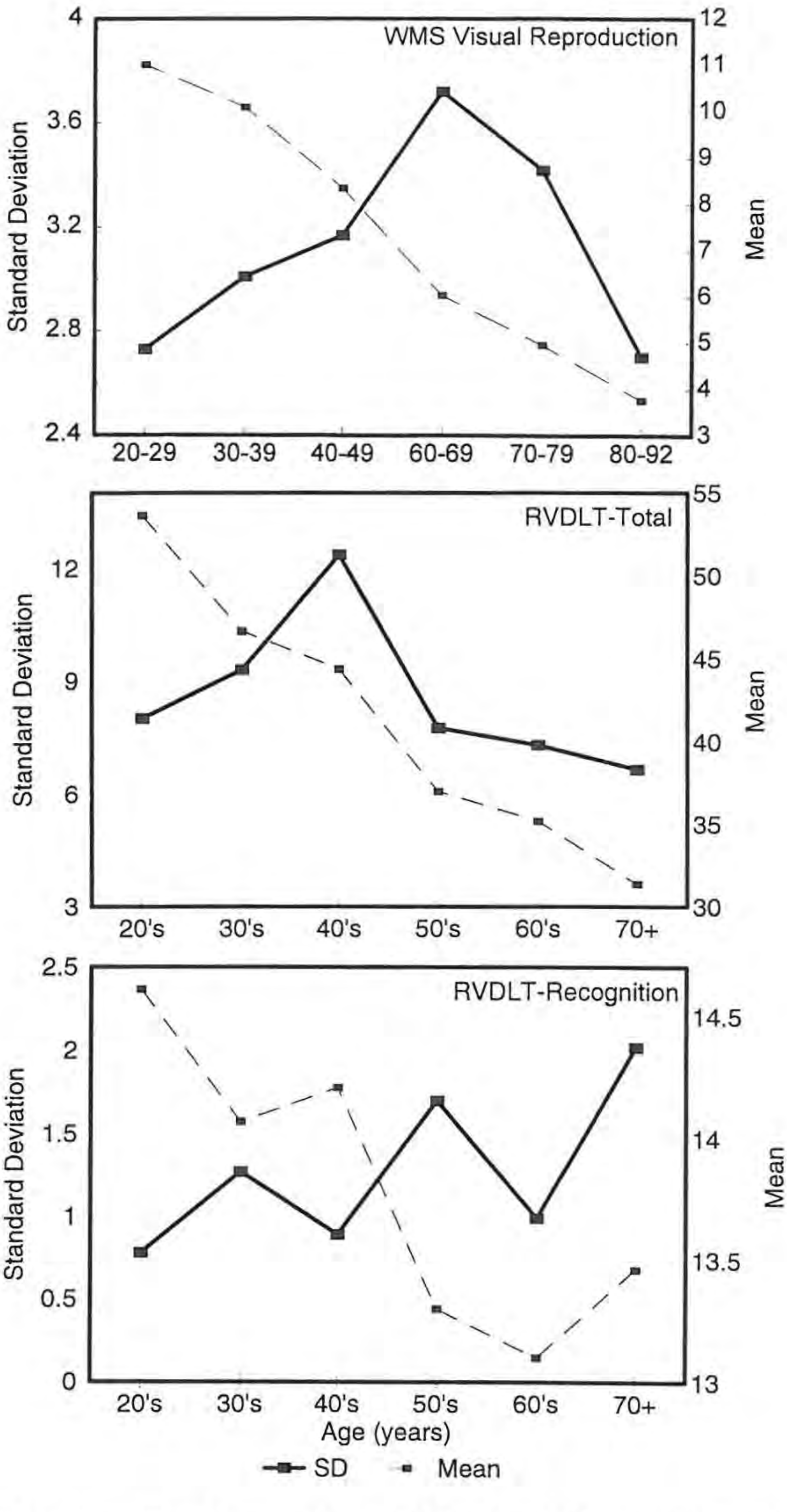


Figure 7.16. Wechsler Memory Scale (WMS) Visual Reproduction and Rey Visual Design Learning Test (RVDLT): Standard deviations and mean scores across adult age groups for the Wechsler Memory Scale (WMS) Visual Reproduction (data in Lezak, 1983), Rey Visual Design Learning Test (RVDLT) Total recall and Recognition (Strauss & Spreen, unpublished study in Spreen & Strauss, 1991), illustrating a non-linear inverted-U pattern of variability in the later adult years in association with declining mean scores (upper two graphs), and an erratic pattern of standard deviations and means for RVDLT Recognition (bottom graph).

Over a number of visual memory data sets three clear shuttle shift effects were detected in the variability data, for educational level, gender and age-sensitivity of tasks. The data for the BVRT-R supported a shuttle shift effect for gender, in that for males variability effects for both educational conditions occurred earlier than for females (left versus right shuttle effect, respectively), and also for educational level in that for both gender conditions variability effects occurred earlier for the lower educational condition of 'no degree' than for the higher educational condition of a 'college degree' (left versus right shuttle effect, respectively). A shuttle shift effect for differential age-sensitivity of tasks was also supported in that a non-linear pattern of an initial increase and subsequent decrease in variability in association with aging was apparent earlier for the relatively more age-sensitive task of the RVDLT (a speeded and learning visual reproduction task) compared with all the other non-speeded and non-learning visual reproduction tasks, once again demonstrating a left versus right shuttle effect, respectively.

All of the above variability effects occurred in association with a significant decline in average performance in association with older age. In the one instance where there were erratic trends in the pattern of variability effects (RVDLT recognition), this occurred in the presence of the only test for which there was also an erratic up-down pattern in the mean scores across a very narrow range of scores, indicating that the recognition task is not particularly age sensitive.

Table 7.5. MEMORY (VISUAL): Mean Scores and Comparisons of Standard Deviations across all Age Groups.

WMS Visual Reproduction (Collated data in Spreen & Strauss 1991).

AGE GROUP	18-29	30-49	50-64	65-69	70-74	75-79	80+	df	B[1]	p
n	97	81	51	49	74	40	13	-	-	-
WMS Vis Rep-Immed										
Mean	10.48	10.10	8.73	6.0	5.1	4.9	3.3	-	-	-
SD	1.93	2.55	2.59	2.1	2.0	2.0	2.3	6	11.89	<0.01**
WMS Vis Rep-Del										
Mean	9.84	9.26	7.35	5.4	4.3	4.2	2.8	-	-	-
SD	2.21	2.74	2.46	2.5	2.3	1.9	1.9	6	9.39	<0.01**

WMS Visual Reproduction (van Gorp et al. 1990).

AGE GROUP	57-65	66-70	71-75	76-85	df	B	p
n	28	45	57	26	-	-	-
WMS Vis Rep-Immed							
Mean	9.60	8.67	7.79	4.09	-	-	-
SD	2.76	3.81	2.99	2.26	3	9.10	<0.01**
WMS Vis Rep-Del							
Mean	6.60	6.33	4.38	1.64	-	-	-
SD	2.46	3.83	3.93	1.69	3	24.38	<0.01**

WMS Visual Reproduction (Collated data in Lezak 1983)

AGE GROUP	20-29	30-39	40-49	60-69	70-79	80-92	df	B	p
n	50	53	46	70	46	115	-	-	-
WMS Vis Rep-Immed									
Mean	11.00	10.09	8.35	6.03	4.95	3.76	-	-	-
SD	2.73	3.01	3.17	3.72	3.42	2.70	5	11.61	<0.01**

[1]Bartlett's test Statistic

Table 7.5 *continued*

Table 7.5 MEMORY (VISUAL): Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

Rey Complex Figure Recall Delayed (Collated data in Spreen & Strauss 1991)								
AGE GROUP	16-30	31-44	50-59	60-69	70+	df	B	p
n	67	26	14	13	10	-	-	-
Mean	22.70	19.50	18.82	16.65	11.80	-	-	-
SD	7.00	6.70	7.37	8.70	6.20	4	1.62	0.5010

Rey Complex Figure Recall Delayed (van Gorp et al. 1990)							
AGE GROUP	57-65	66-70	71-75	76-85	df	B	p
n	28	45	57	26	-	-	-
Mean	14.45	14.13	11.13	8.41	-	-	-
SD	5.33	7.83	6.66	5.86	3	5.58	0.0690^{+}

Rey Complex Figure Recall Delayed (Troyer et al. 1994)						
AGE GROUP	60-69	70-79	80-91	df	B	p
n	19	17	15	-	-	-
Mean	18.90	15.70	13.90	-	-	-
SD	5.90	6.00	4.10	2	2.38	0.3044

Table 7.5 *continued*

Table 7.5 MEMORY (VISUAL): Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

BVRT-R Males (Robertson-Tehabo & Arenberg 1989, in Spreen & Strauss 1991)										
AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	B	p
BVRT-R No Degree										
n	45	57	51	42	28	47	15	-	-	-
Mean	2.02	3.56	3.75	5.07	6.25	7.72	11.13	-	-	-
SD	1.53	3.04	2.46	2.79	2.41	3.92	4.29	6	44.43	<0.01**
BVRT-R College Degree										
n	92	181	155	156	123	139	40	-	-	-
Mean	2.47	2.49	2.94	3.67	4.66	6.03	8.15	-	-	-
SD	1.87	1.87	1.89	2.57	2.81	3.20	4.25	6	113.69	<0.01**

BVRT-R Females (Robertson-Tehabo & Arenberg 1989, in Spreen & Strauss 1991)										
AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	B	p
BVRT-R No degree										
n	19	27	23	35	42	32	9	-	-	-
Mean	2.90	3.04	3.87	5.11	6.02	7.53	9.44	-	-	-
SD	2.08	2.21	1.82	1.95	2.58	3.37	3.84	6	19.51	<0.01**
BVRT-R College Degree										
n	37	70	28	38	56	43	13	-	-	-
Mean	2.70	2.67	2.61	3.68	4.57	6.79	7.62	-	-	-
SD	1.82	2.11	2.17	1.99	2.17	2.97	4.43	6	28.74	<0.01**

Table 7.5 *continued*

Table 7.5 MEMORY (VISUAL): Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

RVDLT (Strauss & Spreen unpublished study, in Spreen & Strauss 1991)									
AGE GROUP	20's	30's	40's	50's	60's	70+	df	B	p
n	23	14	14	10	10	13	-	-	-
RVDLT-Trial 1									
Mean	6.09	5.36	5.64	5.00	4.70	4.29	-	-	-
SD	2.11	2.10	1.91	1.56	1.16	1.64	5	4.88	0.1129
RVDLT-Trial 5									
Mean	13.48	11.64	10.93	9.30	8.20	7.46	-	-	-
SD	1.70	2.21	2.92	2.16	2.62	1.66	5	7.01	0.0408*
RVDLT-Total									
Mean	53.65	46.64	44.36	37.00	35.20	31.38	-	-	-
SD	8.02	9.30	12.36	7.77	7.32	6.67	5	6.38	0.0550⁺
RVDLT-Recog									
Mean	14.61	14.07	14.21	13.30	13.10	13.46	-	-	-
SD	0.78	1.27	0.89	1.70	0.99	2.02	5	19.76	<0.01**

Significant difference (*$p<0.05$; **$p<0.01$); Approaching significance ($^{+}p<0.10$).

Table 7.6. MEMORY (VISUAL): Pairwise comparisons of Means between the oldest and youngest Age Groups, and of Standard Deviations between the oldest Age Group and each younger Age Group unless otherwise specified.

WMS Visual Reproduction (Collated data in Spreen & Strauss 1991)										
AGE GROUP	18-29	30-49	50-64	65-59	70-74	75-79	80+	df	t/F	p
n	97	81	51	49	74	40	13	-	-	-
WMS Vis Rep-Immed										
Mean	10.48	-	-	-	-	-	3.3	108	12.32	<0.001***
SD	-	-	-	-	-	2.0	2.3	39,12	0.76	0.2454
SD	-	-	-	-	2.0	-	2.3	73,12	0.76	0.2247
SD	-	-	-	2.1	-	-	2.3	48,12	0.83	0.3108
SD	-	-	2.59	-	-	-	2.3	50,12	1.27	0.3405
SD	-	2.55	-	-	-	-	2.3	80,12	1.23	0.3642
SD	1.93	-	-	-	-	-	2.3	96,12	0.70	0.1700
WMS Vis Rep-Del										
Mean	9.84	-	-	-	-	-	2.8	108	10.95	<0.001***
SD	-	-	-	-	-	1.9	1.9	39,12	1.00	0.5331
SD	-	-	-	-	2.3	-	1.9	73,12	1.47	0.2371
SD	-	-	-	2.5	-	-	1.9	48,12	1.73	0.1508
SD	-	-	2.46	-	-	-	1.9	50,12	1.68	0.1655
SD	-	2.74	-	-	-	-	1.9	80,12	2.08	0.0759
SD	2.21	-	-	-	-	-	1.9	96,12	1.35	0.2900

Table 7.6 *continued.*

Table 7.6. MEMORY (VISUAL): Pairwise comparisons of Means and of Standard Deviations between Age Groups *continued.*

WMS Visual Reproduction (Collated data in Spreen & Strauss 1991)										
AGE GROUP	18-29	30-49	50-64	65-59	70-74	75-79	80+	df	t/F	p
Forwards Comparisons										
WMS Vis Rep-Immed										
SD	1.93	2.55	-	-	-	-	-	96,80	0.57	0.0046**
SD	1.93	-	2.59	-	-	-	-	96,50	0.56	0.0069**
SD	1.93	-	-	2.1	-	-	-	96,48	0.84	0.2398
SD	1.93	-	-	-	2.0	-	-	96,73	0.93	0.3692
SD	1.93	-	-	-	-	2.0	-	96,39	0.93	0.3807
SD	1.93	-	-	-	-	-	2.3	96,12	0.70	0.1700
WMS Vis Rep-Del										
SD	2.21	2.74	-	-	-	-	-	96,80	0.65	0.0094*
SD	2.21	-	2.46	-	-	-	-	96,50	0.81	0.1838
SD	2.21	-	-	2.5	-	-	-	96,48	0.78	0.1532
SD	2.21	-	-	-	2.3	-	-	96,73	0.92	0.3546
SD	2.21	-	-	-	-	1.9	-	96,39	1.35	0.1450
SD	2.21	-	-	-	-	-	1.9	96,12	1.35	0.2900

Table 7.6 *continued.*

Table 7.6. MEMORY VISUAL: Pairwise comparisons of Means and of Standard Deviations between Age Groups *continued.*

WMS Visual Reproduction (Van Gorp et al. 1990)

AGE GROUPS	57-65	66-70	71-75	76-85	df	t/F	p
n	28	45	57	26	-	-	-
WMS Vis Rep-Immed							
Mean	9.60	-	-	4.09	52	7.99	<0.001***
SD	-	-	2.99	2.26	56,25	1.75	0.0633
SD	-	3.81	-	2.26	44,25	2.84	0.0033**
SD	2.76	-	-	2.26	27,25	1.49	0.1592
WMS Vis Rep-Del							
Mean	6.60	-	-	1.64	52	8.57	<0.001***
SD	-	-	3.93	1.69	56,25	5.41	<0.001***
SD	-	3.83	-	1.69	44,25	5.14	<0.001***
SD	2.46	-	-	1.69	27,25	2.12	0.0315
Additional Comparison							
SD	2.46	3.83	-	-	27,44	0.41	0.0083**

WMS Visual Reproduction (Collated data in Lezak 1983)

AGE GROUP	20-29	30-39	40-49	60-69	70-79	80-92	df	t/F	p
n	50	53	46	70	46	115	-	-	-
Mean	11.00	-	-	-	-	3.76	163	15.78	<0.001***
SD	-	-	-	-	3.42	2.70	45,114	1.60	0.0235
SD	-	-	-	3.72	-	2.70	69,114	1.90	<0.001***
SD	-	-	3.17	-	-	2.70	45,114	1.38	0.0887
SD	-	3.01	-	-	-	2.70	52,114	1.24	0.1693
SD	2.73	-	-	-	-	2.70	49,114	1.02	0.4509

Table 7.6 *continued.*

Table 7.6. MEMORY VISUAL: Pairwise comparisons of Means and of Standard Deviations between Age Groups *continued.*

Rey Complex Figure Recall Delayed (Van Gorp et al. 1990)

AGE GROUPS	57-65	66-70	71-75	76-85	df	t/F	p
n	28	45	57	26	-	-	-
Mean	14.45	-	-	8.41	52	3.97	<0.001***
SD	-	-	6.66	5.86	56,25	1.29	0.2448
SD	-	7.83	-	5.86	44,25	1.79	0.0617
SD	5.33	-	-	5.86	27,25	0.83	0.3143
Forwards Comparisons							
SD	5.33	7.83	-	-	27,44	0.46	0.0182*
SD	5.33	-	6.66	-	27,56	0.64	0.1040
SD	5.33	-	-	5.86	27,25	0.83	0.3143

BVRT-R: Male/No degree (Robertson-Tehabo & Arenberg 1989, in Spreen & Strauss 1991)

AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	t/F	p
n	45	57	51	42	28	47	15	-	-	-
Mean	2.02	-	-	-	-	-	11.13	58	-12.25	<0.001***
SD	-	-	-	-	-	3.92	4.29	46,14	0.84	0.3094
SD	-	-	-	-	2.41	-	4.29	27,14	0.32	0.0049**
SD	-	-	-	2.79	-	-	4.29	41,14	0.42	0.0164*
SD	-	-	2.46	-	-	-	4.29	50,14	0.33	0.0019**
SD	-	3.04	-	-	-	-	4.29	56,14	0.50	0.0354
SD	1.53	-	-	-	-	-	4.29	44,14	0.13	<0.001***
Progressive Comparisons										
SD	1.53	3.04	-	-	-	-	-	44,56	0.25	<0.001***
SD	-	3.04	2.46	-	-	-	-	56,50	1.53	0.0649
SD	-	-	2.46	2.79	-	-	-	50,41	0.78	0.1969
SD	-	-	-	2.79	2.41	-	-	41,27	1.34	0.2132
SD	-	-	-	-	2.41	3.92	-	27,46	0.38	0.0042**
SD	-	-	-	-	-	3.92	4.29	46,14	0.83	0.3094

Table 7.6 *continued*

Table 7.6. MEMORY VISUAL: Pairwise comparisons of Means and of Standard Deviations between Age Groups *continued.*

BVRT-R: Male/College Degree (Robertson-Tehabo & Arenberg, in Spreen & Strauss 1991)

AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	t/F	p
n	92	181	155	156	123	139	40	-	-	-
Mean	2.47	-	-	-	-	-	8.15	130	-10.69	<0.001***
SD	-	-	-	-	-	3.20	4.25	138,39	0.57	0.0090*
SD	-	-	-	-	2.81	-	4.25	122,39	0.44	<0.001***
SD	-	-	-	2.57	-	-	4.25	155,39	0.37	<0.001***
SD	-	-	1.89	-	-	-	4.25	154,39	0.20	<0.001***
SD	-	1.87	-	-	-	-	4.25	180,39	0.19	<0.001***
SD	1.87	-	-	-	-	-	4.25	91,39	0.19	<0.001***
Progressive Comparisons										
SD	1.87	1.87	-	-	-	-	-	91,180	1.00	0.4920
SD	-	1.87	1.89	-	-	-	-	180,154	0.98	0.4438
SD	-	-	1.89	2.57	-	-	-	154,155	0.54	<0.001***
SD	-	-	-	2.57	2.81	-	-	155,122	0.84	0.1466
SD	-	-	-	-	2.81	3.20	-	122,138	0.77	0.0712
SD	-	-	-	-	-	3.20	4.25	138,39	0.57	0.0090*

Table 7.6 *continued*

Table 7.6. MEMORY VISUAL: Pairwise comparisons of Means and of Standard Deviations between Age Groups *continued.*

BVRT-R: Female/No Degree (Robertson-Tehabo & Arenberg, in Spreen & Strauss 1991)

AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	t/F	p
n	19	27	23	35	42	32	9	-	-	-
Mean	2.90	-	-	-	-	-	9.44	26	-5.89	<0.001***
SD	-	-	-	-	-	3.37	3.84	31,8	0.77	0.2806
SD	-	-	-	-	2.58	-	3.84	41,8	0.45	0.4603
SD	-	-	-	1.95	-	-	3.84	34,8	0.26	0.0043**
SD	-	-	1.82	-	-	-	3.84	22,8	0.22	0.0052**
SD	-	2.21	-	-	-	-	3.84	26,8	0.33	0.0155*
SD	2.08	-	-	-	-	-	3.84	18,8	0.29	0.0146*
Progressive Comparisons										
SD	2.08	2.21	-	-	-	-	-	18,26	0.89	0.4017
SD	-	2.21	1.82	-	-	-	-	26,22	1.47	0.1791
SD	-	-	1.82	1.95	-	-	-	22,34	0.87	0.3731
SD	-	-	-	1.95	2.58	-	-	34,41	0.57	0.0484
SD	-	-	-	-	2.58	3.37	-	41,31	0.59	0.0546
SD	-	-	-	-	-	3.37	3.84	31,8	0.77	0.2806

Table 7.6 *continued.*

Table 7.6. MEMORY VISUAL: Pairwise comparisons of Means and of Standard Deviations between Age Groups *continued.*

BVRT-R Female/College Degree (Robertson-Tehabo & Arenberg in Spreen & Strauss 1991)

AGE GROUP	20's	30's	40's	50's	60's	70's	80's	df	t/F	p
n	37	70	28	38	56	43	13	-	-	-
Mean	2.70	-	-	-	-	-	7.62	48	-5.61	<0.001***
SD	-	-	-	-	-	2.97	4.43	42,12	0.45	0.0280
SD	-	-	-	-	2.17	-	4.43	55.12	0.24	<0.001***
SD	-	-	-	1.99	-	-	4.43	37.12	0.20	<0.001***
SD	-	-	2.17	-	-	-	4.43	27.12	0.24	<0.001***
SD	-	2.11	-	-	-	-	4.43	69.12	0.23	<0.001***
SD	1.82	-	-	-	-	-	4.43	36.12	0.17	<0.001***
Progressive Comparisons										
SD	1.82	2.11	-	-	-	-	-	36,39	0.74	0.1674
SD	-	2.11	2.17	-	-	-	-	69,27	0.95	0.4125
SD	-	-	2.17	1.99	-	-	-	27,37	0.19	0.3147
SD	-	-	-	1.99	2.17	-	-	37,55	0.84	0.2914
SD	-	-	-	-	2.17	2.97	-	55,42	0.53	0.0146*
SD	-	-	-	-	-	2.97	4.43	42,12	0.45	0.0280

Table 7.6 *continued*

Table 7.6 **MEMORY (VISUAL):** Pairwise comparisons of Means and of Standard Deviations between Age Groups *continued.*

RVDLT (Strauss & Spreen unpublished study, in Spreen & Strauss 1991) AGE GROUP	20's	30's	40's	50's	60's	70+	df	t/F	p
n	23	14	14	10	10	13	-	-	-
RVDLT-Trial 5									
Mean	13.48	-	-	-	-	7.46	34	10.29	<0.001***
SD	-	-	-	-	2.62	1.66	9,12	2.49	0.0713
SD	-	-	-	2.16	-	1.66	9,12	1.69	0.1487
SD	-	-	2.92	-	-	1.66	13,12	3.09	0.0296
SD	-	2.21	-	-	-	1.66	13,12	1.77	0.1632
SD	1.70	-	-	-	-	1.66	22,12	1.05	0.4832
RVDLT-Total									
Mean	53.65	-	-	-	-	31.38	34	8.48	<0.001***
SD	-	-	-	-	7.32	6.67	9,12	1.20	0.3736
SD	-	-	-	7.77	-	6.67	9,12	1.36	0.3048
SD	-	-	12.36	-	-	6.67	13,12	3.43	0.0200*
SD	-	9.30	-	-	-	6.67	13,12	1.94	0.1296
SD	8.02	-	-	-	-	6.67	22,12	1.45	0.2576
RVDLT-Recog									
Mean	14.61	-	-	-	-	13.46	34	2.48	0.0099**
SD	-	-	-	-	0.99	2.02	9,12	0.74	0.0198*
SD	-	-	-	1.70	-	2.02	9,12	0.71	0.3070
SD	-	-	0.89	-	-	2.02	13,12	0.19	0.0031**
SD	-	1.27	-	-	-	2.02	13,12	0.40	0.0550
SD	0.78	-	-	-	-	2.02	22,12	0.15	0.001***

Table 7.6 *continued*

Table 7.6 MEMORY (VISUAL): Pairwise comparisons of Means and of Standard Deviations between Age Groups *continued.*

RVDLT (Strauss & Spreen unpublished study, in Spreen & Strauss 1991)									
AGE GROUP	20's	30's	40's	50's	60's	70+	df	t/F	p
Forwards Comparisons									
RVDLT-Trial 5									
SD	1.70	2.21	-	-	-	-	22,13	0.59	0.1344
SD	1.70	-	2.92	-	-	-	22,13	0.34	0.0123*
SD	1.70	-	-	2.16	-	-	22,9	0.62	0.1722
SD	1.70	-	-	-	2.62	-	22,9	0.42	0.0473
SD	1.70	-	-	-	-	1.66	22,12	1.05	0.4832
RVDLT-Total									
SD	8.02	9.30	-	-	-	-	22,13	0.74	0.2614
SD	8.02	-	12.36	-	-	-	22,13	0.42	0.0355
SD	8.02	-	-	7.77	-	-	22,9	1.07	0.4873
SD	8.02	-	-	-	7.32	-	22,9	1.20	0.4061
SD	8.02	-	-	-	-	6.67	22,12	1.45	0.2576
RVDLT-Recog									
SD	0.78	1.27	-	-	-	-	22,13	0.38	0.0212
SD	0.78	-	0.89	-	-	-	22,13	0.77	0.2832
SD	0.78	-	-	1.70	-	-	22,9	0.21	0.0014*
SD	0.78	-	-	-	0.99	-	22,9	0.62	0.1733
SD	0.78	-	-	-	-	2.02	22,12	0.15	<0.001***

Significant difference (*p<0.10; **p<0.05; ***p<0.01) as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

7.4 LANGUAGE SKILLS

For the modality of language skills the data sets submitted to analysis are collated in Table A (Appendix pp.10-11), and a summary of the educational and/or IQ levels of the age groups appear in Table B (Appendix p.19). The following tests and accompanying sources of normative data sets were included in the analysis (ordered for discussion as they appear in the tables):

(i) **FAS Word Association Test (Oral):** Two data sets (Read, 1987, in Spreen & Strauss, 1991; Axelrod & Henry, 1992).

(ii) **Animal Name Fluency:** One data set (Read, 1987, in Spreen & Strauss, 1991).

(iii) **Boston Word Naming:** Two data sets (Van Gorp, Satz, Kiersch, & Henry, 1986, in Spreen & Strauss, 1991; Au, Joung, Nicholas, Obler, Kass, & Albert, 1995).

7.4.1 Comparisons of standard deviations across age groups

For the *FAS Word Association Oral* (see Figure 7.17 based on data in Tables 7.7 and 7.8), the first data set for the two educational conditions of <12 and >13 years of education (Read, 1987, in Spreen & Strauss, 1991), revealed a significant effect for the comparisons of standard deviations across all groups for both educational conditions ($p=0.0371$ and $p=0.0034$, respectively). For the <12 years educational condition neither the backwards nor the forwards comparisons revealed the source of the significance detected on the overall analysis (there were no significant differences between any of the pairwise comparisons). This appears to be due to the unusual occurrence of a relatively inflated standard deviation in the youngest 50-54 year-old age group in the <12 years educational condition. An additional spot comparison for this subset of data between the 55-59 year old and 70-74 year-old age groups revealed a significantly higher standard deviation for the 70-74 year old group when compared with the 55-59 year old group ($p=0.0061^{*}$). For the >13 years educational condition backwards comparisons revealed that there was a significantly higher standard deviation for the oldest 75+ age group compared with the 70-74, 65-69 and 60-64 year-old age groups ($p=0.0104^{*}$, 0.0097^{**} and 0.0005^{***}, respectively), and no significant difference

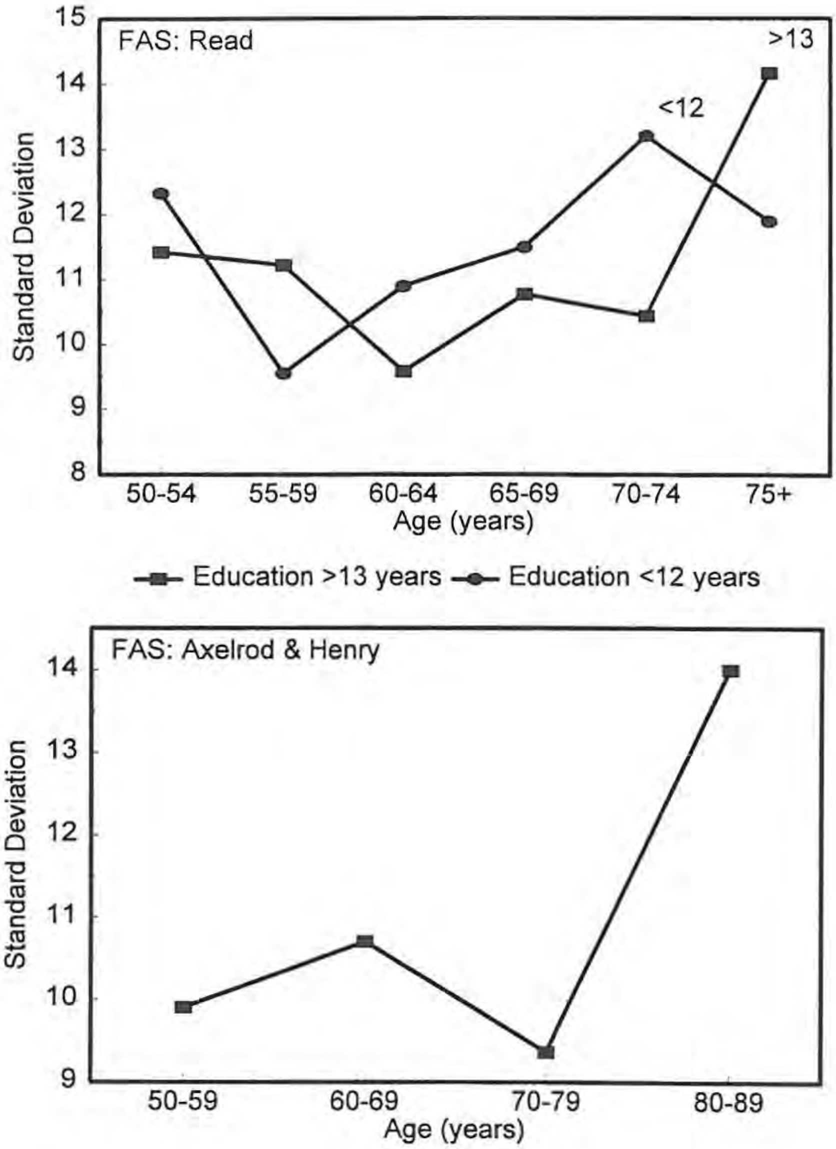


Figure 7.17. FAS Word Association (Oral): Standard deviations across age groups for two data sets (Read, 1987, in Spreen & Strauss, 1991; Axelrod & Henry, 1992). The points at which variability peaks progressively for the low and high educational conditions in the Read data set are denoted by < 12 and > 13, respectively (upper graph).

between the 75+ age group compared with the youngest 55-59 and 50-54 year-old age groups. Together these results suggest some fluctuations in standard deviations in the 50's age groups which were inconsistent between the two educational conditions. However, from the 60's age groups upwards there was a significant pattern of progressively increasing standard deviations for both educational conditions. A left versus right shuttle effect was apparent in the later age stages (see demarcations indicating educational conditions of the data sets on Figure 7.17, upper graph) in that the significant increase in variability in these later years occurred a decade earlier for the < 12 year old educational condition (for which there was a left shuttle effect) than for the > 13 year old educational condition (for which there was a right shuttle effect. Moreover for the < 12 year educational condition which was shifted relatively to the left, there was the possible indication of an emergent non-linear pattern with a subsequent decline in standard deviations following the initial significant increase, although the tendency was not significant.

For the second data set (Axelrod & Henry, 1992) (see Figure 7.17, lower graph), the comparison of standard deviations across all groups did not reach significance (0.1632). However the study was based on much smaller samples of only 20 subjects per group (compared with groups ranging in age from 42 to 70 subjects per group in the Read study), which could explain why the effect in the Axelrod and Henry study did not reach significance. Moreover the trends were highly consistent with a sharp increase in variability in the very late post mid-70's years as was evident in the Read > 13 years educational condition. Conceptually this is consistent with the fact that the Axelrod and Henry study was based on four groups well-matched for a relatively high level of education (15, 14, 15, and 15 years of education per group) (see Appendix B, p.19), and thus on the educational dimension was comparable with the Read high education (> 13 years) data set. Accordingly the Read > 13 year data set and the Axelrod and Henry data set both showed the same pattern of a late old-age increase in variability (around the late 70's and 80's age stages). In neither instance was there the indication of a tendency for an initial increase in the early 70's age stage together with a subsequent tendency for a decrease in standard deviations in later old age as was present with the Read < 12 year educational condition.

The trends on the Axelrod and Henry data set do not provide support for an inflation in the earlier 50's age group such was noted in the Read study for the > 13 year educational condition. This points towards the likelihood of a sample artifact rather than a robust variability effect for this age group in the Read study. One possibility is that unlike for the Axelrod and Henry study (for which, as noted earlier, the four groups were well matched for educational level), the Read age groups were not well-matched for education *within* the broader < 12 and > 13 years educational conditions. Thus disproportionately distributed groups for education may have caused the erratic pattern in the earlier age groups in the Read study.

For *Animal Name Fluency* (Read, 1987, in Spreen & Strauss, 1991) (see Figure 7.18, based on data in Tables 7.7 and 7.8), the trends for both educational conditions presented an erratic and inconsistent pattern across the age groups within a very narrow range of measurement units from 3.58 to 5.14. The comparisons of standard deviations across all age groups revealed no significant effect for the < 12 years educational condition, and a significant effect for the > 13 years educational condition (0.0032). Backwards comparisons for the > 13 years educational condition revealed a significantly higher standard deviation between the 75+ age group and the 70-74 year-old age group (0.0078**), and no other significant differences in the standard deviations between the groups which supports a significant drop in variability just for the 70-74 year-old age stage. It appears likely that the statistical potency of the isolated dip in the 70-74 year-old age group for the > 13 years educational condition was an artifact of an otherwise exceptionally narrow range of standard deviations, and if this dip was removed neither subset of data for Animal Name Fluency would have revealed any significant effect for variability across age groups. Thus in particular it is probable that an overall absence of variability effects might reflect the more true picture for this test which would emerge if the standard deviations were calculated across age groups for wider age ranges such as, for example, 10 rather than only 5 years per group.

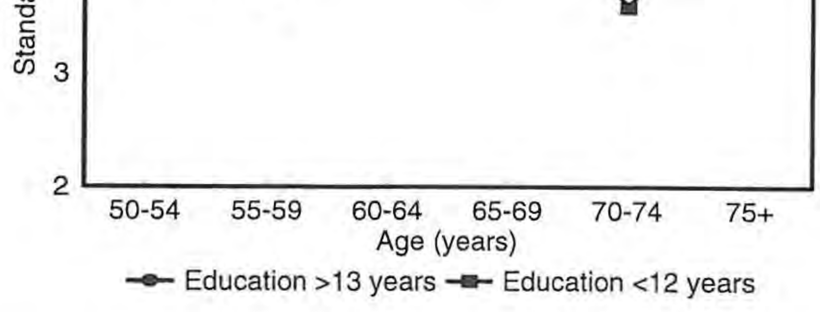


Figure 7.18. Animal Name Fluency: Standard deviations across age groups for two levels of education (data from Read, 1987, in Spreen & Strauss, 1991).

For *Boston Word Naming* (see Figure 7.19, based on data in Tables 7.7 and 7.8), for both studies included in the analysis (van Gorp et al., 1986, in Spreen & Strauss; Au et al., 1995) the comparisons of standard deviations across age groups revealed a highly significant effect ($p=<0.01$ in each instance), which supported a significant pattern of progressively increasing standard deviations across the adult age groups for both studies. For the van Gorp et al. data set, backwards comparisons revealed no significant difference between the 80-85 year-old age group and the 75-79 and 70-74 year-old age groups, but highly significant increases in the standard deviations between the 80-85 year-old age group and the 65-69 and 60-64 year-old age groups ($p=<0.001^{***}$, and 0.0017^{***}, respectively) which indicates that the significant increase occurs by the 70-74 year-old age group. For the Au et al. data set, backwards comparisons between the oldest 70's age group revealed highly significant increases in the standard deviation in this group compared with each of the younger 60's, 50's and 30's age groups (0.0028^{***}, $<0.001^{***}$, $<0.001^{***}$, respectively); progressive forwards comparisons revealed a significantly higher standard deviation in the 50's age group compared with the 30's age group ($p=0.0081^{**}$), no significant difference between the 60's and 50's age group, and a significantly higher standard deviation in the 70's age groups compared with the 60's age group ($p=0.0028^{**}$) which indicates an initial increase in variability by the 50's age group, and a second significant increase by the 70's age group. The additional significant increase at the 70's age decade for the Au et al. data set, is commensurate with the pattern noted for the van Gorp et al. data set, supporting a consistent pattern of progressively increasing variability in association with age particularly by the 70's for the Boston Word Naming Test.

7.4.2 Comparisons of Means

For the modality of language skills (see Table 7.7 for overall trends, and Table 7.8 for pairwise statistical comparisons), depending on the test there were both somewhat mixed or clearcut patterns of means scores in association with variability trends.

For the verbal fluency tests, as with the patterns of variability effects for these tests, there was a somewhat mixed picture. For the Read FAS Word Association Oral test < 12 and > 13 years conditions (see Figure 7.20 upper two graphs), in association with an erratic pattern of variability in the years up to the 60's followed by a clear pattern of variability by the 70's, there was a similarly erratic pattern of mean scores up to the 60's followed by a clear pattern

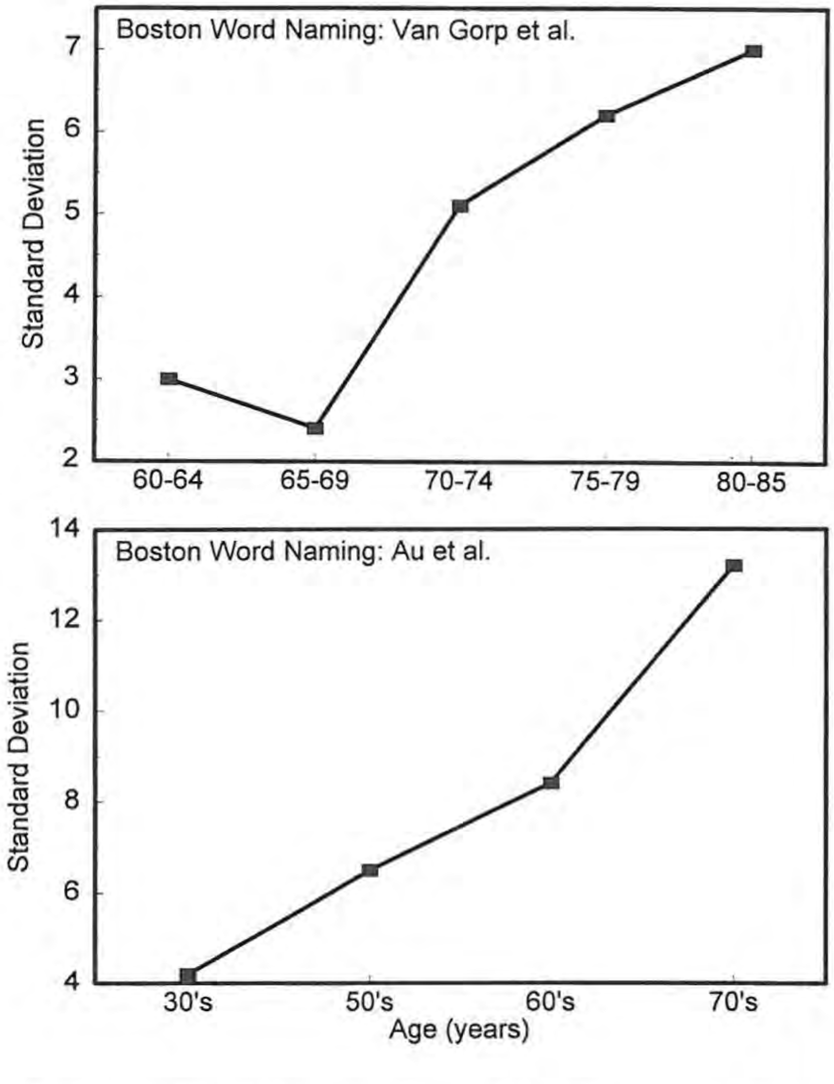


Figure 7.19. Boston Word Naming: Standard deviations across age groups for two data sets (van Gorp et al., 1986; Au et al., 1995).

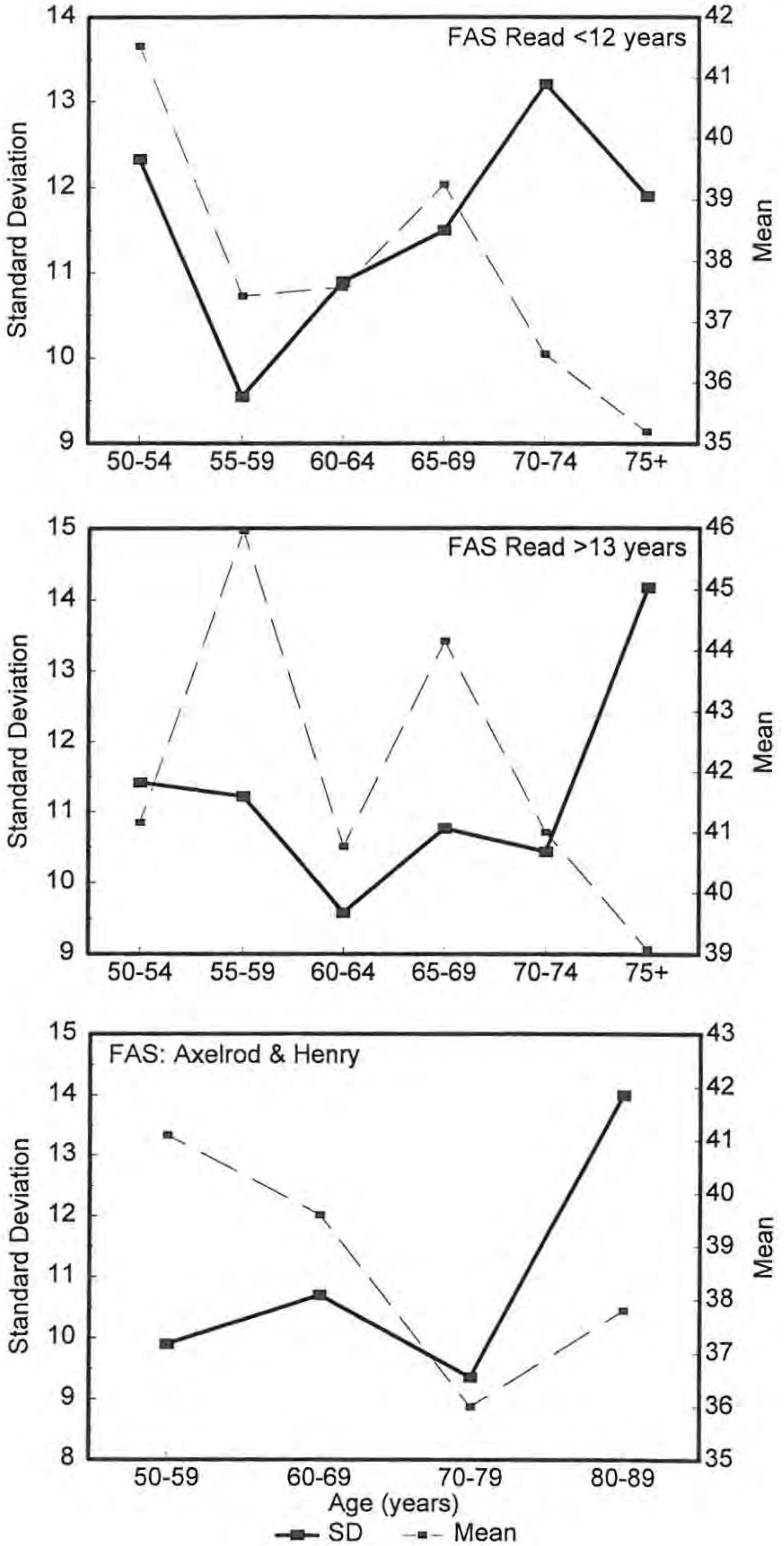


Figure 7.20. FAS Word Association (Oral): Standard deviations and mean scores across age groups for two educational conditions (data from Read, 1987, in Spreen & Strauss, 1991) (upper two graphs), and for data from Axelrod and Henry, 1992 (lower graph).

of declining mean scores by the 70's age groups. For the < 12 years educational condition, the pairwise comparisons revealed significantly poorer performance in the oldest age group compared with the youngest age group (0.0062**), but for the > 13 years condition there was no significant difference between the mean scores of the oldest and youngest age group. The particularly erratic up-down pattern in the mean scores for the FAS Word Association > 13 years educational condition until the 60's age group could lend support to the proposal made above that the erratic pattern in the variability data in the pre-60's age groups might be accounted for by a problematic sampling effect.

For the Axelrod and Henry FAS Word Association data set (see Figure 7.20 bottom graph), where, unlike in the case of Read's data sets, it is known that the age groups were well-matched for educational level, a sampling artifact is less of a possibility. Yet, in spite of this, there was a clear picture of progressively declining mean scores from the 50's through to the 70's, there was an upward curve in the 80's age group, all in association with progressively increasing standard deviations in association with age. (Since the Axelrod & Henry data set did not reveal any significant overall effect for variability, again the mean scores were not submitted to statistical analysis for this data set, and thus it is not possible to comment on the significance of the mean score trends). Thus generally, taking all the FAS data sets into consideration from both the Read and Axelrod and Henry studies (see Figure 7.20) the picture is *not* one of a very clear and consistent decline in average performance across the adult years. This suggests that there is a lot of inherent variability in performance on this particular task regardless of age stage and the extent to which groups are well matched for educational level.

For Animal Name Fluency, for which in particular there were no clear variability effects, there was however a consistent trend of steadily declining mean scores across age groups (see Table 7.7) and the mean score of the oldest age group was significantly lower than that of the youngest age group for the > 13 years subset of data that was statistically analyzed ($p = < 0.001^{***}$). (Since the variability effects for the < 12 years condition were not significant, pairwise comparisons of mean scores were not conducted for this subset of Animal Name Fluency data). In contrast to indications of particularly high inherent variability in the FAS Word Association test per se, it appears from the steadily declining mean in association with minimal alterations in variability across groups for the Animal Name Fluency test, that there is an unusual *absence* of inherent variability in the Animal

Naming Fluency test. This can be explained by the fact that an animal naming task taps material which would be *very* commonly available to all people regardless of levels of education, socio-economic status, gender and other possible influences. Thus it would not be likely to elicit such variable effects in interaction with these variables and with age, as would the more educationally dependent FAS verbal fluency task. However, since none of the verbal fluency tasks were tapped below the 50's age stage, it is not possible to know whether effects would occur across the entire age range even for Animal Naming Fluency.

In contrast to the mixed findings for the verbal fluency tasks, for the Boston Word Naming Test (see Figure 7.21) there was a highly consistent trend across both data sets (van Gorp et al. 1986, and Au et al. 1995) of steadily declining mean scores in association with progressively increasing variability, from younger to older age groups. For each data set, the mean of the oldest age group was significantly lower than that of the youngest age group, indicating significantly poorer performance in the oldest age groups for both studies ($<0.0069^{**}$ and $<0.001^{***}$, respectively).

7.4.3 Language skills: Summary of effects

Generally for language skills, as tapped by the FAS Word Association test and the Boston Word Naming Test, there was support for a sharp significant increase in variability by the 70's age decade in association with declining mean performance (that is the initial differentiation stage of the shuttle bulge effect), with no evidence of a significant subsequent decline in variability by this age stage for any of the tests or within-test conditions. In the only study for which age groups under 50 were tapped (Au et al. 1995, for the Boston Word Naming Test) it appeared that there was an earlier initial increase in variability by the 50's age decade, followed by a yet further sharp increase by the 70's.

For the Read (1991) FAS Word Association test a clear shuttle shift effect was detected in the variability data for educational level in the later age stages in that for the <12 educational condition the significant increase in variability occurred earlier than for the >13 years educational condition (left versus right shuttle effect, respectively). Further it appeared that the initial signs of a subsequent decrease in variability in association with aging was apparent for the <12 years educational condition which was not apparent for the >13 years educational condition, which further supported a left versus right shuttle effect, respectively,

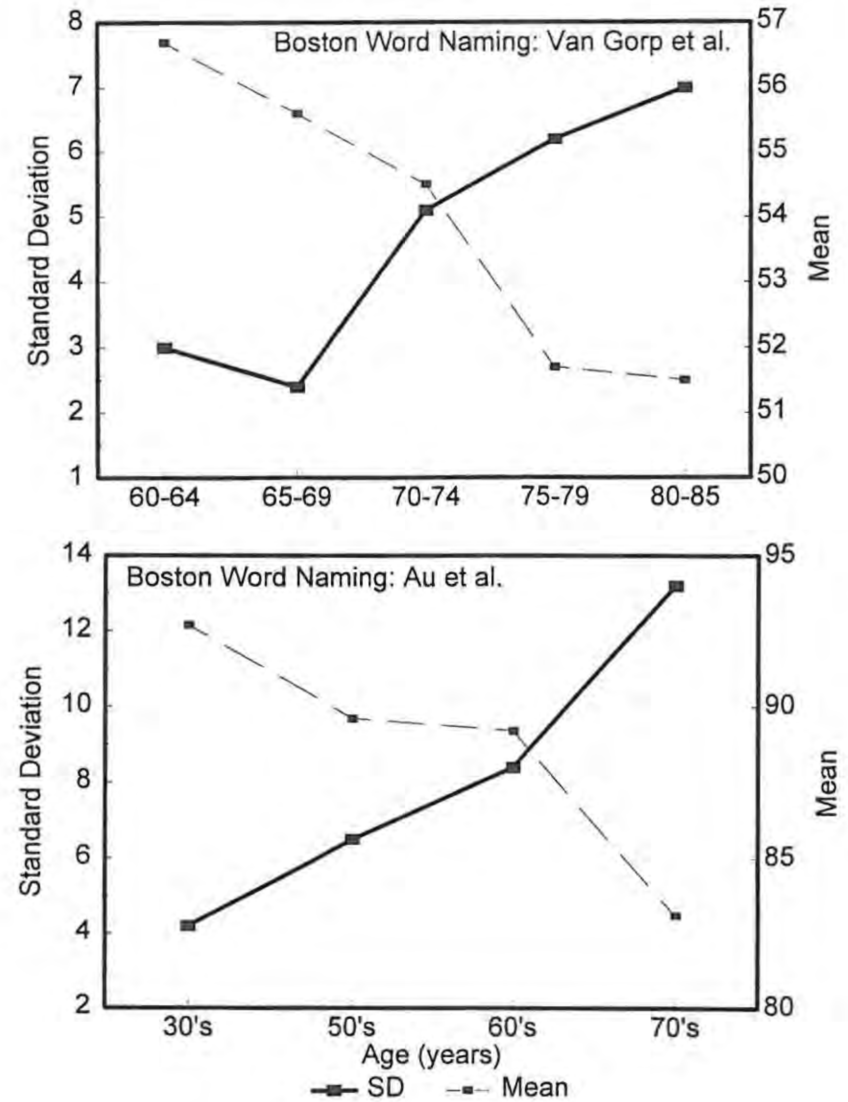


Figure 7.21. Boston Word Naming: Standard deviations and mean scores across age groups for two data sets (van Gorp et al., 1986; Au et al., 1995).

for these two subsets of FAS Word Association data. In addition this tendency suggests that were later age stages tapped for the Word Association test (i.e. into the 80's and later age stages), that a significant dedifferentiation effect would eventually become apparent for the low educational condition and subsequently that one would become apparent also for the higher educational condition.

All of these variability effects were in association with an overall decline in average ability between the younger and older age stages, although for the FAS Word Association test mean scores were very variable in the earlier old age groups (that is the 50's and 60's). For the

Animal Name Fluency task variability was relatively stable across the adult age range from 50 through to 70 in association with a steady decline in mean performance which could indicate a very marked absence of inherent variability for this task per se due to its common accessibility for all people; however, it is also possible that an increase in variability would be apparent between the 50+ age stage and the earlier adult years for this task had data with respect to earlier age groups been available for comparative purposes.

Table 7.7. LANGUAGE SKILLS: Mean Scores and Comparisons of Standard Deviations across all Age Groups.

FAS Word Association Oral (Read 1987, in Spreen & Strauss 1991)

AGE GROUP	50-54	55-59	60-64	65-69	70-74	75+	df	B[1]	p
FAS Education <12 years									
n	42	67	70	56	55	55	-	-	-
Mean	41.52	37.42	37.57	39.25	36.47	35.20	-	-	-
SD	12.33	9.55	10.90	11.50	13.21	11.90	5	7.20	0.0371*
FAS Education >13 years									
n	56	78	87	90	60	59	-	-	
Mean	41.16	45.96	40.76	44.16	41.00	39.08	-	-	-
SD	11.42	11.22	9.58	10.77	10.44	14.17	5	12.14	0.0034*

FAS Word Association Oral (Axelrod & Henry, 1992)

AGE GROUP	50-59	60-69	70-79	80-89	df	B	p
n	20	20	20	20	-	-	-
Mean	41.1	39.6	36.0	37.8	-	-	-
SD	9.9	10.7	9.3	14.0	3	3.82	0.1632

Animal Name Fluency (Read 1987, in Spreen & Strauss 1991)

AGE GROUP	50-54	55-59	60-64	65-69	70-74	75+	df	B	p
Animal Name Fluency Education <12 years									
n	42	67	70	56	55	55	-	-	-
Mean	18.64	18.81	18.61	17.75	16.35	15.09	-	-	-
SD	4.06	3.83	4.49	4.02	4.36	4.25	5	2.14	0.4082
Animal Name Fluency Education >13 years									
n	56	78	87	90	60	59	-	-	-
Mean	19.98	20.88	19.97	19.13	18.28	17.07	-	-	-
SD	4.08	5.14	5.00	5.01	3.58	4.93	5	12.27	0.0032*

[1] Bartlett's test statistic

Table 7.7 *continued*

Table 7.7. LANGUAGE SKILLS: Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

Boston Word Naming (Van Gorp et al. 1986, in Spreen & Strauss 1991)								
AGE GROUP	60-64	65-69	70-74	75-79	80-85	df	B	p
n	15	37	47	23	14	-	-	-
Mean	56.70	55.60	54.50	51.70	51.50	-	-	-
SD	3.00	2.40	5.10	6.20	7.00	4	35.47	<0.01**

Boston Naming Test: Experimental Version (Au et al. 1995)							
AGE GROUP	30's	50's	60's	70's	df	B	p
n	33	32	41	39	-	-	-
Mean	92.70	89.60	89.20	83.10	-	-	-
SD	4.20	6.50	8.40	13.20	3	43.56	<0.01**

Significant difference (*p<0.05; **p<0.01).

Table 7.8. LANGUAGE SKILLS: Pairwise Comparisons of Means between the oldest and youngest Age Groups and of Standard Deviations between the oldest Age Group and each younger Age Group unless otherwise specified.

FAS Word Association Oral/Education <12 years (Read 1987, in Spreen & Strauss 1991).									
AGE GROUP	50-54	55-59	60-64	65-69	70-74	75+	df	t/F	p
n	42	67	70	56	55	55	-	-	-
Mean	41.52	-	-	-	-	35.20	95	2.55	0.0062**
SD	-	-	-	-	13.21	11.90	54,54	1.23	0.2227
SD	-	-	-	11.50	-	11.90	55,54	0.93	0.4005
SD	-	-	10.90	-	-	11.90	69,54	0.84	0.2442
SD	-	9.55	-	-	-	11.90	66,54	0.64	0.0443
SD	12.33	-	-	-	-	11.90	41,54	1.07	0.3994
Forwards Comparisons									
SD	12.33	9.55	-	-	-	-	41,66	1.67	0.0317
SD	12.33	-	10.90	-	-	-	41,69	1.28	0.1809
SD	12.33	-	-	11.50	-	-	41,55	1.15	0.3117
SD	12.33	-	-	-	13.21	-	41,54	0.87	0.3252
SD	12.33	-	-	-	-	11.90	41,54	1.07	0.3994
Progressive Comparisons									
SD	12.33	9.55	-	-	-	-	41,66	1.67	0.0317
SD	-	9.55	10.90	-	-	-	66,69	0.77	0.1406
SD	-	-	10.90	11.50	-	-	69,55	0.90	0.3344
SD	-	-	-	11.50	13.21	-	55,54	0.76	0.1543
SD	-	-	-	-	13.21	11.90	54,54	1.23	0.2227
Additional Comparison									
SD	-	9.55	-	-	13.21	-	66,54	0.52	0.0061**

Table 7.8 *continued*

Table 7.8. LANGUAGE SKILLS: Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

FAS Word Association Oral/Education >13 years (Read 1987, in Spreen & Strauss 1991)

AGE GROUP	50-54	55-59	60-64	65-69	70-74	75+	df	t/F	p
n	56	78	87	90	60	59	-	-	-
Mean	41.16	-	-	-	-	39.08	113	0.86	0.1361
SD	-	-	-	-	10.44	14.17	59,58	0.54	0.0104*
SD	-	-	-	10.77	-	14.17	89,58	0.58	0.0097**
SD	-	-	9.58	-	-	14.17	86,58	0.46	0.0005***
SD	-	11.22	-	-	-	14.17	77,58	0.63	0.0277
SD	11.42	-	-	-	-	14.17	55,58	0.65	0.0545

Animal Name Fluency/Education >13 years (Read 1987 in Spreen & Strauss 1991)

AGE GROUP	50-54	55-59	60-64	65-69	70-74	75+	df	t/F	p
n	56	78	87	90	60	59	-	-	-
Mean	19.98	-	-	-	-	17.07	113	3.44	<0.001***
SD	-	-	-	-	3.58	4.93	59,88	0.53	0.0078**
SD	-	-	-	5.01	-	4.93	89,58	1.03	0.4533
SD	-	-	5.00	-	-	4.93	86,58	1.03	0.4597
SD	-	5.14	-	-	-	4.93	77,58	1.09	0.3709
SD	4.08	-	-	-	-	4.93	55,58	0.68	0.0797

Table 7.8 *continued*

Table 7.8. **LANGUAGE SKILLS:** Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

Boston Word Naming (van Gorp et al. 1986, in Spreen & Strauss 1991)								
AGE GROUP	60-64	65-69	70-74	75-79	80-85	df	t/F	p
n	15	37	47	23	14	-	-	-
Mean	56.7	-	-	-	51.5	27	2.63	<0.0069**
SD	-	-	-	6.2	7.0	22,13	0.78	0.2978
SD	-	-	5.1	-	7.0	46,13	0.53	0.0579
SD	-	2.4	-	-	7.0	36,13	0.12	<0.001***
SD	3.0	-	-	-	7.0	14,13	0.18	0.0017***

Boston Word Naming (Au et al. 1995)							
AGE GROUP	30's	50's	60's	70's	df	t/F	p
n	33	32	41	39	-	-	-
Mean	92.70	-	-	83.1	70	4.01	<0.001***
SD	-	-	8.4	13.2	40,38	0.41	0.0028***
SD	-	6.5	-	13.2	31,38	0.24	<0.001***
SD	4.2	-	-	13.2	32,38	0.10	<0.001***
Progressive Comparisons							
SD	4.2	6.5	-	-	32,31	0.42	0.0081**
SD	-	6.5	8.4	-	31,40	0.60	0.0712
SD	-	-	8.4	13.2	40,38	0.41	0.0028***

Significant Difference (*p <0.10; **p<0.05; ***p<0.01) as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

7.5 VISUAL SKILLS

For the modality of visual skills the data sets submitted to analysis are collated in Table A (Appendix pp.11-13), and a summary of the educational and/or IQ levels of the age groups appear in Table B (Appendix p.20). The following tests and accompanying sources of normative data sets were included in the analysis (ordered for discussion as they appear in the tables):

(i) **Right/Left Orientation:** One data set (Spreen & Strauss, unpublished study, in Spreen & Strauss, 1991).

(ii) **Embedded Figures Test:** One data set (Spreen & Strauss, unpublished study, in Spreen & Strauss, 1991).

(iii) **Rey-Osterrieth Complex Figure (RCF) Copy:** Two data sets (data collated from Kolb & Whishaw, 1985 and Strauss & Spreen, unpublished study, in Spreen & Strauss, 1991; Van Gorp, Satz., & Mitrushina, 1990).

(iv) **Trail Making Test:** Two data sets (Van Gorp, Satz., & Mitrushina, 1990; Stuss, Stethem, & Poirier, 1987).

(v) **Symbol Digit Modalities:** One data set (Centofanti, unreferenced study, in Lezak, 1983).

7.5.1 Comparisons of standard deviations across age groups

The visual tests can be subdivided conceptually between the relatively pure visuo-perceptual tasks comprising the Right/Left Orientation and Embedded Figures Tests, and the tests which focus on sustained visual search and tracking abilities at speed comprising the Trail Making Test and the Digit Symbol Modalities Test.

As a cluster the relatively pure visuo-perceptual tests are conspicuous for their highly consistent *absence* of regular or recognizable patterns in the progression of variability across the age groups. For *Right/Left Orientation* (Spreen & Strauss, unpublished study, in Spreen

& Strauss, 1991) (see Table 7.9), although the trend was non-linear with an apparent peak in variability for the middle groups, the comparisons of standard deviations across all groups very clearly revealed no significant effect (p=0.6289). For the *Embedded Figures Test* (Spreen & Strauss, unpublished study) (see Figure 7.22 based on data in Tables 7.9 and 7.10), there was an inconsistent pattern between the two submodalities of Total Correct and Total Credit with declining standard deviations across age groups for Total Correct, and an erratic up-down pattern for Total Credit.

In both instances there was a significant effect across all groups (p=0.0333, and p=<0.01). For Total Correct, backwards comparisons revealed no significant differences between the oldest 70+ age group and the 60's and 50's age group, but a significantly lower standard deviation between the 70+ age group than the 40's, 30's and 20's age groups (p=0.0090**, p=0.0090** p=0.0064**) which supported an unusual pattern of progressively declining variability across age groups. For Total Credit, backwards comparisons revealed no significant difference between the oldest 70+ age group and the 60's and 40's age groups, but a significantly higher standard deviation in the 70+ age group compared with the 50's, 30's and 20's age groups (p=<0.0017**, p=<0.001, respectively). This supported a zig-zag pattern of an initial increase in variability in the 40's age group which disappears in the 50's age group, but is followed by relatively high variability in the 60's and 70+ age groups.

In the case of both data sets for the *Rey Osterrieth Complex Figure (RCF) Copy* (data in Spreen & Strauss, 1991, and van Gorp et al., 1990) (see Figure 7.23, based on data in Tables 7.9, and 7.10) the comparisons of standard deviations across all groups revealed a highly significant effect (p=<0.01 in each instance). An erratic zig-zag pattern was supported for the Spreen and Strauss data set in that backwards comparisons revealed no significant difference between the 70+ age group and the 60-69 year-old age group, a significantly higher standard deviation for the 70+ age group compared with the 50-59 and 16-30 year-old age groups (p=0.0019***, p=0.0027**, respectively), and a significantly lower standard deviation for the 70+ age group than the 31-44 year-old age group (p=0.0073**). Another unusual pattern was evident for the van Gorp et al. data set in that backwards comparisons revealed a U-shaped progression in the variability data, with significantly higher standard deviation between the oldest 76-85 year-old age group than the 71-75 and 66-70 year-old age groups (p=0.0006*** and 0.0010***, respectively) and no significant difference between the 76-85 and 57-65 year-old age groups.

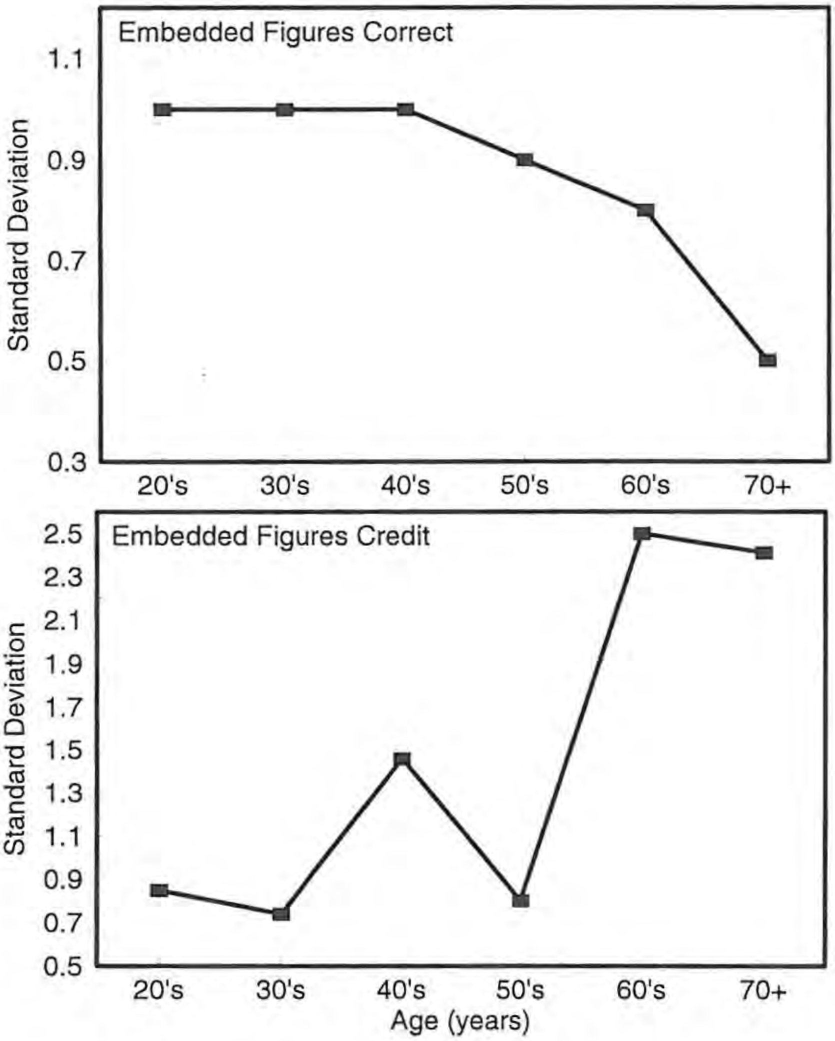


Figure 7.22. Embedded Figures Test: Standard deviations across age groups for Total Correct and Total Credit (data from Spreen & Strauss, unpublished study, in Spreen & Strauss, 1991).

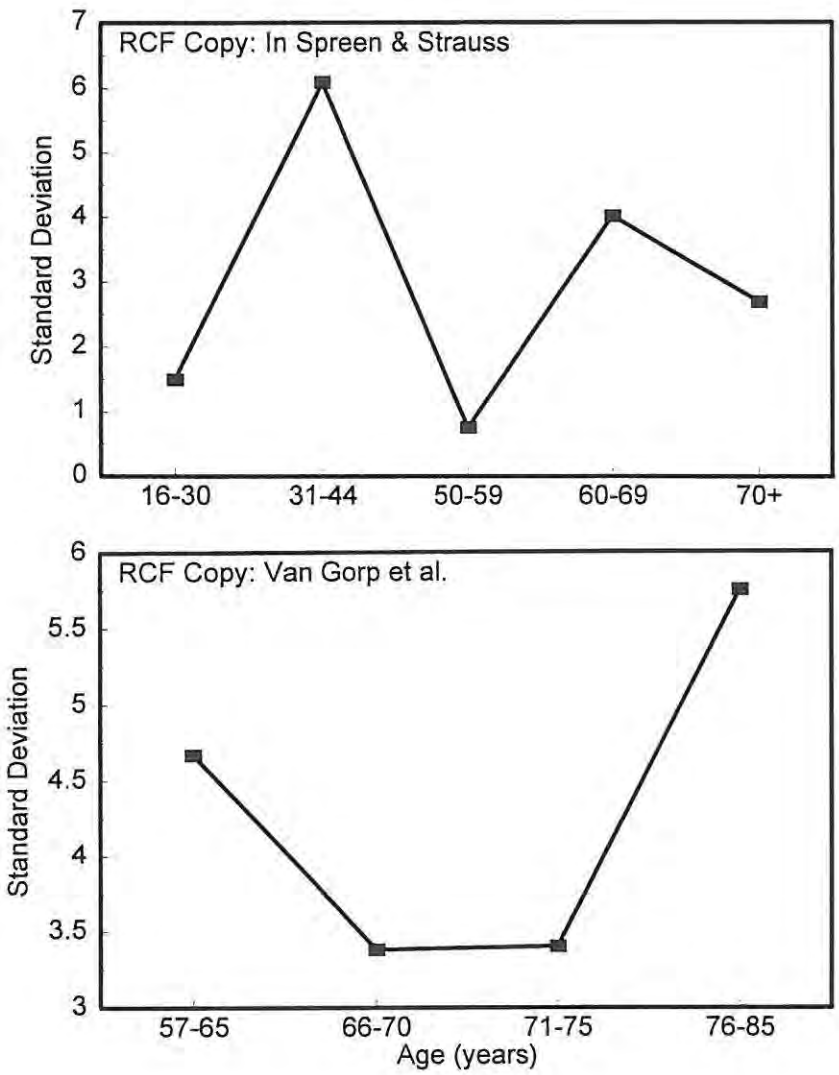


Figure 7.23. Rey Complex Figure (RCF) Copy: Standard deviations across age groups for two data sets (collated data in Spreen & Strauss, 1991; Van Gorp, Satz, & Mitrushina, 1990).

As a group (compared with the cluster of pure visuo-spatial skills) the visuo-perceptual speeded tasks comprising the *Trail Making Test Parts A and B*, and the *Symbol Digit Modalities Test Oral and Written*, present a much more coherent picture of variability patterns across age groups (see Figures 7.24 and 7.25, based on data in Tables 7.9 and 7.10). Each of the three data sets (van Gorp et al.1990, Stuss et al. 1987, and Centofanti, unreferenced study in Lezak, 1983, respectively) revealed highly consistent within-study patterns of effects for the two levels of presentations in each of the tests. In all cases the comparisons of standard deviations across groups revealed highly significant effects ($p < 0.01$ in each instance).

For the first Trail Making Test study (van Gorp et al. 1990) (See Figure 7.2, upper graph), a pattern of significantly increased variability between the younger and older age groups was supported in that backwards comparisons for Trails A revealed a significantly higher standard deviation for the oldest 76-85 year-old age group compared with the youngest 57-65 year-old age group ($p=0.001^{***}$), and no significant difference between any of the other pairs. Backwards comparisons for Trails B revealed a significantly higher standard deviation for the 76-85 year-old age group and each of the younger age groups ($p= < 0.001^{***}$, $p=0.0166^{**}$, and $p= < 0.001^{***}$). Overall this indicated that there was a sustained significant increase in variability for both Trails A and B from the 66-70 age group. Progressive advancing comparisons revealed that the tendency for a dip in the standard deviations which occurred for both Trails A and B, was not significant for Trails A, but was significant for Trails B (there was a significantly lower standard deviation in the 71-75 year-old age group compared with the 66-70 year-old age group, $p=0.0060^{**}$). However, scrutiny of the demographic data for the van Gorp et al. data set, (see Appendix Table B, note 5, p.21) indicates that *there is a clear peak of at least 10 Verbal and 10 Performance IQ points for the 71-75 year-old age group compared with all the other age groups*, and hence the dip in the overall pattern of progressively increasing variability across the age groups for the van Gorp et al. data can be seen as an artifact of this disproportionately high level of education in this age group, and does not intrinsically negate the overall effect of increasing variability from younger to older age groups for this test.

For the second Trail Making Test Study (Stuss et al. 1987) (See Figure 7.24, lower graph), progressive comparisons for Trails A revealed no significant differences between advancing pairs until the 50-59 year-old age group which was significantly higher than the 40-49 year-

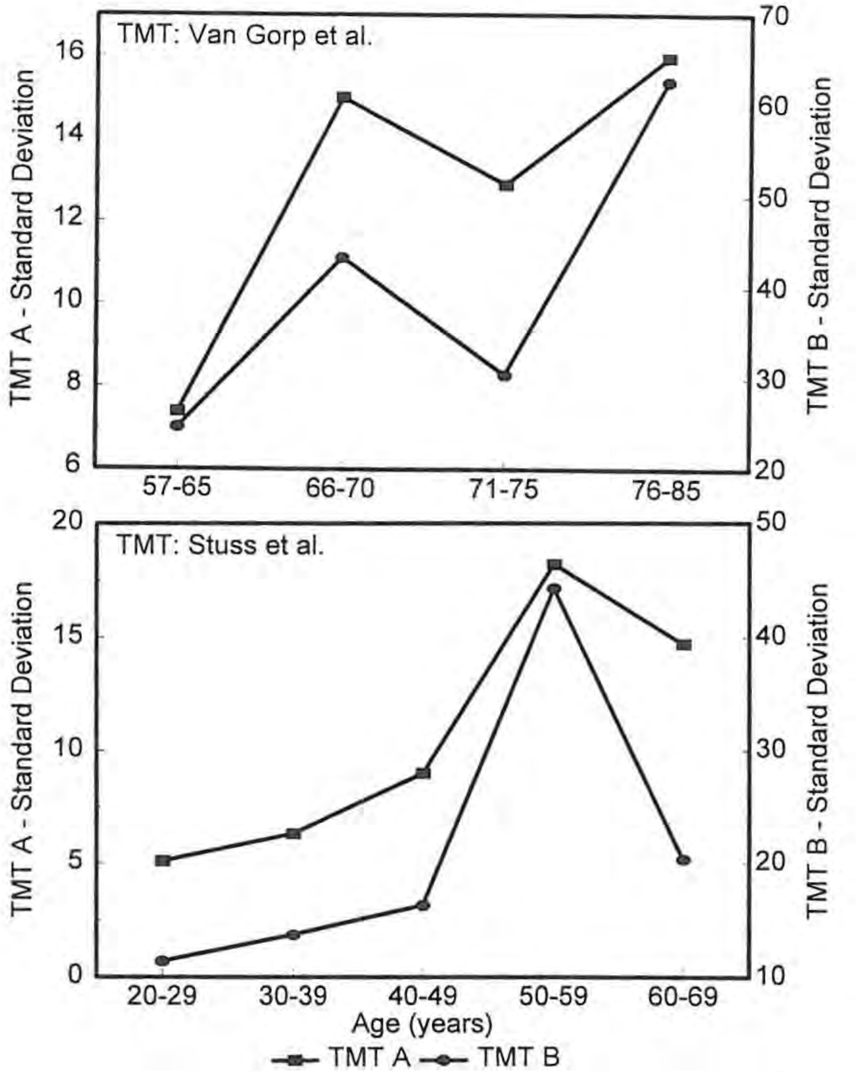


Figure 7.24. Trail Making Test Parts A and B: Standard deviations across age groups for two data sets (Van Gorp, Satz., & Mitrushina, 1990; Stuss, Stethem, & Poirier, 1987).

old age group (p=0.0238*), and no significant difference between the 50-59 year-old and 60-69 year-old age groups. Similarly progressive comparisons for Trails B revealed no significant differences between any of the advancing age groups until the 50-59 year-old age group which was significantly higher than the 40-49 year-old age group (p=0.0032**), but in contrast to Trails A there was a significantly lower standard deviation between the 60-69 year-old age group than the 50-59 year-old age group. Thus at first glance the Stuss et al. data set appeared to reveal a significant non-linear inverted-U pattern of variability across the age groups for Trails B which occurred at the relatively early age stage of the 50's age decade. However, scrutiny of the demographic data for the Stuss et al. data set (see Appendix Table B, p.20) indicates that *the 50-59 year-old age group has an educational level of only 12 years compared with levels of 16, 17, and 14 years for each of the other age* *groups*. Hence the peak at the 50's age stage caused the artificial impression of an early dedifferentiation effect in the 60's age group for Trails B, and can be better understood as an artifact of the disproportionately low level of education for this group. If the peak in the pattern of variability at the 50's age group were removed, a pattern of progressively increasing variability up to the 60's age group would be revealed for both Trails A and Trails B which would be entirely consistent with that revealed on the Trail Making Test for the van Gorp et al. data. (It is of note that the disproportionate educational levels across groups in the Stuss et al. and van Gorp et al. data do not appear to have caused obvious aberrant patterns of effects for the PASAT and the Wechsler Memory Scale subtests, respectively, due to unbalanced levels of education across groups. This suggests that variability is particularly sensitive to educational effects on the Trail Making Test).

For the *Symbol Digit Modalities Test* (Centofanti, in Lezak, 1983) (see Figure 7.25, based on data in Tables 7.9 and 7.10), progressive comparisons revealed that for both the oral and written versions of the test there was no significant increase between any of the advancing age groups until the 65-74 year-old group which was significantly higher than the 55-64 year old group ($p=0.0036^{**}$ and $p=0.0130^{*}$). This indicates that there was a clear-cut pattern of increased variability across the adult age range which took effect in the 65-74 year-old age group on this test, which is consistent with the overriding pattern that was established above for the Trail Making Test.

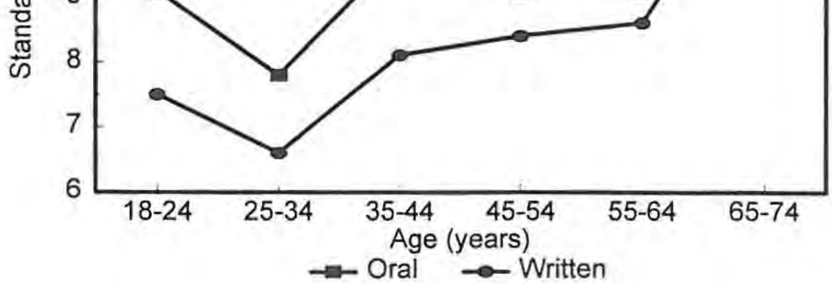


Figure 7.25. Symbol Digit Modalities (oral and written): Standard deviations across age groups (data from Centofanti, unreferenced study, in Lezak, 1983).

7.5.2 Comparisons of means

For the set of relatively pure visuo-perceptual skills which as a cluster showed significant yet erratic variability effects (see Table 7.9, for overall trends, and Table 7.10 for pairwise statistical comparisons) there was also a somewhat erratic tendency for the means. For the Embedded Figures Test Total Correct (see Figure 7.26, upper graph) there was a fluctuating tendency across the age groups with a significantly increased mean in the oldest 70+ age group relative to the youngest age group ($p = < 0.001^{***}$) indicating an *improved* score for the oldest age group. For the Embedded Figures Test Total Credit (see Figure 7.26, lower graph) there was an overall decline in mean scores across age groups with a significantly lower mean score in the oldest 70+ age group relative to the youngest age group ($p = 0.0069^{**}$). It is of note that for the Embedded Figures Total Correct submodality, this highly erratic pattern of mean scores occurred in association with a highly unusual pattern of steadily declining variability in association with aging. Further, for the Embedded Figures Total Credit submodality (see Figure 7.26, lower graph), it is of note that once there were signs of clearly declining mean scores from the 50's age group onwards, the erratic pattern of variability was no longer present and the declining mean score tendency occurred in association with the recognizable pattern of a clear pattern of significantly increased variability. It appears likely that the difference in effects between the two submodalities is explicable in terms of age-sensitivity of the tasks. Whereas both tasks involve visual search within a specified time limit, the score for the Total Correct submodality is much less heavily weighted on the ability to perform at speed than the Total Credit submodality. As age advanced in the latter task, which is the more speed-dependent and therefore arguably more age-sensitive task, there was clear evidence of a declining mean and associated enhanced variability.

For the Rey Complex Figure Copy in association with the inconsistent patterns of variability for both data sets (see Figure 27) there was also a fluctuating pattern in the progression of the mean scores across age groups. For the Spreen and Strauss data set, the mean for the oldest 70+ age group was significantly lower than the youngest 16-30 year-old age group ($< 0.001^{***}$) indicating diminished ability in the oldest age group, although the interim pattern was highly fluctuating with a steady decline only becoming apparent after the 50's age group. For the van Gorp et al. data, although there was the trend of consistently declining mean scores from the 66-70 year-old age group, the mean for the oldest 76-85 year-old age group

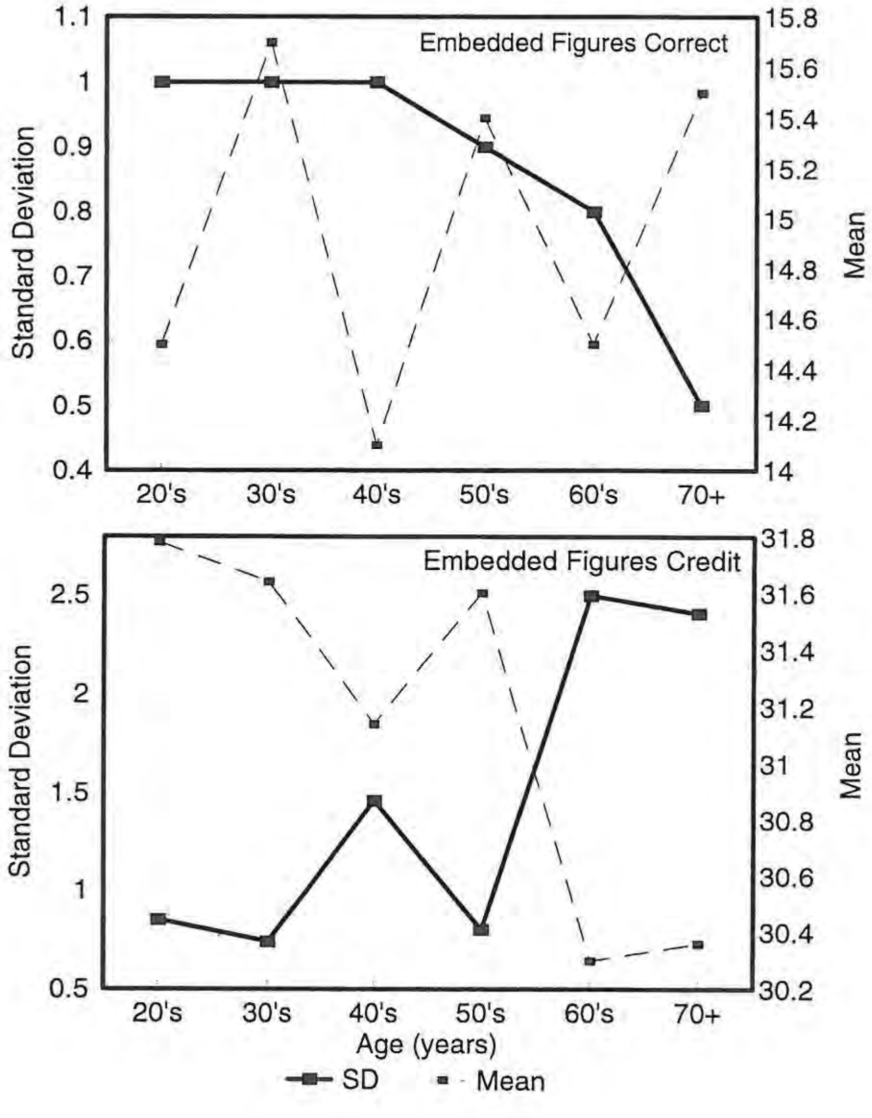


Figure 7.26. Embedded Figures Test: Standard deviations and mean scores across age groups for Total Correct and Total Credit.

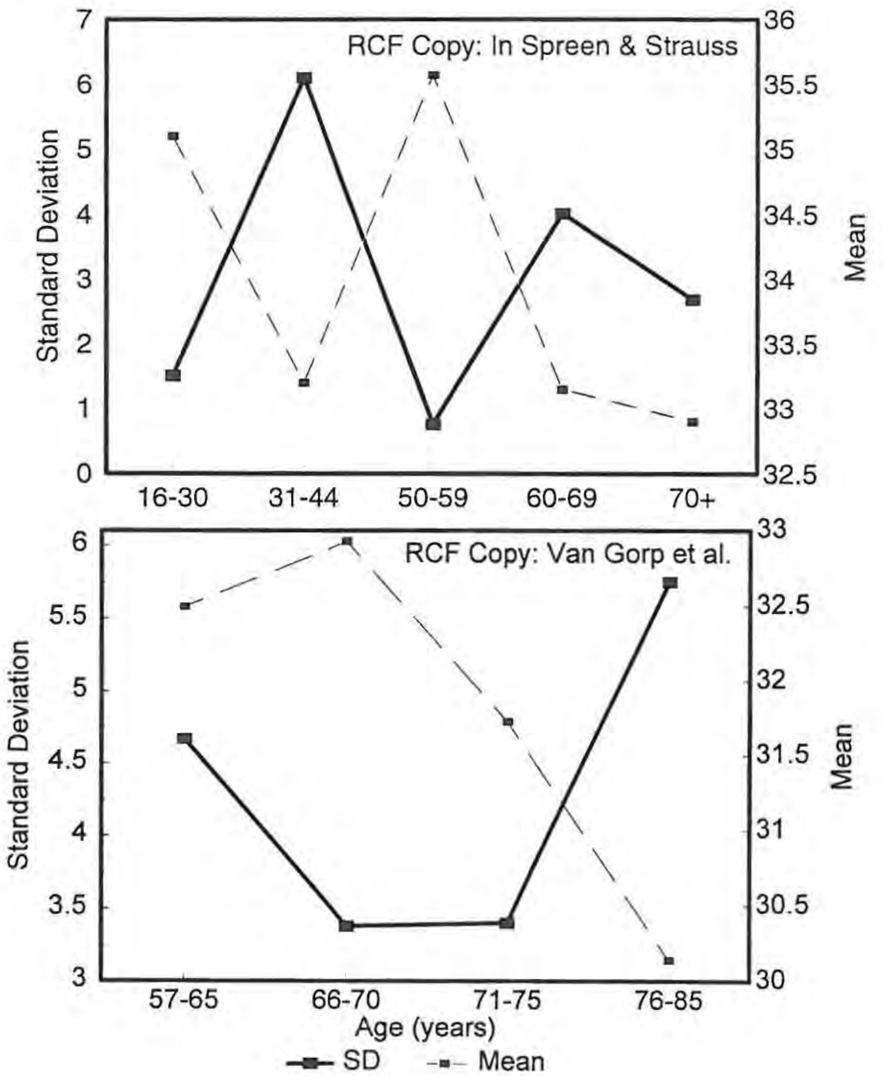


Figure 7.27. Rey Complex Figure (RCF) Copy: Standard deviations and mean scores across age groups for two data sets (collated data in Spreen & Strauss, 1991; Van Gorp, Satz, & Mitrushina, 1990).

was not significantly lower than the youngest 57-65 year-old age group (0.0514). What is very clear from both the Embedded Figures and RCF Copy data sets (see Figures 7.26 and 7.27), is that whilst the mean fluctuated or failed to show clear indications of decline, there were also erratic and fluctuating tendencies in the variability. On the other hand, once the mean showed a clear-cut decline consistently there were indications of significant increases in variability in association with aging.

For the cluster of timed sustained visual search and tracking tasks, in association with progressively increasing variability across the adult age groups, there was a highly consistent pattern of significant increases in the timed score means (indicating a decrease in the levels of ability) between the oldest and youngest age groups for Trail Making Test A and B for both the van Gorp et al., and Stuss et al. data sets ($p = < 0.001^{***}$ in all instances), and decreases in means (indicating a decrease in the levels of ability) between the oldest 65-74 year-old age groups and the youngest 18-24 year-old age groups for the Symbol Digit Modalities Test Oral and Written versions ($p = < 0.001^{***}$ in both instances). These consistent effects are selectively illustrated for the Trail Making Test B (van Gorp et al.) and Symbol Digit Modalities (Centofanti) in Figure 7.28.

For the Trail Making Test the influence of unmatched IQ or educational levels across groups as was proposed above, is generally supported by an associated disruption in the expected pattern of declining mean performance across the age groups (see Table 7.9). For the Trail Making Test van Gorp et al. data set, the proposed artifactual effect on the variability pattern of a disproportionately higher level of IQ in the 71-75 year-old age group, is reflected by a relative improvement in mean performance (decreased time score)in that age group for part B of the test (see Figure 7.28, upper graph). Similarly for the Stuss et al. Trail Making Test data set (see data in Table 7.9), the proposed artifactual effect on the variability pattern of a disproportionately lower level of education in the 50's age group is clearly confirmed by a relative slump in mean performance (higher time score) compared with the 60's age group for both parts A and B of the test.

7.5.3 Visual skills: Summary of effects

Generally for the relatively pure visuo-perceptual skills, as tapped by the Right-Left Orientation and Embedded Figures Tests, there was an erratic pattern of variability trends in association with age. However, the absence of a clear pattern of variability effects was always in association with the absence of a clear declining trend in average ability. Hence this did not stand in contradiction to the tenets of the shuttle model within which it is proposed that a variability bulge will occur in association with an overall aging effect. In other words the haphazard pattern of variability occurred *consistently* whenever there was

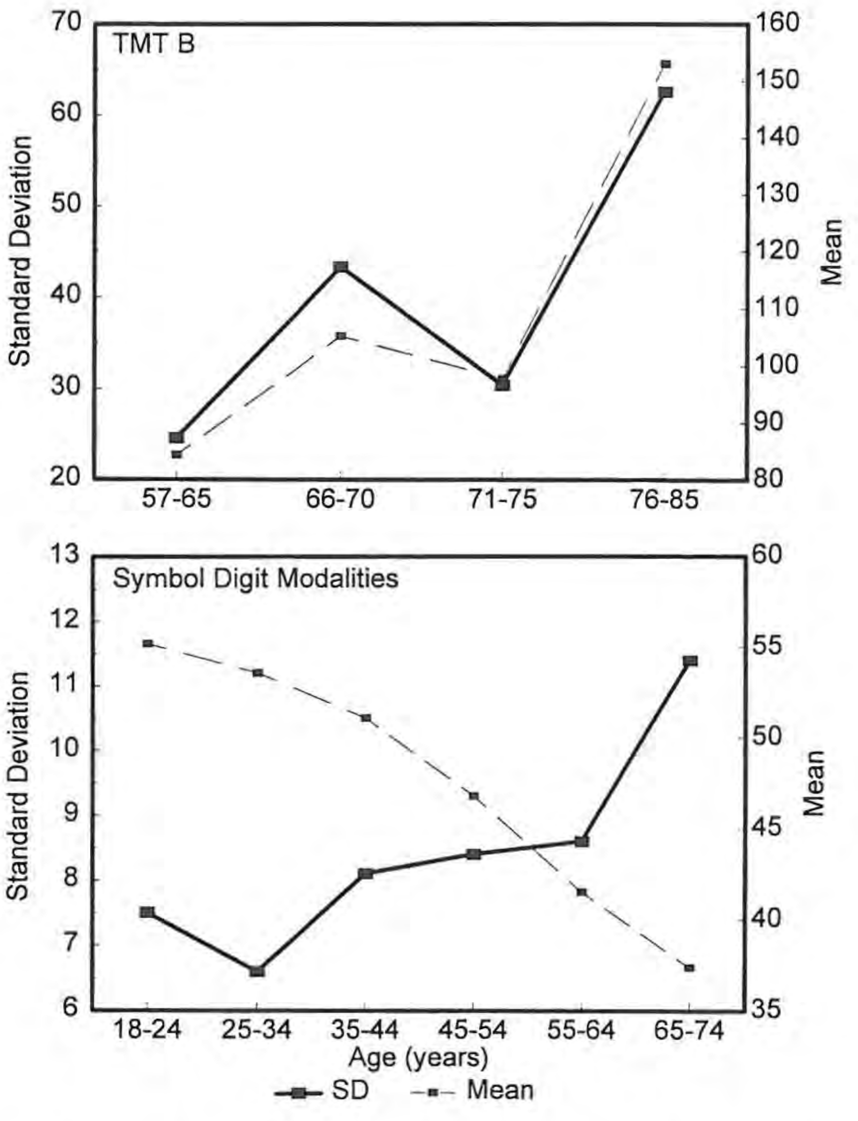


Figure 7.28. Trail Making Test (TMT) B and Symbol Digit Modalities: Standard deviations and mean scores across age groups for data from Van Gorp, Satz, & Mitrushina, 1990 (TMT B) and from Centofanti, unreferenced study, in Lezak, 1983 (Symbol Digit Modalities).

evidence that the test modality was not showing signs of being particularly age sensitive. Where a speed factor is called upon more heavily in a task (as for the Embedded Figures Total Credit submodality), rendering the task a greater degree of age-sensitivity, a clearly declining mean score was apparent in association with a significant increase in variability (that is the initial differentiation stage of the shuttle bulge effect), by the 60's age stage.

For the visual tracking and sustained visual attention tasks as tapped by the Trail Making Test and the Symbol Digit Modalities test, taking into consideration likely sample artifacts, generally there was consistent support for a significant increase in variability by the mid to late 60's age stage in association with declining mean performance, with no clear support for a significant subsequent decline in variability by this age stage for any of the tests or within-test conditions. (This is bearing in mind the argument posed above, which is that the Stuss et al. 1987 data set probably revealed the trend of an overall increase up to the 60's age group rather than an inverted-U effect, due to an artifactual rise in variability effected by a lower educational level in the 50's age group).

Thus the visual tasks generally (including both the visual perceptual tasks and speeded visual tracking tasks) provided support for the presence of the initial differentiation phase of the shuttle bulge effect in association with declining mean performance. Where tasks were not clearly age-sensitive, consistently also there were no clear-cut variability effects. Since there were no indications of the earlier onset of variability for the Trails A relative to the presumably more age-sensitive Trails B tasks in either case, there was no clear evidence of a shuttle shift effect for level of task challenge for either of the Trail Making data sets. Rather the variability effects for both studies appeared to occur simultaneously across the two submodalities of the task.

Table 7.9. VISUAL SKILLS: Mean Scores and Comparisons of Standard Deviations across all Age Groups.

Right/Left Orientation (Spreen & Strauss unpublished study, in Spreen & Strauss 1991)

AGE GROUP	20-30	50-59	60-69	70-79	80+	df	B[1]	p
n	20	19	27	23	15	-	-	-
Mean	17.82	17.16	16.78	15.17	15.00	-	-	-
SD	2.40	2.89	2.98	2.66	2.70	4	1.12	0.6289

Embedded Figures Test (Spreen & Strauss unpublished study, in Spreen & Strauss 1991)

AGE GROUP	20's	30's	40's	50's	60's	70's	df	B	p
n	23	14	14	10	10	14	-	-	-
Embedded Figures-Total Correct									
Mean	14.5	15.7	14.1	15.4	14.5	15.5	-	-	-
SD	1.0	1.0	1.0	0.9	0.8	0.5	5	7.43	0.0333*
Embedded Figures-Total Credit									
Mean	31.78	31.64	31.14	31.60	30.30	30.36	-	-	-
SD	0.85	0.74	1.46	0.80	2.50	2.41	5	37.18	<0.01**

Rey-Osterrieth Complex Figure Copy (Collated data in Spreen & Strauss 1991)

AGE GROUP	16-30	31-44	50-59	60-69	70+	df	B	p
n	67	26	14	13	10	-	-	-
Mean	35.10	33.20	35.57	33.15	32.90	-	-	-
SD	1.50	6.10	0.76	4.02	2.69	4	106.19	<0.01**

Rey-Osterrieth Complex Figure Copy (Van Gorp et al. 1990)

AGE GROUP	57-65	66-70	71-75	76-85	df	B	p
n	28	45	57	26	-	-	-
Mean	32.50	32.93	31.73	30.14	-	-	-
SD	4.67	3.38	3.40	5.75	3	14.46	<0.01**

[1] Bartlett's test Statistic.

Table 7.9 *continued*

Table 7.9. VISUAL SKILLS: Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

Trail Making Test (van Gorp et al. 1990) AGE GROUP			57-65	66-70	71-75	76-85	df	B	p
n			28	45	57	26	-	-	-
TMT A									
Mean			41.50	43.20	50.08	59.73	-	-	-
SD			7.38	14.98	12.88	15.95	3	16.25	<0.01**
TMT B									
Mean			84.40	105.20	97.79	153.09	-	-	-
SD			24.60	43.43	30.40	62.60	3	30.59	<0.01**

Trail Making Test (Stuss et al. 1987) AGE GROUP		20-29	30-39	40-49	50-59	60-69	df	B	p
n		10	10	10	10	10	-	-	-
TMT A									
Mean		18.5	21.9	29.2	38.5	37.5	-	-	-
SD		5.1	6.3	9.0	18.2	14.7	4	18.55	<0.01**
TMT B									
Mean		41.6	46.3	64.1	83.1	73.3	-	-	-
SD		11.4	13.7	16.3	44.3	20.3	4	22.63	<0.01**

Symbol Digit Modalities (Centofanti unreferenced study, in Lezak 1983) AGE GROUPS	18-25	25-34	35-44	45-54	55-64	65-74	df	B	p
n	69	72	76	75	67	61	-	-	-
Symbol Digit-Oral									
Mean	62.7	61.2	59.7	54.5	48.4	46.2	-	-	-
SD	9.1	7.8	9.7	9.1	9.1	12.8	5	18.55	<0.01**
Symbol Digit-Written									
Mean	55.2	53.6	51.1	46.8	41.5	37.4	-	-	-
SD	7.5	6.6	8.1	8.4	8.6	11.4	5	22.64	<0.01**

Significant difference (*p<0.05; **p<0.01).

Table 7.10. VISUAL SKILLS: Pairwise Comparisons of Means between the oldest and youngest Age Groups and of Standard Deviations between the oldest Age Group and each younger Age Group unless otherwise specified.

Embedded Figures Test (Spreen & Strauss unpublished study, in Spreen & Strauss 1991)

AGE GROUP	20's	30's	40's	50's	60's	70+	df	t/F	p
n	23	14	14	10	10	14	-	-	-
Embedded Figures-Total Correct									
Mean	14.5	-	-	-	-	15.5	35	-3.47	<0.001***
SD	-	-	-	-	0.8	0.5	9,13	2.56	0.0651
SD	-	-	-	0.9	-	0.5	9,13	3.24	0.0283
SD	-	-	1.0	-	-	0.5	13,13	4.00	0.0090**
SD	-	1.0	-	-	-	0.5	13,13	4.00	0.0090*
SD	1.0	-	-	-	-	0.5	22,13	4.00	0.0064**
Embedded Figures-Total Credit									
Mean	31.78	-	-	-	-	30.36	35	35	0.0069**
SD	-	-	-	-	2.50	2.41	9,13	1.08	0.4385
SD	-	-	-	0.80	-	2.41	9,13	0.11	0.0017**
SD	-	-	1.46	-	-	2.41	13,13	0.37	0.0411
SD	-	0.74	-	-	-	2.41	13,13	0.09	<0.001***
SD	0.85	-	-	-	-	2.41	22,13	0.12	<0.001***

Rey Osterrieth Complex Figure Copy (Collated data in Spreen & Strauss 1991)

AGE GROUP	16-30	31-44	50-59	60-69	70+	df	t/F	p
n	67	26	14	13	10	-	-	-
Mean	35.10	-	-	-	32.90	75	3.85	<0.001***
SD	-	-	-	4.02	2.69	12,9	2.23	0.1170
SD	-	-	0.76	-	2.69	13,9	7.98	0.0019***
SD	-	6.10	-	-	2.69	25,9	5.14	0.0073**
SD	1.5	-	-	-	2.69	66,9	0.31	0.0027**

Table 7.10 *continued.*

Table 7.10. VISUAL SKILLS: Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

Rey-Osterrieth Complex Figure Copy (van Gorp et al. 1990)							
AGE GROUP	57-65	66-70	71-75	76-85	df	t/F	p
n	28	45	57	26	-	-	-
Mean	32.50	-	-	30.14	52	1.66	0.0514
SD	-	-	3.40	5.75	56,25	0.35	0.0006***
SD	-	3.38	-	5.75	44,25	0.35	0.0010***
SD	4.67	-	-	5.75	27,25	0.66	0.1459

Trail Making Test (van Gorp et al. 1990)							
AGE GROUP	57-65	66-70	71-75	76-85	df	t/F	p
n	28	45	57	26	-	-	-
TMT A							
Mean	41.50	-	-	59.73	52	-5.46	<0.001***
SD	-	-	12.88	15.95	56,25	0.65	0.0930
SD	-	14.98	-	15.95	44,25	0.88	0.3497
SD	7.38	-	-	15.95	27,25	0.21	<0.001***
TMT B							
Mean	84.40	-	-	153.09	52	-5.38	<0.001***
SD	-	-	30.40	62.60	56,25	0.24	<0.001***
SD	-	43.43	-	62.60	44,25	0.48	0.0166**
SD	24.60	-	-	62.60	27,25	0.15	<0.001***
Progressive Comparisons							
TMT A							
SD	7.38	14.98	-	-	27,44	0.24	0.001***
SD	-	14.98	12.88	-	44,56	1.35	0.1421
SD	-	-	12.88	15.95	56,25	0.65	0.0929
TMT B							
SD	24.60	43.43	-	-	27,44	0.32	<0.001***
SD	-	43.43	30.40	-	44,56	2.04	0.0060**
SD	-	-	30.40	62.60	56,25	0.24	<0.001***

Table 7.10. *continued.*

Table 7.10. VISUAL SKILLS: Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

Trail Making Test (Stuss et al. 1987) AGE GROUP	20-29	30-39	40-49	50-59	60-69	df	t/F	p
n	10	10	10	10	10	-	-	-
TMT A								
Mean	18.5	-	-	-	37.3	18	-3.82	<0.001***
SD	-	-	-	18.2	14.7	9,9	1.53	0.2673
SD	-	-	9.0	-	14.7	9,9	0.37	0.0800
SD	-	6.3	-	-	14.7	9,9	0.18	0.0094**
SD	5.1	-	-	-	14.7	9,9	0.12	0.0021***
TMT B								
Mean	41.6	-	-	-	73.3	18	-4.31	<0.001***
SD	-	-	-	44.3	20.3	9,9	4.76	0.0147*
SD	-	-	16.3	-	20.3	9,9	0.65	0.2617
SD	-	13.7	-	-	20.3	9,9	0.46	0.1285
SD	11.4	-	-	-	20.3	9,9	0.32	0.0504
Progressive Comparisons								
TMT A								
SD	5.1	6.3	-	-	-	9,9	0.66	0.2695
SD	-	6.3	9.0	-	-	9,9	0.49	0.1514
SD	-	-	9.0	18.2	-	9,9	0.25	0.0238*
SD	-	-	-	18.2	14.7	9,9	1.53	0.2673
TMT B								
SD	11.4	13.7	-	-	-	9,9	0.69	0.2964
SD	-	13.7	16.3	-	-	9,9	0.49	0.1508
SD	-	-	16.3	44.3	-	9,9	0.14	0.0032**
SD	-	-	-	44.3	20.3	9,9	4.76	0.0147*

Table 7.10 *continued*

Table 7.10. VISUAL SKILLS: Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

Symbol Digit Modalities (Centofanti, unreferenced study in Lezak 1983)									
AGE GROUP	18-24	25-34	35-44	45-54	55-64	65-74	df	t/F	p
n	69	72	76	75	67	61	-	-	-
Symbol Digit-Oral									
Mean	62.7	-	-	-	-	46.2	128	8.54	<0.001***
SD	-	-	-	-	9.1	12.8	66,60	0.51	0.0036**
SD	-	-	-	9.1	-	12.8	74,60	0.51	0.0070**
SD	-	-	9.7	-	-	12.8	75,60	0.57	0.0115*
SD	-	7.8	-	-	-	12.8	71,60	0.37	<0.001***
SD	9.1	-	-	-	-	12.8	68,60	0.51	0.0033**
Symbol Digit-Written									
Mean	55.2	-	-	-	-	37.4	128	10.63	<0.001***
SD	-	-	-	-	8.6	11.4	66,60	0.57	0.0130*
SD	-	-	-	8.4	-	11.4	74,60	0.54	0.0063**
SD	-	-	8.1	-	-	11.4	75,60	0.50	0.0026**
SD	-	6.6	-	-	-	11.4	71,60	0.34	<0.001***
SD	7.5	-	-	-	-	11.4	68,60	0.43	0.0005***
Progressive Comparisons									
Symbol Digit-Oral									
SD	9.1	7.8	-	-	-	-	68,71	1.36	0.1001
SD	-	7.8	9.7	-	-	-	71,75	0.65	0.0327
SD	-	-	9.7	9.1	-	-	75,74	1.14	0.2917
SD	-	-	-	9.1	9.1	-	74,66	1.00	0.5018
SD	-	-	-	-	9.1	12.8	66,60	0.51	0.0036**
Symbol Digit-Written									
SD	7.5	6.6	-	-	-	-	68,71	1.29	0.1439
SD	-	6.6	8.1	-	-	-	71,75	0.66	0.0416
SD	-	-	8.1	8.4	-	-	75,74	0.93	0.3769
SD	-	-	-	8.4	8.6	-	74,66	0.95	0.4205
SD	-	-	-	-	8.6	11.4	66,60	0.57	0.0130*

Significant difference (*p<0.10; **p<0.05; ***p<0.01) as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

7.6 PROBLEM SOLVING

For the modality of problem solving the data sets submitted to analysis are collated in Table A (Appendix pp.13-15), and a summary of the educational and/or IQ levels of the age groups appear in Table B (Appendix p.20). The following tests and accompanying sources of normative data sets were included in the analysis (ordered for discussion as they appear in the tables):

> **(i) Wisconsin Card Sorting:** Three data sets (Data collated from Heaton, 1981, and Strauss & Spreen, unpublished study, in Spreen & Strauss, 1991; Beatty, 1993; Axelrod & Henry, 1992).
>
> **(ii) Stroop Test:** Two data sets (Data collated from Regard, 1981, and Spreen & Strauss, unpublished study, in Spreen & Strauss, 1991; Troyer, Graves, & Cullum, 1994).
>
> **(iii) Similarities:** One data set (Axelrod & Henry, 1992).

7.6.1 Comparisons of standard deviations across age groups

For the *Wisconsin Card Sorting Test* (see Figure 7.29, based on data in Tables 7.11 and 7.12), consistently across both modalities of the test (categories achieved and perseverative errors) there was a somewhat erratic pattern in the variability across the earlier adult years for the first two data sets (data in Spreen & Strauss, 1991; Axelrod & Henry, 1992), although there was a consistent trend across all three data sets (data in Spreen & Strauss, 1991; Axelrod & Henry, 1992; Beatty, 1993) for variability to increase in later old age. The comparisons of standard deviations across all age groups revealed a highly significant effect ($p = < 0.01$) for both categories achieved and perseverative errors in each of the three data sets.

For the Spreen and Strauss data set, backwards comparisons revealed a significantly higher standard deviation for the 80+ age groups compared with the <40 age group for both categories achieved and perseverative errors ($p=0.0146^*$ and $p=0.0111^*$, respectively), and no significant differences for any of the other pairwise comparisons except that for

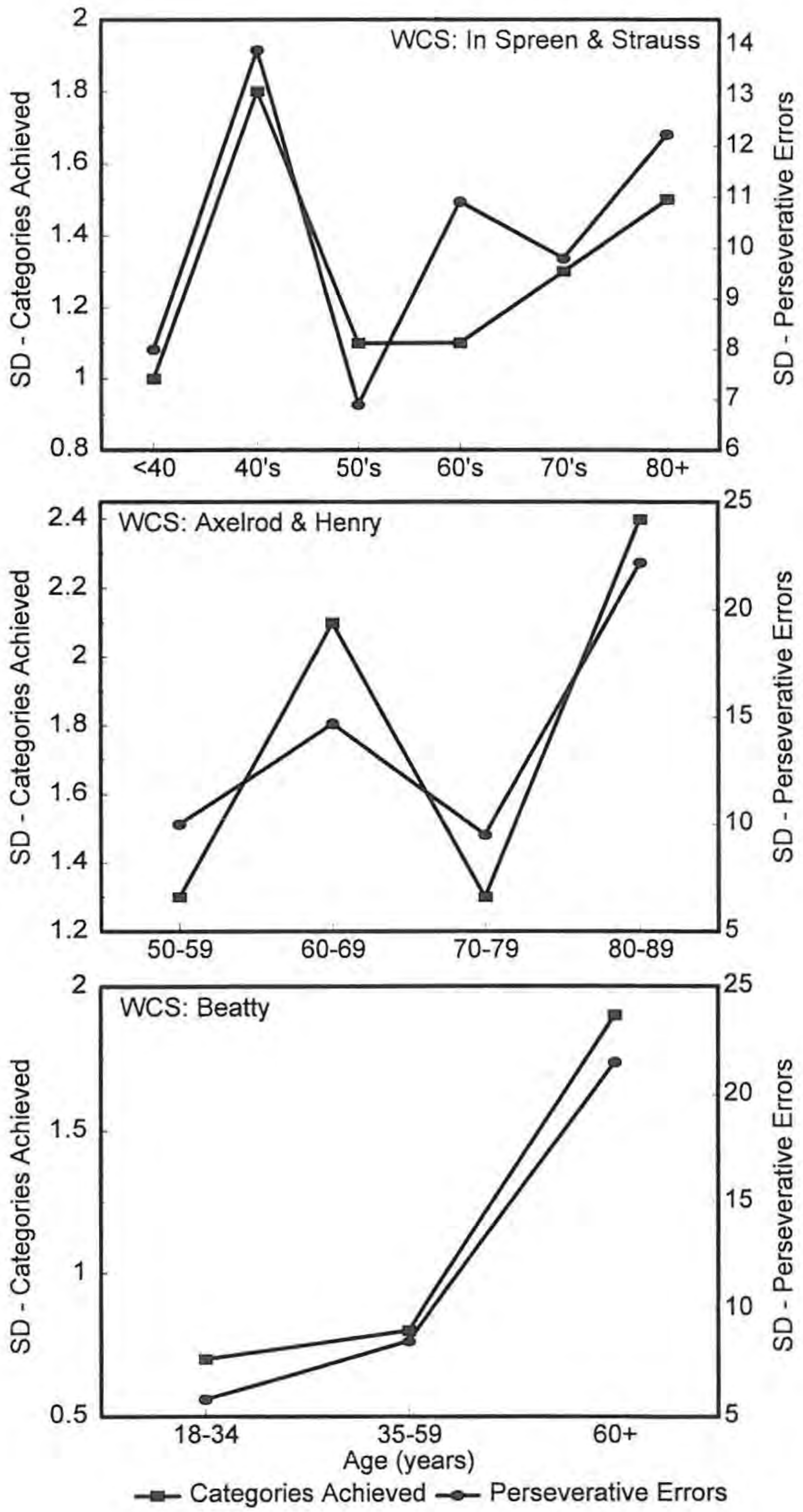


Figure 7.29. Wisconsin Card Sorting (WCS): Standard deviations across age groups for three data sets (Collated data in Spreen & Strauss, 1991; Axelrod & Henry, 1992; Beatty, 1993).

perseverative errors there was also a significantly higher standard deviation in the 80+ age group than in the 50's age group ($p=0.0196^{*}$). Overall this revealed a significant pattern of increased variability for both submodalities of the test which occurred by the 40's age group, although there was also the tendency for a dip at the 50's age group which reached significance for perseverative errors only. For the Axelrod and Henry data set backwards comparisons revealed a significantly higher standard deviation in the 80-89 year-old age group compared with the 70-79 and 50-59 year-old age groups for categories achieved ($p=0.0052^{**}$, $p=0.0052^{**}$) as well as for perseverative errors ($p=<0.001^{***}$, $p=<0.001^{***}$), and no significant difference between the 80-89 and 60-69 year-old age groups for either submodality of the test. These results supported a zig-zag pattern of increasing and decreasing variability with relatively high variability at the 60's age interval which was present also for the 80's age group, although there was a significant dip in standard deviations in the 70-79 year-old age group.

In contrast to the somewhat erratic pattern of standard deviations in the first two studies, for the Beatty data set there was a very clear and highly significant pattern of increased standard deviations in the oldest 60+ age group compared with the 35-59 and 18-34 year-old age groups for both submodalities of the test ($p<0.001^{***}$ in all instances), indicating a linear pattern of increased variability in association with age which was present in the years following 60. Considering this smooth variability pattern for the Beatty data on the Wisconsin test, the erratic patterns of variability for the first two studies need to be examined for the contribution of possible artifactual sampling effects. The data in Spreen and Strauss were collated from two separate pools of data - Heaton (1981) for the <60 years, and Strauss & Spreen, (unpublished study) for the 60+ years, and both subsets of data came from samples with subjects who were well-educated or who were of above average IQ (see Appendix Tables A and B, p.13, and p.20). Thus the initial zig-zag pattern between the <40, 40's and 50's age groups is not explicable purely in terms of artifactual effects of two broad data sets with grossly disparate levels of education/IQ, since the data from these three age groups came from the same 'well-educated' Heaton subset. It is possible that within the Heaton subset there were discrepancies in the educational levels of the groups since these were not specifically reported. However an erratic pattern was also evident for the Axelrod and Henry data set for which educational levels of each group were well-matched (15, 14, 15, 15 years of education per group). Thus it does not appear that sampling effects are contributing to the erratic patterns. Rather it appears that there is wide variability inherent in the Wisconsin

Card Sorting task itself, which is evident when variability figures are described across relatively narrow age ranges such as decades in the case of the first two studies (see Figure 7.*, Spreen & Strauss, and Axelrod & Henry, upper two graphs), but which is smoothed out when the data are described across very broad age ranges well in excess of a decade such as was the case for the last study (Figure 7, Beatty, bottom graph). It would appear that if variability figures for the first two data sets were calculated in this manner over wider age ranges, the erratic patterns in the earlier years would average out to reveal an overall steady increase in variability between the early to late adult years as was evident for the Beatty data set.

For the *Stroop Test* including all submodalities (see Figure 7.30, based on data in Tables 7.11 and 7.12), the comparisons of standard deviations across all groups revealed a highly significant effect for both data sets (data in Spreen & Strauss, 1991; Troyer et al., 1994)($p = < 0.01$ in all instances). For Stroop dots, backwards comparisons revealed no significant difference between the standard deviations of the oldest 80+ and 70-79 year-old age groups, but a significantly higher standard deviation between the 80+ age group and each of the other younger age groups ($p = < 0.001^{***}$, 0.0099^{***}, 0.001^{***}). For Stroop words, backwards comparisons revealed no significant difference in the standard deviation of the oldest 80+ age group and those of the 70-79, 60-69 and 50-59 year-old age groups, but a significantly higher standard deviation between the 80+ age group than in the 20-35 year-old age group. For Stroop colour, backwards comparisons revealed a significantly higher standard deviation between the oldest 80+ age group and each of the younger age groups ($p = < 0.001^{***}$ in all instances). For Stroop interference backwards comparisons revealed a significantly higher standard deviation for the 80-91 year-old age group and each of the younger age groups ($p = 0.0429^{*}$, $< 0.001^{***}$). Thus overall for the Stroop test across all submodalities and studies, there was a highly consistent linear pattern of progressively increasing variability which became apparent in the middle years for words and dots, was clearly apparent for all submodalities by the 70's, and was sustained at a high level by the 80's age stage. There was no indication of any significant dedifferentiation effect.

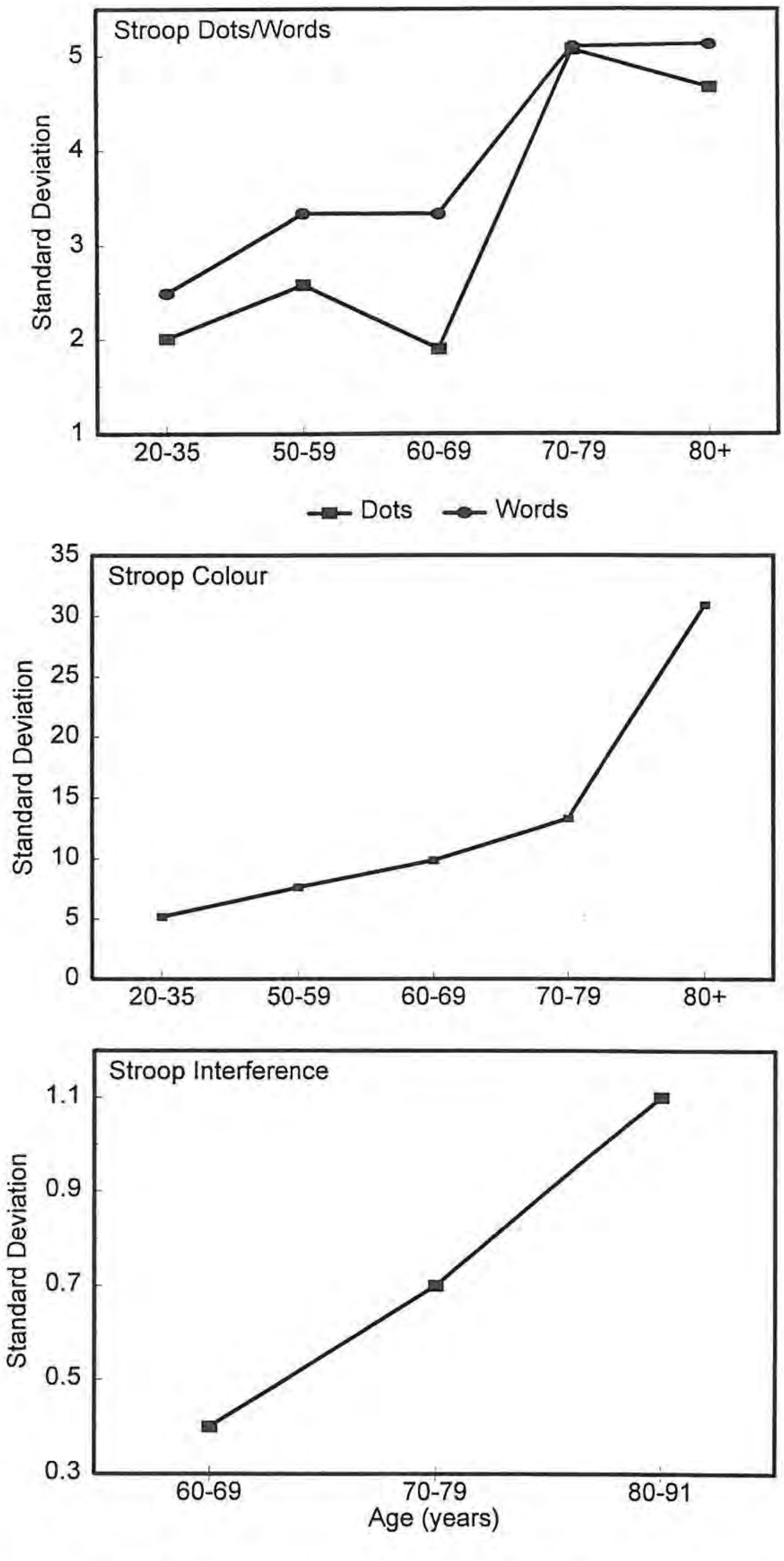


Figure 7.30. Stroop Test: Standard deviations across age groups for Dots, Words, Colour (collated data in Spreen & Strauss, 1991), and Interference (data from Troyer, Graves, & Cullum, 1994).

For *Similarities* (Axelrod & Henry) (see Figure 7.31 based on data contained in Tables 7.11 and 7.12), the variability pattern revealed was identical with that for the Stroop test. There was a highly significant effect for the comparisons of standard deviations across all groups ($p = < 0.01$), and backwards comparisons revealed a significantly higher standard deviation in the 80-89 age group compared with all the younger age groups ($p=0.0245^{*}$, 0.0035^{**}, $p = < 0.001^{***}$) which supported a linear pattern of steadily increasing variability from younger to older age groups, with a prominently significant effect in evidence by the 80's age decade.

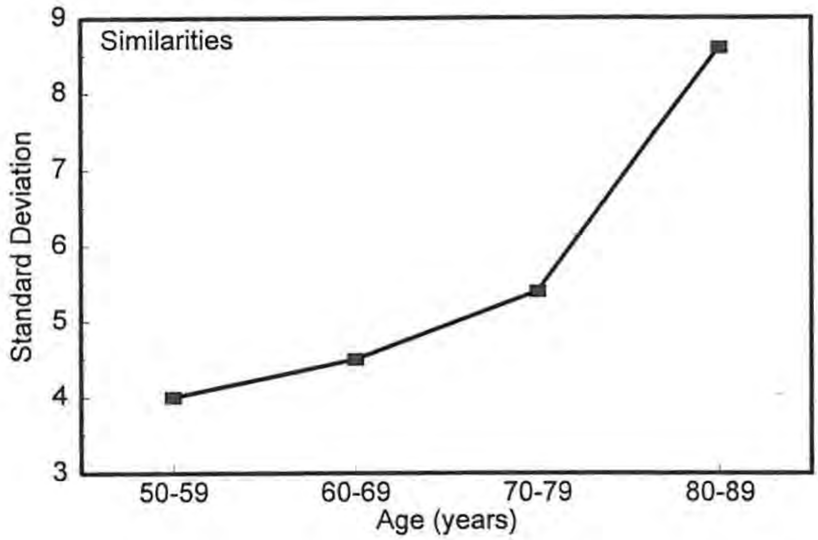


Figure 7.31. Similarities: Standard deviations across age groups (data from Axelrod & Henry, 1992).

7.6.2 Comparisons of means

For the modality of problem solving (see Table 7.11 for overall trends, and Table 7.12 for pairwise statistical comparisons) in all instances there was a clear trend of impaired ability in the oldest age groups compared with the youngest age groups as reflected in the mean scores which showed a decline for the Wisconsin Card Sorting Test (categories achieved) and Similarities, and an overall increase for the Wisconsin Card Sorting Test (perseverative errors) and all submodalities of the Stroop Test. For the Wisconsin Card Sorting Test (categories achieved) the mean scores of the oldest age groups were significantly lower than those of the youngest age groups for all data sets ($p = < 0.001^{***}$, $p=0.0027^{***}$, $p=0.0049^{***}$), and for the Wisconsin Card Sorting Test (perseverative errors) the mean scores of the oldest age groups were significantly higher than those of the youngest age groups for all data sets ($p= < 0.001^{***}$, $p= < 0.001^{***}$, $p=0.0027^{***}$, $p= < 0.001^{***}$, $p=0.0049^{***}$, $p=0.0175^{**}$). For the Stroop Test the timed score means of the oldest age groups were significantly higher than those of the youngest age groups across all the modalities presented ($p < 0.001^{***}$ in all instances). For Similarities the mean of the oldest age group was significantly lower than the mean of the youngest age group ($p=0.0032^{**}$). The latter effects are selectively illustrated in Figure 7.32, with steadily declining mean scores for the Beatty (1993) Wisconsin Card Sorting Test Categories and for Similarities (top and bottom graphs, respectively), and increasing mean time scores for the Stroop Colour test (middle graph), all in association with progressively increasing variability.

Although there was an ultimate decline in ability between the youngest and oldest age groups as reflected in the comparisons of mean scores for all the problem solving tests that were submitted to analysis, of note were the initial fluctuations in mean scores which occurred in association with erratic variability patterns for the earlier age groups of two of the Wisconsin Card Sorting data sets (Spreen & Strauss; Axelrod & Henry). These effects which support the postulate posed above of inherently wide variability in performance for the Wisconsin Card Sorting test, are selectively illustrated for the WCS Categories achieved (see Figure 7.33). However, as can be seen from Figure 7.33, in the later age groups the mean scores were no longer erratic and simultaneously the haphazard pattern of variability also disappeared. For both data sets in the later years a linear pattern of significantly increasing variability emerged. Thus across the entire spectrum of problem solving tasks, wherever the mean indicated a steady decline in ability in association with older age, this invariably occurred together with a linear pattern of significantly increasing variability.

7.6.3 Problem solving: Summary of effects

Generally for problem solving, as tapped by the Wisconsin Card Sorting, Stroop and Similarities tests, there was support for a sharp significant increase in variability in the later adult years compared with the younger adult years in association with a significantly diminished level of performance between the oldest and youngest age stages as reflected by the mean scores (that is the initial differentiation stage of the overall shuttle bulge effect).

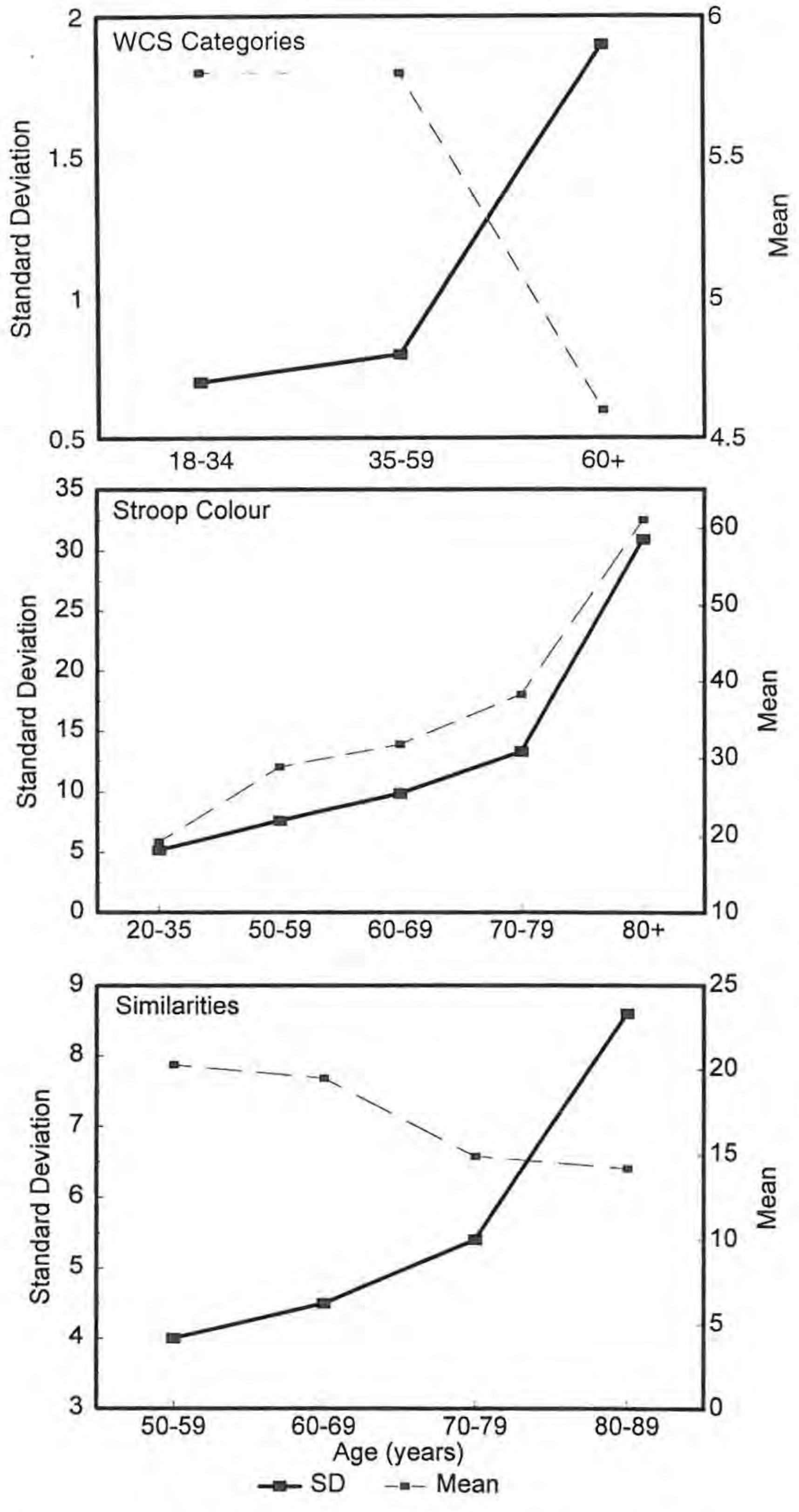


Figure 7.32. Wisconsin Card Sorting (WCS), Stroop Colour and Similarities: Standard deviations and mean scores across age groups (data from Beatty, 1993; collated data in Spreen & Strauss, 1991; and data from Troyer, Graves, and Cullum, 1994, respectively).

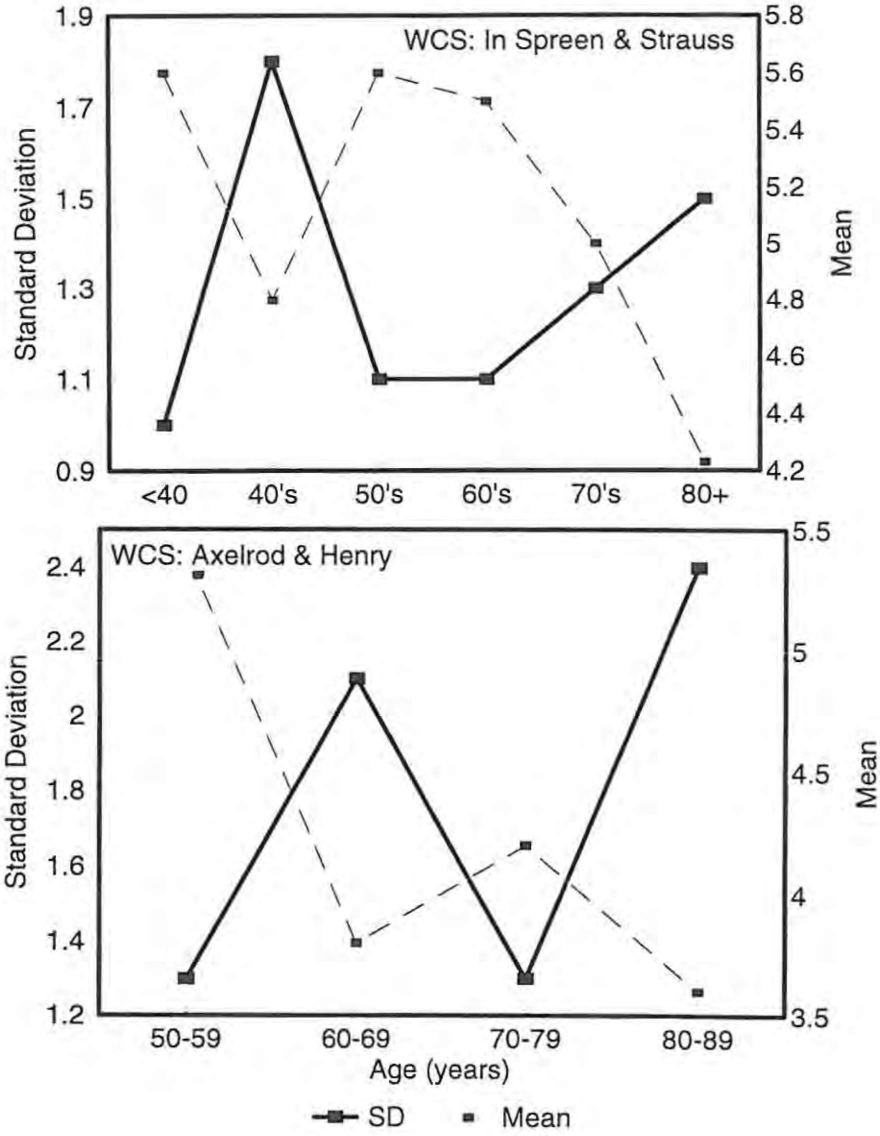


Figure 7.33. Wisconsin Card Sorting (WCS) Categories achieved: Standard deviations and mean scores across age groups for two data sets (Collated data in Spreen & Strauss, 1991; Axelrod & Henry, 1992).

In all instances the significant increase in variability was clearly apparent by the 80's age stage, and in no case was there any indication of a subsequent decrease in variability. In some instances the progressive increase in variability in association with age was apparent by the 70's and even the 60's age stages. Erratic patterns in variability in the earlier age stages for the Wisconsin Card Sorting test occurred in association with equally erratic patterns in the mean scores. Hence this did not stand in contradiction to the basic points of the shuttle model in which it is proposed that the variability bulge occurs in association with the presence of decremental aging effects. Since samples were reportedly relatively well-matched for educational level it appears that the Wisconsin Card Sorting test is subject to wide fluctuations in variability independent of age effects such that, only when variability was examined across very wide age ranges as in the Beatty 1993 study, did the overriding steady increase in variability due to age become apparent.

Table 7.11. PROBLEM SOLVING: Mean Scores and Comparisons of Standard Deviations across all Age Groups.

Wisconsin Card Sorting (Collated data in Spreen & Strauss 1991)

AGE GROUP	<40	40's	50's	60's	70's	80+	df	B[1]	p
n	100	19	16	28	19	13	-	-	-
WCS-Categories Achieved									
Mean	5.6	4.8	5.6	5.5	5.0	4.23	-	-	-
SD	1.0	1.8	1.1	1.1	1.3	1.5	5	15.62	<0.01**
WCS-Perseverative Errors									
Mean	10.4	16.0	11.3	12.25	15.9	25.77	-	-	-
SD	8.0	13.9	6.9	10.91	9.8	12.23	5	17.21	<0.01**

Wisconsin Card Sorting (Axelrod and Henry 1992)

AGE GROUP	50-59	60-69	70-79	80-89	df	B	p
n	20	20	20	20	-	-	-
WCS-Categories Achieved							
Mean	5.4	3.8	4.2	3.6	-	-	-
SD	1.3	2.1	1.3	2.4	3	11.08	<0.01**
WCS-Perseverative Errors							
Mean	14.0	22.4	25.5	33.0	-	-	-
SD	10.0	14.7	9.5	22.2	3	18.17	<0.01**

Wisconsin Card Sorting (Beatty 1993)

AGE GROUP	18-34	35-59	60+	df	B	p
n	21	21	21	-	-	-
WCS-Categories Achieved						
Mean	5.8	5.8	4.6	-	-	-
SD	0.7	0.8	1.9	2	24.42	<0.01**
WCS-Perseverative Errors						
Mean	9.9	9.9	20.5	-	-	-
SD	5.8	8.5	21.5	2	35.20	<0.01**

[1] Bartlett's test statistic

Table 7.11 *continued*

Table 7.11. PROBLEM SOLVING: Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

Stroop Test (Collated data in Spreen & Strauss 1991)								
AGE GROUP	20-35	50-59	60-69	70-79	80+	df	B	p
n	40	19	28	24	15	-	-	-
Stroop-Dots								
Mean	10.10	13.74	12.71	15.00	18.87	-	-	-
SD	2.01	2.58	1.90	5.07	4.67	4	42.74	<0.01**
Stroop-Words								
Mean	12.00	16.58	16.32	19.04	24.13	-	-	-
SD	2.49	3.34	3.33	5.10	5.13	4	19.74	<0.01**
Stroop-Colour								
Mean	19.25	28.90	31.82	38.33	61.13	-	-	-
SD	5.18	7.62	9.86	13.29	30.94	4	88.81	<0.01**

Stroop Test (Troyer et al. 1994)						
AGE GROUP	60-69	70-79	80-91	df	B	p
n	19	17	15	-	-	-
Stroop-Interference						
Mean	2.2	2.6	3.2	-	-	-
SD	0.4	0.7	1.1	2	14.86	<0.01**

Similarities (Axelrod & Henry 1992)							
AGE GROUP	50-59	60-69	70-79	80-89	df	B	p
n	20	20	20	20	-	-	-
Mean	20.3	19.5	14.9	14.2	-	-	-
SD	4.0	4.5	5.4	8.6	3	13.88	<0.01**

Significant difference (**p<0.01).

Table 7.12. PROBLEM SOLVING: Pairwise Comparisons of Means between the oldest and youngest Age Group, and of Standard Deviations between the oldest Age Group and each younger Age Group unless otherwise specified.

Wisconsin Card Sorting (Collated data in Spreen & Strauss 1991)									
AGE GROUP	<40	40's	50's	60's	70's	80+	df	t/F	p
n	100	19	16	28	19	13	-	-	-
WCS-Categories Achieved									
Mean	5.6	-	-	-	-	4.23	111	4.36	<0.001***
SD	-	-	-	-	1.3	1.5	18,12	0.75	0.2831
SD	-	-	-	1.1	-	1.5	27,12	0.54	0.0882
SD	-	-	1.1	-	-	1.5	15,12	0.54	0.1280
SD	-	1.8	-	-	-	1.5	18,12	1.44	0.2627
SD	1.0	-	-	-	-	1.5	99,12	0.44	0.0146*
WCS-Perseverative Errors									
Mean	10.4	-	-	-	-	25.77	111	-6.09	<0.001***
SD	-	-	-	-	9.8	12.23	18,12	0.64	0.1919
SD	-	-	-	10.91	-	12.23	27,12	0.79	0.2984
SD	-	-	6.9	-	-	12.23	15,12	0.32	0.0196*
SD	-	13.9	-	-	-	12.23	18,12	1.29	0.6691
SD	8.0	-	-	-	-	12.23	99,12	0.43	0.0111*

Wisconsin Card Sorting (Axelrod and Henry 1992)							
AGE GROUP	50-59	60-69	70-79	80-89	df	t/F	p
n	20	20	20	20	-	-	-
WCS-Categories Achieved							
Mean	5.4	-	-	3.6	38	2.95	0.0027***
SD	-	-	1.3	2.4	19,19	0.29	0.0052**
SD	-	2.1	-	2.4	19,19	0.77	0.2831
SD	1.3	-	-	2.4	19,19	0.29	0.0052**
WCS-Perseverative Errors							
Mean	14.0	-	-	33.0	38	-3.49	<0.001***
SD	-	-	9.5	22.2	19,19	0.18	<0.001***
SD	-	14.7	-	22.2	19,19	0.44	0.0401
SD	10.0	-	-	22.2	19,19	0.18	<0.001***

Table 7.12 *continued*

Table 7.12. PROBLEM SOLVING: Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

Wisconsin Card Sorting (Beatty 1993)						
AGE GROUP	18-34	35-59	60+	df	t/F	p
n	21	21	21	-	-	-
WCS-Categories Achieved						
Mean	5.8	-	4.6	40	2.67	0.0049***
SD	-	0.8	1.9	20,2	0.18	<0.001***
SD	0.7	-	1.9	20,2	0.14	<0.001***
WCS-Perserverative Errors						
Mean	9.9	-	20.5	40	-2.18	0.0175**
SD	-	8.5	21.5	20,2	0.16	<0.001***
SD	5.8	-	21.5	20,2	0.07	<0.001***

Stroop Test (Collated Data in Spreen & Strauss 1991)								
AGE GROUP	20-35	50-59	60-69	70-79	80+	df	t/F	p
n	40	19	28	24	15	-	-	-
Stroop-Dots								
Mean	10.10	-	-	-	18.87	53	-9.80	<0.001***
SD	-	-	-	5.07	4.67	23,14	1.18	0.3834
SD	-	-	1.90	-	4.67	27,14	0.17	<0.001***
SD	-	2.58	-	-	4.67	18,14	0.31	0.0099**
SD	2.01	-	-	-	4.67	39,14	0.19	<0.001***
Stroop-Words								
Mean	12.00	-	-	-	24.13	53	-11.81	<0.001***
SD	-	-	-	5.10	5.13	23,14	0.99	0.4747
SD	-	-	3.33	-	5.13	27,14	0.42	0.0262
SD	-	3.34	-	-	5.13	18,14	0.42	0.0443
SD	2.49	-	-	-	5.13	39,14	0.24	<0.001***
Stroop-Colour								
Mean	19.25	-	-	-	61.13	53	-8.38	<0.001***
SD	-	-	-	13.29	30.94	23,14	0.19	<0.001***
SD	-	-	9.86	-	30.94	27,14	0.10	<0.001***
SD	-	7.62	-	-	30.94	18,14	0.06	<0.001***
SD	5.18	-	-	-	30.94	39,14	0.03	<0.001***

Table 7.12 *continued*

Table 7.12. PROBLEM SOLVING: Pairwise Comparisons of Means and of Standard Deviations between Age Groups *continued.*

Stroop Test (Troyer et al. 1994) AGE GROUP	60-69	70-79	80-91	df	t/F	p
n	19	17	15	-	-	-
Stroop- Interference						
Mean	2.2	-	3.2	32	-3.68	<0.001***
SD	-	0.7	1.1	16,14	0.41	0.0429*
SD	0.4	-	1.1	18,14	0.13	<0.001***

Similarities (Axelrod & Henry 1992) AGE GROUP	50-59	60-69	70-79	80-89	df	t/F	p
n	20	20	20	20	-	-	-
Mean	20.3	-	-	14.2	38	2.88	0.0032***
SD	-	-	5.4	8.6	19,19	0.39	0.0245*
SD	-	4.5	-	8.6	19,19	0.27	0.0035**
SD	4.0	-	-	8.6	19,19	0.22	<0.001***

Significant difference (*p<0.10; **p<0.05; ***p<0.01) as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

7.7 MOTOR ACTIVITY

For the modality of motor activity the data submitted to analysis appear in Table A (Appendix pp.16-17), and the educational/IQ level of the samples appears in Table B (Appendix p.21). The following test and accompanying normative data set was included in the analysis:

> *(i) Finger Tapping Preferred and Non-Preferred Hands:* One data set (Bornstein, 1985, in Spreen & Strauss, 1991).

7.7.1 Comparisons of standard deviations across age groups

For the Finger Tapping data set (Bornstein, 1985) the comparisons of standard deviations across all age groups revealed *no significant differences for either males or females in any of the grade <12 or grade >12 educational conditions.* For males on Finger Tapping Preferred (grade >12) (see Tables 7.13 and 7.14) the all groups analyses revealed an effect which approached significance (p=0.0807). The follow-up backwards comparisons for this subset of data revealed that there was a significantly lower standard deviation in the oldest 60-69 year-old age groups relative to the 40-59 and 20-39 year-old age groups (0.0125** and 0.0430*) which indicated that there was a decline in variability for the 60-69 year-old age group relative to the two younger age-groups without any indication of a prior increase. However this was an isolated trend which was unsupported by regular patterns identified in the broader analysis across all the other functional modalities included in the analysis, and nor was it commensurate with any other trends in the entire Bornstein Finger Tapping data set. For the rest of the subsets for males on both Finger Tapping Preferred and Non-Preferred Hands test there was fairly consistent trend in both educational conditions (grade <12 years and grade >12 years) for *lower* standard deviations in the middle 40-59 year-old age group relative to the younger 20-39 year-old age group and the older 60-69 year-old age group, the effect of which did not approach significance in any of the subsets of data. For females across all the subsets, when compared with males, there was an entirely contrasting trend for a peak in standard deviations in the middle 40-59 year-old age group which in no case even remotely approached significance.

Overall these results for the Finger Tapping Test tend to support the absence of significant alterations in variability by the 60's age decade. Since no data sets beyond the 60's age decade were analyzed it is not possible to know whether changes might have emerged were older age groups submitted for analysis, or whether there would have continued to be no differences detected across groups.

7.7.2 Comparisons of means

For the modality of hand motor activity (see Table 7.13 for overall trends, and Table 7.14 for pairwise statistical comparison) in all instances there was a trend of consistently declining mean scores. For males in the grade >12 years educational condition which was subjected to statistical analysis, there was a significantly lower mean score for the 60's age group compared with that of the 20's age group ($<0.001^{***}$). This occurred in association with a strong trend of a reduction in variability in the 60's age group which, however, as discussed above, only approached significance on the overall age groups analysis and was not supported by any other trends.

7.7.3 Motor activity: Summary of effects

Generally across the subsets of data for motor activity on the Finger Tapping Test for males and females in two educational conditions of grade <12 and grade >12 years, there were no consistent trends or significant changes in variability in evidence by the 60's age decade. The general absence of any significant alterations in variability by the 60's age decade occurred in association with the consistent trend of declining mean performance between older and younger age groups. However the data submitted for analysis were very limited in that they were derived from only one study which tapped age groups only as far as the 60's age stage. Thus the absence of any notable variability effects did not negate the possibility that shuttle bulge and shuttle shift variability effects might occur in this functional domain, were data with older age groups submitted to analysis.

Table 7.13. MOTOR ACTIVITY: Mean Scores and Comparisons of Standard Deviations across all Age Groups.

Finger Tapping - Males (Bornstein 1985, in Spreen & Strauss 1991) AGE GROUP	20-39	40-59	60-69	df	B[1]	p
Finger Tapping Pref <12 years ed						
n	21	13	16	-	-	-
Mean	49.7	42.3	39.1	-	-	-
SD	6.0	5.2	5.7	2	0.29	0.8648
Finger Tapping Pref >12 years ed						
n	86	17	23	-	-	-
Mean	48.4	43.3	43.0	-	-	-
SD	6.5	7.9	4.7	2	5.03	0.0807^{+}
Finger Tapping Non-Pref <12 years ed						
n	21	13	16	-	-	-
Mean	47.0	39.8	35.2	-	-	-
SD	5.5	3.6	5.2	2	2.47	0.2915
Finger Tapping Non-Pref >12 years ed						
n	86	17	23	-	-	-
Mean	44.8	39.5	39.3	-	-	-
SD	6.4	5.8	6.2	2	0.26	0.8801

[1]Bartlett's test statistic.

Table 7.13 *continued*

Table 7.13. **MOTOR ACTIVITY:** Mean Scores and Comparisons of Standard Deviations across all Age Groups *continued.*

Finger Tapping - Females (Bornstein 1985, in Spreen & Strauss 1991) AGE GROUP	20-39	40-59	60-69	df	B	p
Finger Tapping Pref < 12 years ed						
n	13	22	22	-	-	-
Mean	45.2	36.3	29.7	-	-	-
SD	6.0	7.8	6.2	2	1.51	0.4700
Finger Tapping Pref > 12 years ed						
n	49	43	34	-	-	-
Mean	44.3	40.5	32.2	-	-	-
SD	5.8	7.1	6.0	2	2.05	0.3589
Finger Tapping Non-Pref < 12 years ed						
n	13	22	22	-	-	-
Mean	40.7	35.2	29.8	-	-	-
SD	5.0	5.8	5.6	2	0.32	0.8509
Finger Tapping Non-Pref > 12 years ed						
n	49	43	34	-	-	-
Mean	40.6	37.8	32.0	-	-	-
SD	5.6	6.0	4.9	2	1.47	0.4791

Approaching significance ($^{+}p<0.10$).

Table 7.14. MOTOR ACTIVITY: Comparisons of Means between the oldest and youngest Age Groups, and of Standard Deviations between the oldest Age Group and each younger age group unless otherwise specified.

Finger Tapping - Males (Bornstein 1985, in Spreen & Strauss 1991)						
AGE GROUP	20-39	40-59	60-69	df	t/F	p
Finger Tapping Pref > 12 years ed						
n	86	17	23	-	-	-
Mean	48.5	-	43.0	107	3.80	<0.001***
SD	-	7.9	4.7	16,22	2.83	0.0125**
SD	6.5	-	4.7	85,22	1.91	0.0430*
Additional Comparison						
SD	6.5	7.9	-	85,16	0.68	0.1276

Significant difference (*p<0.10; **p<0.05; ***p<0.01) as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

7.8 OVERALL SUMMARY OF EFFECTS

Considering the results over the seven functional modalities of attention, verbal and visual memory, language and visual skills, problem solving and motor activity, the following general indications have emerged:

(i) For all modalities where there were clear indications of consistently declining average mean performances in association with age, there were almost invariably also significant increases in variability most often by the 60's age stage. For speeded attention and visual memory tasks the increase was always by the 60's age decade; for verbal memory, language and problem solving the increase sometimes occurred only by the 70's or 80's age stages. Where the speed factor was escalated for both an attention task and a visual memory task the increase in variability shifted from the later post 60's age onset, and occurred as early as the 40's age stage (that is the increased challenge in the speed factor produced a left shuttle effect). For a single instance of a language task (word naming) the initial increase occurred as early as the 50's age stage, with a second increase by the 70's age stage.

(ii) Following the significant increase in variability in association with steadily declining mean scores and older age, there was an additional subsequent significant decline in variability in later old age which was highly task and domain specific. Almost invariably for attention, verbal and visual short-term memory there was a significant dedifferentiation effect by the 70' or 80's age stages, and where the speed factor was escalated for attention and visual short-term memory, the dedifferentiation effect occurred as early as the 50's age decade (that is the increased challenge by virtue of the speed factor produced a left shuttle effect). On the other hand, for language and visuo-perceptual tracking skills and problem solving there were almost invariably no significant dedifferentiation effects, even though with the exception of Digit Symbol Modalities, all of these skills were tapped up to as late as the 80's age stage.

(iii) Where there were shifts in the variability effects for different submodalities within tasks or across comparable tasks, these occurred with highly consistent indications of an *earlier* onset of the variability bulge (that is a *left shuttle effect*) for:

higher level of task challenge (where the speed factor in the attention and visual memory modalities was enhanced); for *male gender* (as noted for the RAVLT-Trial 1 verbal short-term memory task, the RAVLT verbal recognition task, and the BVRT-R visual memory task); for *greater age-sensitivity of task* (as noted above with the escalation of the speed factor in tasks, and also for submodalities within the RAVLT and BSRT verbal memory tasks); and for *lower levels of education* (as noted for the BVRT-R visual memory task and the FAS Word Association test).

(iv) In a number of instances there were highly significant erratic variability effects. These were for a recognition task and pure visuo-perceptual tasks with a low level or absent speed factor. Consistently in all these instances there were also no clear-cut steady declines in average performances as reflected in markedly fluctuating average task performance, indicating that erratic variability effects occurred regularly in association with relatively non-age-sensitive tasks.

(v) There were very few instances for which there were no significant variability effects in association with steadily declining mean scores, or of trends that were not highly commensurate across comparable submodalities. There was a single data set for the Wechsler Memory Scale Logical Memory (Abikoff et al., 1987), which stood in contrast to two other data sets for Logical Memory for which there *were* commensurate non-linear inverted-U variability effects, implying that the latter were the more robust findings. The second occasion where no clear-cut variability effects were in evidence in association with declining average scores, was for Animal Naming, which was only tapped from the 50's age stage onwards. Hence, considering the indication from another word finding language task that variability may increase as early as the 50's age stage in this modality, the absence of evidence of an increase in variability for this particular Animal Naming study does not rule out the possibility that variability increases earlier for an animal naming task. Alternatively, the naming of such familiar items may be a task which is unusually free from variability effects regardless of age stage. Finally, in association with declining means, generally there was an absence of variability effects for hand motor activity tapped only up to the 60's age decade. However, were older age groups tapped, it is conceivable that variability effects might still be revealed for this task.

(vi) There was one instance of separate data sets for the same task, which, when tapped across the same overall age span, revealed apparently contrasting significant variability effects: the Stuss (1987) PASAT data set revealed a dedifferentiation effect, which was not apparent in the Stuss (1988) PASAT data set as reported in Spreen and Strauss (1991). As discussed earlier (p.197, third paragraph, continuing p.198), the Stuss (1988) data comprise much larger age groupings than were used in the Stuss (1987) report, and it is conceivable that this may have obscured the appearance of the dedifferentiation effect in the latter data set.

Thus in sum, from this broad meta-analysis of cognitive aging studies, it appears that the occurrence of a significant increase in variability in association with declining average cognitive ability and the progression of older age, with or without a subsequent significant decrease in variability depending on the task and functional domain (that is the shuttle bulge effect), is an event which is equally as robust as the well-documented finding of significant linear decline of average cognitive ability due to aging. Generally, from the graphical material presented, it was evident that the significant onset of the variability bulge (the initial differentiation effect) presented coincidentally with the beginning of marked decline in average performance. On occasion, however, the commencement of increased variability preceded or followed the onset of diminished average performance by one to two decades. Finally, there was support for the shuttle shift effect of an earlier onset of variability in cognitive test performance (that is a left shuttle effect) for lower levels of education, higher levels of task challenge and male gender, and conversely there was a later onset of variability (that is a right shuttle effect) for higher levels of education, lower levels of task challenge and female gender.

It was considered that the few instances of apparently contradictory indications across data sets for the same tests or modalities were *not* sufficient in themselves to support a fundamental *absence* of coherent variability patterns. This was in that either this occurred as an exception against the background of a more robust series of expected variability findings, or it was possible to identify plausible explanations for such discrepancies. These included the range of the particular age ranges tapped, very different clusterings of age stages, and obviously discrepant educational levels between groups.

8

CHAPTER 8: RESULTS PHASE 2

The results for the limited number of data sets included in Phase 2 of the study are presented together as a group in subsections dedicated to the three broad categories within which the statistical analyses were conducted: *Analyses of subject characteristics, Between-groups analyses,* and *Within-groups analyses*. The results for all the tests are summarized in tables within each category of analysis, and are presented together (followed by histograms where relevant) at the end of each subsection. Figures are presented at the next available opportunity after they are first mentioned in the text. Thus page references will *not* be given for tables and figures which, when alluded to, follow on obviously within each subsection in the chapter. However, as for the rest of the thesis, page numbers *will* be given when cross-reference is made to tables (or figures) which are contained outside of the chapter, and/or are not easily accessible within the chapter.

As in the Phase 1 results section (due to the large amount of data submitted to analysis) except where particular emphasis is required, only significant p-values are reiterated in the text. For purposes of cross-referencing these values are duplicated in the text as they appear in the tables: For the all-groups comparisons, the p-values require no adjustment in terms of the levels of significance which they denote; for the multiple pairwise comparisons, the levels of significance as determined by Bonferroni's adjustment for simultaneous testing are indicated in the text (as in the results tables) by means of asterisks that denote the following levels of significance: $^{*}p<0.10$; $^{**}p<0.05$; and $^{***}p<0.01$ (see Table E, Appendix, p.25). As noted in the method (section 6.2.3), in accordance with Miller (1981), the overall level of significance for the multiple comparisons was set consistently at $p=0.10$ for both Phase 1 and Phase 2 of the study in order to guard against the probability of a Type II error (that is failure to identify the nature of significant differences detected on the analyses across all groups) in view of the large numbers of comparison groups for most of the data sets.

The data sets which were submitted to analysis for Phase 2 of the study are collated in Table C (Appendix pp.22-23), and a synopsis of the educational and/or IQ levels of the age groups appear in Table D (Appendix p.24). In summary the following tests (and accompanying sources of normative data sets) were included in the analysis:

(i) Trail Making Test Parts A and B: Data set for the 60-69, 70-79 and 80-89 year-old age groups from Cornfield and Shuttleworth-Jordan (1996), for which both educational and NART IQ levels per age group were available;

(ii) Digit Symbol and Digit Symbol Incidental Recall: Data set for the 20-39, 40-59, 60-69, 70-79 and 80-89 year-old age groups from Bode and Shuttleworth-Jordan (1993) and Shuttleworth-Jordan and Bode (1995a), for which educational levels per age group were available.

(iii) Finger Tapping Test Preferred and Non-Preferred Hands: Data set for the 20-39, 40-59, 60-69, 70-79 and 80-89 year-old age groups from Shuttleworth-Jordan and Bode (1996), for which educational levels per age group were available.

8.1 ANALYSES OF SUBJECT CHARACTERISTICS

The analyses of subject characteristics for each of the three data sets appear in Tables 8.1, 8.2 and 8.3. As expected there were highly significant differences in the mean scores between age groups for age ($p = < 0.01$, all tests). Further (as is desirable for a study investigating test variability), for the Trail Making Test (see Table 8.1) there was no significant difference in the standard deviations between age groups for age indicating that the spread and range of the age of subjects was equivalent between groups. However, for the Digit Symbol, Incidental Recall and Finger Tapping tests (see Tables 8.2 and 8.3), there were significant differences in the standard deviations between all five age groups for age ($p = < 0.01$ in each instance) due to the wider age range in the younger two age groups compared with that in the older three age groups (see distribution diagram of age grouped by age for the basic five-group sample, Histogram 8.1, p.326). When the younger two age groups were excluded from the statistical analyses for these three tests (see Tables 8.2 and 8.3), there were no significant differences in the standard deviations between the three oldest age groups for age.

In the case of all test samples (See Tables 8.1, 8.2, and 8.3), there were no significant differences in the mean scores or standard deviations between age groups for years of education. For the Trail Making Test, in addition, there were no significant differences in the mean scores or standard deviations between age groups for NART IQ.

Table 8.1. Trail Making Test: Comparisons of Subject Characteristics (Range, Mean and Standard Deviation) for Age, Years of Education and NART IQ across three Age Groups.

AGE GROUP	60-69	70-79	80-89	df	F	p
n	33	38	34	-	-	-
Female/Male	26/7	24/14	20/14	-	-	-
Age						
Mean	64.76	74.32	83.35	2,102	399.09	<0.01**
SD	2.80	2.83	2.42	2,102	0.51	0.6025
Range	60-69	70-79	80-89	-	-	-
Education						
Mean	14.61	15.32	14.97	2,102	0.56	0.5715
SD	3.04	2.80	2.59	2,102	0.22	0.8003
Range	10-27	10-22	10-21	-	-	-
NART IQ						
Mean	128.36	128.82	128.15	2,102	0.05	0.9481
SD	9.04	7.09	10.45	2,102	1.39	0.2541
Range	111-144	116-142	96-144	-	-	-

Note: For Total Sample 60-89 years: N=105 (Female 70/Male 35); Educational Mean = 14.98 (SD=2.80); NART IQ Mean=128.46 (SD=8.82). Significant difference (**$p<0.01$).

Table 8.2. Digit Symbol and Incidental Recall: Comparisons of Subject Characteristics (Range, Mean and Standard Deviation) for Age and Years of Education across five Age Groups, and for Age across the three older Age Groups.

AGE GROUP	20-39	40-59	60-69	70-79	80-89	df	F	p
n	23	21	28	31	28	-	-	-
Female/Male	15/8	15/6	24/4	21/10	18/10	-	-	-
Age								
Five Group Comparison								
Mean	29.39	47.76	64.36	74.87	83.32	4,126	676.10	<0.01**
SD	6.45	5.76	2.84	2.81	2.94	4,126	16.15	<0.01**
Three Group Comparison								
Mean	-	-	64.36	74.87	83.32	2,84	355.11	<0.01**
SD	-	-	2.84	2.81	2.94	2,84	1.12	0.3307
Range	21-39	40-59	60-69	70-79	80-88	-	-	-
Education								
Mean	15.91	15.10	14.07	14.97	14.82	4,126	1.71	0.1516
SD	1.91	2.79	2.26	2.82	2.68	4,126	1.92	0.1117
Range	2-19	11-22	10-18	10-22	10-21	-	-	-

Note. For total sample 20-89 years: N=131 (Female 93/Male 38); Educational Mean=14.93 (SD=2.49)
Significant difference (**p <0.01).

Table 8.3. Finger Tapping Preferred and Non-Preferred Hands: Comparisons of Subject Characteristics (Range, Mean and Standard Deviation) for Age and Years of Education across five Age Groups, and for Age across the three older Age Groups.

AGE GROUP	20-39	40-59	60-69	70-79	80-89	df	F	p
n	23	21	27	26	26	-	-	-
Female/Male	15/8	15/6	23/4	19/17	16/10	-	-	-
Age								
Five Group Comparison								
Mean	29.39	47.76	64.33	74.73	83.27	4,118	616.47	<0.01**
SD	6.45	5.76	2.90	2.91	2.31	4,118	14.57	<0.01**
Three Group Comparison								
Mean			64.33	74.73	83.27	2,76	322.45	<0.01**
SD			2.90	2.91	2.31	2,76	1.10	0.3375
Range	21-39	40-59	60-69	70-79	80-88	-	-	-
Education								
Mean	15.91	15.10	14.22	14.73	14.73	4,118	1.55	0.1909
SD	1.91	2.79	2.15	2.68	2.77	4,118	2.29	0.0640
Range	12-19	11-22	10-18	10-18	10-21	-	-	-

Note: For total sample 20-89 years: N=123 (Female 88/Male 35); Educational Mean=14.93 (SD=2.46).
Significant difference (** p<0.01).

Histogram 8.1. Digit Symbol and Incidental Recall: Distribution of Age by Age.

```
         TWENTHIR            FORFIFT             SIXTIES             SEVENTIE            EIGHTIES
MIDPOINTS...................+...................+...................+...................+...................+
  91.000)
  87.500)                                                                                ******
  84.000)                                                                                M*********
  80.500)                                                            **                  ************
  77.000)                                                            ************
  73.500)                                                            M************
  70.000)                                        ***                 ****
  66.500)                                        **********
  63.000)                                        M********
  59.500)                    *                   ******
  56.000)                    **
  52.500)                    ****
  49.000)                    M***
  45.500)                    ****
  42.000)                    ****
  38.500)******              **
  35.000)*
  31.500)***
  28.000)M**
  24.500)******
  21.000)****
         LEGEND FOR GROUP MEAN'S:    M - MEAN COINCIDES WITH AN ASTERISK
                                     N - MEAN DOES NOT COINCIDE WITH ANY ASTERISK
```

8.2 BETWEEN-GROUPS ANALYSES

8.2.1 Comparisons of standard deviations and means

For the *Trail Making Test* parts A and B,[16] the comparisons of standard deviations across all groups (see Figure 8.1 based on data contained in Tables 8.4 and 8.5), revealed significant effects for both TMT A and TMT B (0.0403, <0.01, respectively). For TMT A, backwards comparisons revealed a significantly higher standard deviation between the oldest 80-89 year-old age group and the 70-79 year-old age group (0.0407*), and although there was no significant difference between the 80-89 and the 60-69 year-old age groups, from the graphical illustration and the p-value ($p=0.0625$) there was obviously a very strong trend in this direction. Thus overall for TMT A there was a linear pattern of increased variability by the 80's age group. For TMT B, backwards comparisons (see Table 8.4) revealed no significant difference between the standard deviations of the 80-89 year-old and the 70-79 year-old age groups, but a significantly higher standard deviation between the 80-89 year-old age group and the 60-69 year-old age group ($p=0.0021^{**}$), indicating a linear pattern of increased variability by the 70's age group which was sustained in the 80's age group. A left versus right shuttle effect for level of task challenge was clearly evident in that the initial increase in variability occurred in the earlier 70's decade for TMT B (the task of greater challenge, left shuttle effect) when compared with the initial increase in variability which occurred only in the 80's age decade for TMT A (the task of lower challenge, right shuttle effect).

For the Trail Making test the comparisons of means across all groups (see Figure 8.1 based on data contained in Tables 8.4 and 8.5) revealed a highly significant effect for both TMT A and TMT B ($p<0.01$). Backwards comparisons for both TMT A and TMT B revealed that the mean scores of the oldest 80-89 year-old age groups were significantly higher than those of all the younger age groups ($p<0.001^{***}$, in all instances). Thus as is evident from

[16] The between-groups variability has been previously investigated for the same Trail Making Test and Digit Symbol raw score data (Cornfield & Shuttleworth-Jordan, 1996; Bode and Shuttleworth-Jordan, 1993, respectively), and the results of these investigations were discussed in Chapter 4. The analyses have been repeated here for elaboration in terms of the newly postulated variability model, and to serve as the basis for further within-groups analyses in the present investigation which were not previously conducted.

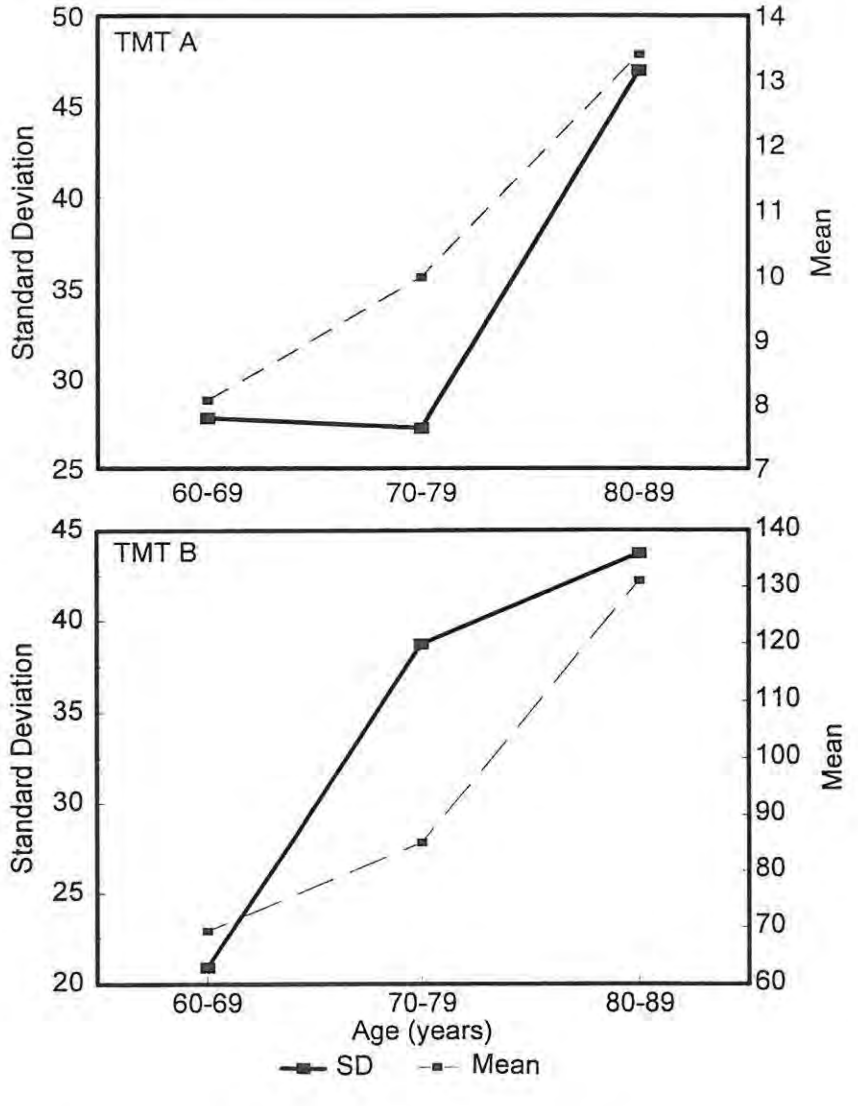


Figure 8.1. Trail Making Test Parts A and B: Standard deviations and mean scores across adult age groups.

Figure 8.1 the significant linear increases in variability for both TMT A and TMT B occurred in association with declining average performance as reflected in significantly increased time scores between the younger and older adult years.

For the *Digit Symbol* test (see Figure 8.2 upper graph, based on data contained in Tables 8.6 and 8.7) the comparisons of standard deviations across all five age groups did not reach significance ($p=0.1351$). However the artifact of significantly increased variability due to age in itself in the younger two age groups (as noted above in the analyses of subject characteristics, section 8.1) is likely to have obscured the presence of a highly significant non-linear inverted-U pattern of variability for this test. This is confirmed in that there was a significant effect for the comparisons of standard deviations across the three older age groups ($p=0.0436$) for which there were evenly spaced age intervals. Further backwards pairwise comparisons (see Table 8.7), with levels of significance adjusted for a three-group comparison, revealed a significantly lower standard deviation in the 80-89 year-old age group compared with each of the 70-79 and 60-69 year-old age groups ($p=0.0133^{**}$, and 0.0277^{*}, respectively). Even allowing for the leniency of a three-group adjustment in the level of significance, there was no significant difference between the 80-89 year old age group and the two youngest 20-39 and 40-49 year-old age groups.

Hence overall for the Digit Symbol test, there was support for a significant non-linear pattern of an initial increase in variability by the 60's age stage which was sustained in the 70's age stage, and which was followed by a subsequent decrease in variability by the 80's age stage. Were younger pre-60 age groups constituted with more restricted age ranges to match the older 60+ age groups, thereby reducing the artificial increase in variability due to age in the younger age stages which was apparent for this particular data set, the effect in all probability would be for the overall non-linear inverted-U effect to become even more pronounced.

For Digit Symbol the comparisons of means across all groups (see Figure 8.2 upper graph, based on data contained in Tables 8.6 and 8.7) revealed a highly significant effect ($p=0.01$), and backwards pairwise comparisons revealed that there was a significantly lower mean score between the 80-89 year-old age group and each of the 60-69, 40-59 and 20-39 year-old age groups ($p<0.001^{***}$, in all instances). Thus the significant non-linear inverted-U pattern of variability for Digit Symbol occurred in association with a significant decline in average performance between the early and late adult years.

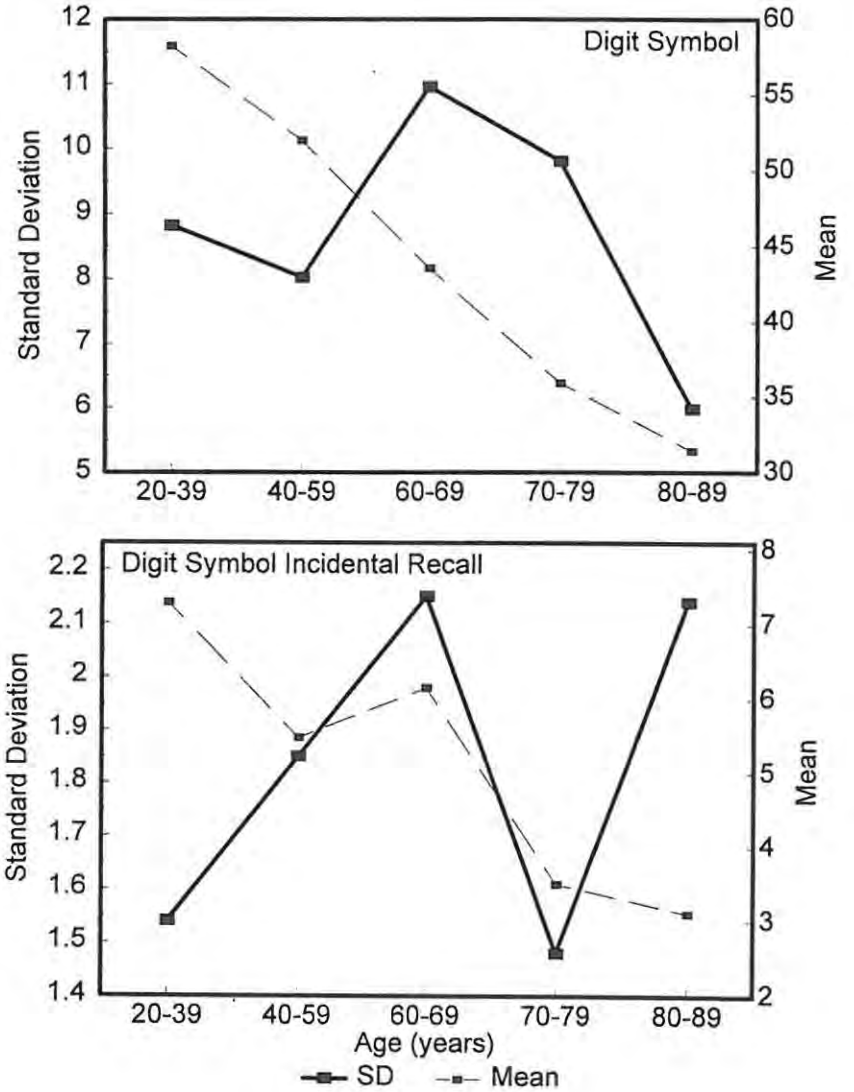


Figure 8.2. Digit Symbol and Incidental Recall: Standard deviations and mean scores across adult age groups.

For *Digit Symbol Incidental Recall* (see Figure 8.2 lower graph, based on data contained in Tables 8.6 and 8.7) the trend was of a haphazard zig-zag progression in the standard deviations across age groups, a pattern which revealed no significant effect in either of the five or three-group sets of comparisons. This lack of a significant effect for variability was confirmed by the backwards comparisons (see Table 8.7) which revealed no significant differences between standard deviations for any of the pairwise comparisons. (The zig-zag pattern depicted graphically appears marked due to the scale in relation to the mean score, but in effect only varies between 2.15 and 1.48).

The absence of an effect for variability for Incidental Recall occurred in association with a highly significant effect of diminishing mean scores in association with age ($p < 0.01$ on the all-groups analysis) (see Figure 8.2 lower graph, based on data contained in Tables 8.6 and 8.7). The backwards comparisons revealed a significantly lower mean score for the 80-89 year-old age group compared with each of the 60-69, 40-59 and 20-39 year-old age groups ($p < 0.001^{***}$, in all cases). Scrutiny of the range of scores for this test reveal that there was a particularly high level of inherent variability across *all* age groups for this task (as noted by Shuttleworth-Jordan & Bode, 1995a), which can account for the lack of a coherent effect for variability on this task in association with older age.

For the *Finger Tapping Test*, there was a consistent trend of increasing variability in association with older age, although the comparisons of standard deviations across all groups (see Figure 8.3 based on data contained in Tables 8.8 and 8.9), revealed significant effects only for the more challenging condition of the non-preferred hand ($p=0.0402$). Backwards comparisons (see Table 8.9) confirmed no significant differences between age groups for the preferred hand although the difference between the oldest 80-89 year-old age group and the youngest 20-39 year-old age group was strongly indicative of the trend for a higher standard deviation in the older age group ($p=0.0676$). For the non-preferred hand backwards comparisons revealed that there was no significant difference between the oldest 80-89 year-old group and the 70-79 and 60-69 year-old groups, the strong indication of a higher standard deviation in the 80-89 year-old group compared with the 40-59 year-old group ($p=0.0269$), and a significantly higher standard deviation in the 80-89 year-old group compared with the 20-39 year-old age group ($p=0.0091^{**}$). The fact that the inherent variability in the two youngest age groups was probably artificially inflated relative to the older age groups due to the artifact of wider age ranges in these youngest groups (see section 8.1, analysis of subject characteristics above), is likely to have significantly *reduced* the chances of detecting an increase in the standard deviations in the older age groups when compared with the younger age groups.

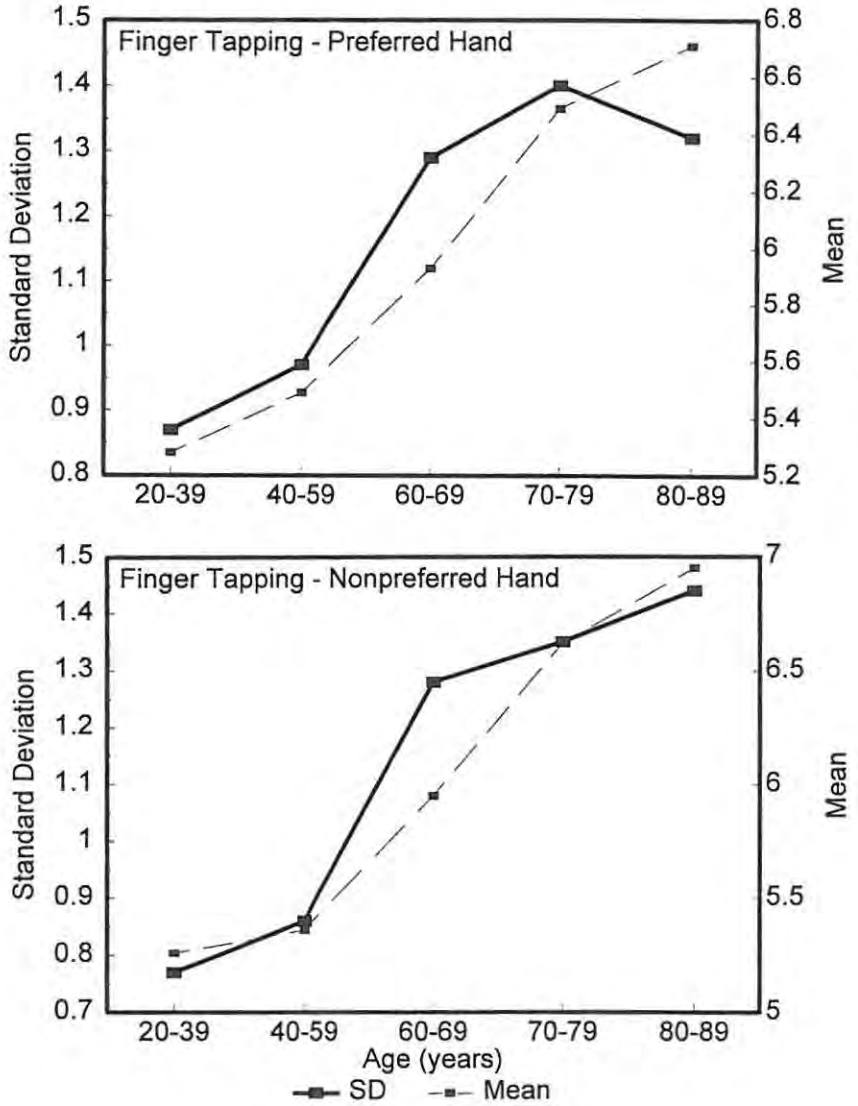


Figure 8.3. Finger Tapping Preferred and Nonpreferred Hands: Standard deviations and mean scores across adult age groups.

Hence the overall indication for the finger tapping test was of a linear pattern of increasing variability in association with older age by the 70's and 80's age decades which was definitely present for the non-preferred hand, and in all probability also for the preferred hand had the younger groups been equivalent in age range.

For Finger Tapping across both the preferred and non-preferred hand conditions (see Figure 8.3, based on data contained in Tables 8.8 and 8.9), the pattern of increasing variability in association with age occurred in parallel with significantly declining average performance as reflected in steadily increasing timed scores between the youngest and oldest years. The comparisons of mean scores across all groups revealed a highly significant effect ($p<0.01$, in both cases); the backwards comparisons confirmed a significantly higher mean score in the 80-89 year-old age group compared with the 40-49 and 20-39 year-old age groups for the preferred hand ($p<0.001^{***}$ in both cases), and a significantly higher mean score between the 80-89 year-old age group and the 60-69, 40-59 and 20-39 year-old age groups ($p<0.001^{***}$ and $p=0.0091^{**}$, respectively).

8.2.2 Summary of between-groups effects

Generally for all the tests included in the Phase 2 between-groups analysis, with the exception of Digit Symbol Incidental Recall (a task which has been previously identified to have the confounding factor of an exceptionally high range of scores regardless of age group), there was evidence of significantly increasing variability in association with older age, either with or without a subsequent decrease in variability, which occurred in parallel with declining average ability. Hence the results provided consistent support for the shuttle bulge effect of an initial differentiation effect, either with or without subsequent dedifferentiation, in association with overall cognitive aging.

For the Trail Making Test the increase in variability occurred by the 70's age decade for the task of greater challenge (TMT B), a decade earlier than for the task of lesser challenge (TMT A) (which constitutes a left shuttle effect and right shuttle effect for TMT A and TMT B, respectively). For both TMT A and B there was no subsequent decrease in variability. For the Finger Tapping Test generally, the indication was of a significant increase in variability by the 70's and 80's decades. In contrast to both the Trail Making and Finger Tapping tests, for Digit Symbol a relatively early significant increase in variability was evident by the 60's age decade followed by a significant decrease by the 80's age decade. This constitutes a task-related left shuttle effect for Digit Symbol in relation to the Trail Making and Finger Tapping tests.

These findings for the Phase 2 between-groups analyses are broadly commensurate with the outcome for the Phase 1 analyses. Generally for both Phase 1 and 2 there was robust evidence for the presence of significantly increasing variability, with or without a subsequent decrease in variability depending on the task (that is the shuttle bulge effect), in association with older age and in parallel with a significant decline in cognitive ability. Where this did not occur it was explicable for idiosyncratic reasons such as erratic average performances, or as in the case of Digit Symbol Incidental Recall, as a consequence of exceptionally high inherent variability of task performance regardless of age stage.

As for the analyses in Phase 1, in the Phase 2 analyses the Trail Making Test revealed a pattern of significantly increasing variability without a subsequent decrease (the initial differentiation stage of the shuttle bulge effect). Also commensurate with the results from Phase 1 was the pattern of no increases in variability for the Finger Tapping test by the 60's, although the Phase 2 data indicated that there was a significant increase in later old age, also without a subsequent decrease (again, the initial differentiation phase of the shuttle bulge effect). Finally, as in Phase 1, there was evidence for the earlier versus later onset of the variability effect (left versus right shuttle effect) for a task of relatively high challenge versus a task of relatively low challenge, respectively. The reason why the challenge-related shuttle effect for the Trail Making Test was in evidence for the Cornfield and Shuttleworth-Jordan data in the Phase 2 analysis, and not for the van Gorp et al. (1990) and Stuss et al. (1987) data in the Phase 1 analyses is not clear. It is possible, however, that well-matched groups for educational and IQ level in the Cornfield and Shuttleworth-Jordan samples (which was clearly not the case for the samples in the van Gorp et al. and Stuss et al. studies) has provided the opportunity for this effect to emerge.

Table 8.4. Trail Making Test: Range and Comparisons of Mean Scores and of Standard Deviations across all Age Groups.

AGE GROUP	60-69	70-79	80-89	df	F	p
n	33	38	34			
TMT A						
Mean	28.77	35.64	47.85	2,102	32.75	<0.01**
SD	7.78	7.63	13.15	2,102	3.31	0.0403*
Range	17-43	23-55	29-79	-	-	-
TMT B						
Mean	69.32	85.00	131.05	2,102	26.84	<0.01**
SD	20.90	38.71	43.74	2,102	5.46	<0.01**
Range	37-123	45-179	80-270	-	-	-

Significant difference (*$p<0.05$; **$p<0.01$).

Table 8.5. Trail Making Test: Pairwise Comparisons of Mean Scores and of Standard Deviations between the oldest Age Group and each younger Age Group.

AGE GROUP	60-69	70-79	80-89	df	t	F	p
n	33	38	34	-	-	-	-
TMT A							
Mean	-	35.64	47.85	70	-4.88	-	<0.001***
SD	-	7.63	13.15	1,70	-	4.35	0.0407*
Mean	28.77	-	47.85	65	-7.25	-	<0.001***
SD	7.78	-	13.15	1,65	-	3.59	0.0625
TMT B							
Mean	-	85.00	131.05	70	-4.74	-	<0.001***
SD	-	38.71	43.74	1,70	-	0.27	0.6072
Mean	69.32	-	131.05	65	-7.33	-	<0.001***
SD	20.90	-	43.74	1,65	-	10.24	0.0021**

Significant difference (*p<0.10; **p<0.05; ***p<0.01) as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

Table 8.6. Digit Symbol and Incidental Recall: Range and Comparisons of Mean Scores and of Standard Deviations across all Age Groups, and across the three older Age Groups.

AGE GROUP	20-39	40-59	60-69	70-79	80-89	df	F	p
n	23	21	28	31	28			
Digit Symbol								
Five Group Comparison								
Mean	58.24	51.95	43.55	35.92	31.38	4,126	38.62	<0.01**
SD	8.82	8.02	10.96	9.81	5.98	4,126	1.79	0.1351
Three Group Comparison								
Mean	-	-	43.55	35.92	31.38	2,84	12.56	<0.01**
SD	-	-	10.96	9.81	5.98	2,84	3.25	0.0436*
Range	33-67	39-67	26-67	20-58	20-46	-	-	-
Inc Recall								
Five Group Comparison								
Mean	7.30	5.38	6.16	3.50	3.11	4,126	24.12	<0.01**
SD	1.54	1.85	2.15	1.48	2.14	4,126	1.66	0.1634
Three Group Comparison								
Mean	-	-	6.16	3.50	3.11	2,84	20.91	<0.01**
SD	-	-	2.15	1.48	2.14	2,84	2.29	0.1071
Range	3-9	2-9	2-9	1-6	0-8	-	-	-

Significant difference (*p<0.05; **p<0.01).

Table 8.7. Digit Symbol and Incidental Recall: Pairwise Comparisons of Mean Scores and of Standard Deviations between the oldest Age Group and each younger Age Group.

AGE GROUP	20-39	40-59	60-69	70-79	80-89	df	t	F	p
n	23	21	28	31	28	-	-	-	-
Digit Symbol									
Five Group Comparison									
Mean	-	-	-	35.92	31.38	57	2.12	-	0.0383
SD	-	-	-	9.81	5.98	1,57	-	6.53	0.0133*
Mean	-	-	43.55	-	31.38	54	5.16	-	<0.001***
SD	-	-	10.96	-	5.98	1,54	-	5.12	0.0277
Mean	-	51.95	-	-	31.38	47	10.29	-	<0.001***
SD	-	8.02	-	-	5.98	1,47	-	3.91	0.0540
Mean	58.24	-	-	-	31.38	49	12.91	-	<0.001***
SD	8.82	-	-	-	5.98	1,49	-	2.20	0.1444
Three Group Comparison									
Mean	-	-	-	35.92	31.38	57	2.12	-	0.0383*
SD	-	-	-	9.81	5.98	1,57	-	6.53	0.0133**
Mean	-	-	43.55	-	31.38	54	5.16	-	<0.001***
SD	-	-	10.96	-	5.98	1,54	-	5.12	0.0277*
Inc Recall									
Mean	-	-	-	3.50	3.11	57	0.83	-	0.4116
SD	-	-	-	1.49	2.14	1,57	-	3.93	0.0522
Mean	-	-	6.16	-	3.11	54	5.33	-	<0.001***
SD	-	-	2.15	-	2.14	1,54	-	0.00	0.9753
Mean	-	5.38	-	-	3.11	47	3.90	-	<0.001***
SD	-	1.85	-	-	2.14	1,47	-	0.30	0.5841
Mean	7.30	-	-	-	3.11	49	7.87	-	<0.001***
SD	1.54	-	-	-	2.14	1,49	-	2.54	0.1171

Significant difference (*p<0.10; **p<0.05; ***p<0.01) as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

Table 8.8. Finger Tapping Preferred and Non-Preferred Hands: Range and Comparisons of Mean Scores and of Standard Deviations across all Age Groups, and across the three older Age Groups.

AGE GROUP	20-39	40-59	60-69	70-79	80-89	df	F	p
n	23	21	27	26	26			
Finger Tapping Pref								
Five Group Comparison								
Mean	5.28	5.49	5.93	6.49	6.71	4,118	6.44	<0.01**
SD	0.87	0.97	1.29	1.40	1.32	4,118	1.69	0.1561
Three Group Comparison								
Mean	-	-	5.93	6.49	6.71	2,76	2.41	0.0969
SD	-	-	1.29	1.40	1.32	2,76	0.27	0.7639
Range	3.9-7.8	3.9-7.8	3-8.6	3-8.3	3.7-9.0			
Finger Tapping Non-Pref								
Five Group Comparison								
Mean	5.26	5.36	5.95	6.62	6.95	4,118	9.52	<0.01**
SD	0.77	0.86	1.28	1.35	1.44	4,118	2.59	0.0402*
Three Group Comparison								
Mean	-	-	5.95	6.62	6.95	2,76	3.80	0.0268*
SD	-	-	1.28	1.35	1.44	2,76	0.24	0.7902
Range	3.9-6-9	3.7-7.2	3.3-8.3	3.1-8.9	4.0-9.9			

Significant difference (*p<0.05; **p<0.01).

Table 8.9. Finger Tapping Preferred and Non-Preferred Hands: Pairwise Comparisons of Mean Scores and of Standard Deviations between Age Groups.

AGE GROUP	20-39	40-59	60-69	70-79	80-89	df	t	F	p
n	23	21	27	26	26	-	-	-	-
Fing Tap Pref									
Mean	-	-	-	6.49	6.71	50	-0.58	-	0.5613
SD	-	-	-	1.40	1.32	1,50	-	0.21	0.6520
Mean	-	-	5.93	-	6.71	51	-2.17	-	0.0345
SD	-	-	1.29	-	1.32	1,51	-	0.08	0.7834
Mean	-	5.49	-	-	6.71	45	-3.59	-	<0.001***
SD	-	0.97	-	-	1.32	1,45	-	1.67	0.2029
Mean	5.28	-	-	-	6.71	47	-4.42	-	<0.001***
SD	0.87	-	-	-	1.32	1,47	-	3.50	0.0676
Fing Tap Non-Pref									
Mean	-	-	-	6.62	6.95	50	-0.85	-	0.4005
SD	-	-	-	1.35	1.44	1,50	-	0.29	0.5910
Mean	-	-	5.95	-	6.95	51	-2.69	-	<0.001***
SD	-	-	1.28	-	1.44	1,51	-	0.41	0.5245
Mean	-	5.36	-	-	6.95	45	-4.45	-	<0.001***
SD	-	0.86	-	-	1.44	1,45	-	5.23	0.0269
Mean	5.26	-	-	-	6.95	47	-5.02	-	<0.001***
SD	0.77	-	-	-	1.44	1,47	-	7.40	0.0091**

Significant difference (*p<0.10; **p<0.05; ***p<0.01) for comparisons of the four younger age groups with the 80-89 year old age group, as determined by Bonferroni's adjustment in the level of significance for simultaneous testing.

8.3 WITHIN-GROUPS ANALYSES

8.3.1 Comparisons of standard deviations and means for levels of education, IQ and gender.

The focus for the analysis of these within-groups statistical comparisons for levels of education, IQ and gender, was specifically on variability effects (as against central tendency effects), in order to locate factors which may have contributed to the increased variability which was noted on the between-groups analyses above (section 8.2). Thus mean score effects were only discussed insofar as they related to significant variability effects within age groups.

8.3.1.1 Education and IQ. For the *Trail Making Test* (see Tables 8.10a and 8.10b), the comparisons of standard deviations within-groups revealed only the following significant effects:

In the 60's age group for TMT B, there was a significantly higher standard deviation for the high (14+ years) educational condition compared with the low (10-13 years) educational condition ($p<0.01$), a finding which was not evident even in the form of a trend in the 60's age group for Trails A, and not replicated for Trails A or B when the data were analyzed for the high (16+ years) versus low (10-15 years) educational conditions. Thus this does not appear to be a robust effect.

On the other hand for the Trail Making Test in the 80's age group, there was a significantly lower standard deviation for the high (14+ years) educational condition compared with the low (10-13 years) educational condition for both TMT A and TMT B ($p=0.0108$ and $p<0.01$, respectively), a finding which was also in evidence when the data were analyzed for the high (16+ years) versus low (10-15 years) educational conditions as a matching directional trend for TMT A, and as a significant effect for TMT B ($p=0.0221$). Furthermore, for the 80's age group, in the absence of any other significant effects for NART IQ, there was a significantly lower standard deviation for the high (130+ IQ condition) compared with the low (< 130 IQ condition) for TMT A ($p=0.0212$) which was present also as a strong directional trend in the 80's age group for TMT B. Consistently these effects occurred in conjunction with a tendency for mean time scores to be lower for the high education/IQ subsets, which was not a consistent trend throughout the comparisons for other age groups. Overall therefore for the Trail Making Test there was the strong indication that (i) the effect of education and IQ on variability and mean scores became a significant feature only in the oldest 80's age stage, and that (ii) lower levels of education and IQ contributed to the enhanced variability and impaired average performance in the 80's age stage and not in the 70's and 60's age stages.

The relatively robust within-groups variability effect for education and IQ in the 80's age group for the Trail Making Test appears to be highly domain specific. For *Digit Symbol* (see Tables 8.12a and 8.12b) in the 60's age group there was only a significantly higher standard deviation for the higher (14+years) educational condition compared with the lower (10-13 years) educational condition ($p=0.0144$) which was supported by a directional trend in the 70's age group, but was not supported by any significant effects for either the 60's or 70's age groups in the comparisons from the other set of educational conditions. For *Finger Tapping* Non-Preferred Hand (see Tables 8.13a and 8.13b) in the 20-39 year-old age group there was only a significantly higher standard deviation in the higher (16+ years) educational condition compared with the lower (10-15 years) educational condition ($p=0.0190$) which was not supported by any other significant findings in the comparisons of standard deviations for the 20-39 year-old age group using the other set of educational conditions, nor by any of the 20-39 year-old age group comparisons for the Finger Tapping Preferred Hand. Thus particularly in view of the fact that the 20-39 year-old age group was not one for which there was increased variability, the latter was neither a robust nor a relevant finding. For Digit Symbol Incidental Recall (see Table 8.15) there was a significantly lower standard deviation for males in the 40-59 year-old age group ($p=0.0255$) which is of no relevance since (as noted in section 8.2.1 above) there was no overall significant effect for variability on this test between-groups. Further to this there were no significant differences between standard deviations for educational conditions within age groups for any of the tests.

8.3.1.2 <u>Gender</u>. For the *Trail making Test* (see Table 8.14) the only significant variability effect occurred in the 80's age group for TMT A in which the standard deviation for males was significantly lower than that for females ($p<0.01$), a finding which was supported by a strong trend in the same direction in the 80's age group for TMT B ($p=0.0698$). This variability effect for gender occurred in association with no significant difference between mean scores in the 80's age group for either TMT A or TMT B ($p=0.4717$, and 0.8054, respectively). It is not possible to discount this finding for the TMT of a significant variability effect for gender in the 80's age group as an artifact of disproportionate numbers of males to females in the gender subsets (14 and 20, respectively). This is because a disproportionately large proportion of females relative to males within groups is true for *both* the other age groups (for which there were *no* signs of a variability effect for gender for either TMT A or TMT B). Furthermore the balance of 14 males relative to 20 females is at its most even in this 80's age group relative to the other groups which comprise the following much lower proportions of males relative to females (7 and 26; 14 and 24; for the age groups 60-69 and 70-79, respectively).

For Digit Symbol Incidental Recall (see Table 8.15) there was a significantly lower standard deviation in the 40-59 year-old age group which is of no relevance since (as noted in section 8.2.1 above), there was no overall significant effect for variability on this test between groups. Further to this there were no significant variability effects for gender for either Digit Symbol (see Table 8.15) or the Finger Tapping Test (see Table 8.16).

Table 8.10a. Trail Making Test: Comparisons of Mean Scores and of Standard Deviations for two Levels of Education (10-13 and 14+ years).

EDUCATION		10-13	14+	df	t/F	p	10-13	14+	df	t/F	p
		TMT A					*TMT B*				
AGE GROUP											
60-89											
	n	32	73	-	-	-	32	73	-	-	-
	Mean	38.19	37.10	103	0.41	0.6810	94.64	95.13	103	-0.05	0.958
	SD	15.54	10.91	1,103	2.72	0.1020	55.76	38.35	1,103	1.91	0.170
60-69											
	n	13	20	-	-	-	13	20	-	-	-
	Mean	30.04	27.94	31	0.75	0.4587	65.96	71.51	25	-0.74	0.4653
	SD	7.92	7.78	1,31	0.54	0.4676	8.38	26.06	1,25	13.42	<0.01**
70-79											
	n	11	27	-	-	-	11	27	-	-	-
	Mean	34.01	36.30	36	-0.84	0.4090	81.96	86.23	36	-0.30	0.7623
	SD	7.81	7.60	1,36	0.00	0.9484	38.43	39.49	1,36	0.07	0.7986
80-89											
	n	8	26	-	-	-	8	26	-	-	-
	Mean	57.20	44.98	32	2.47	0.0190*	158.69	122.55	32	2.15	0.0389*
	SD	17.65	10.22	1,32	7.32	0.0108*	71.23	28.02	1,32	15.56	<0.01**

Significant difference (*p<0.05; **p<0.01).

Table 8.10b. Trail Making Test: Comparisons of Mean Scores and of Standard Deviations for two levels of Education (10-15 and 16+ years).

EDUCATION		10-15	16+	df	t/F	p	10-15	16+	df	t/F	p
		TMT A					*TMT B*				
AGE GROUP											
60-89											
	n	61	44	-	-	-	61	44	-	-	-
	Mean	36.74	38.40	103	-0.68	0.5007	95.90	93.72	103	0.25	0.8043
	SD	12.81	11.98	1,103	0.00	0.9645	48.91	36.94	1,103	1.17	0.2821
60-69											
	n	22	11	-	-	-	22	11	-		
	Mean	29.78	26.74	31	1.06	0.2965	71.66	64.65	31	0.91	0.3727
	SD	7.76	7.79	1,31	0.49	0.4911	22.48	17.36	1,31	0.39	0.5374
70-79											
	n	22	16	-	-	-	22	11	-		
	Mean	33.34	38.81	36	-2.31	0.0269*	82.90	87.88	36	-0.39	0.7005
	SD	5.93	8.71	1,36	3.24	0.0801	37.44	41.47	1,36	0.37	0.5476
80-89											
	n	17	17	-	-	-	17	17	-		
	Mean	50.13	45.57	32	1.01	0.3189	144.09	118.02	32	1.79	0.0820
	SD	14.73	11.33	1,32	1.12	0.2969	54.41	24.98	1,32	5.79	0.0221*

Significant difference (*$p < 0.05$).

Table 8.11. Trail Making Test: Comparisons of Mean Scores and of Standard Deviations for two levels of NART IQ (<130 and 130+).

NART IQ		<130	130+	df	t/F	p	<130	130+	df	t/F	p
		TMT A					*TMT B*				
AGE GROUP											
60-89											
	n	58	47	-	-	-	58	47	-	-	-
	Mean	37.83	36.95	103	0.36	0.7208	99.47	89.44	103	1.16	0.2483
	SD	13.78	10.69	1,103	0.87	0.3540	48.40	37.93	1,103	0.92	0.3395
60-69											
	n	19	14	-	-	-	19	14	-	-	-
	Mean	28.89	28.60	31	0.10	0.9150	70.54	67.66	31	0.39	0.7023
	SD	6.47	9.54	1,31	3.45	0.0730	19.92	22.83	1,31	0.28	0.5979
70-79											
	n	23	15	-	-	-	23	15	-	-	-
	Mean	35.35	36.08	36	-0.28	0.7760	93.10	72.56	36	1.63	0.1110
	SD	7.43	8.17	1,36	0.01	0.9400	40.86	32.62	1,36	0.78	0.3820
80-89											
	n	16	18	-	-	-	16	18	-	-	-
	Mean	51.99	44.17	22	1.79	0.0828	142.99	120.44	32	1.53	0.1357
	SD	16.28	8.44	1,22	5.88	0.0212*	53.24	30.97	1,32	1.89	0.1793

Significant difference (*p<0.05)

Table 8.12a. Digit Symbol and Incidental Recall: Comparisons of Mean Scores and of Standard Deviations for two levels of Education (10-13 and 14+ years).

EDUCATION		10-13	14+	df	t/F	p	10-13	14+	df	t/F	p
		Digit Symbol					*Inc Recall*				
AGE GROUP											
20-89											
	n	41	90	-	-	-	41	90	-	-	-
	Mean	39.09	44.88	129	-2.38	0.0186*	4.54	5.14	129	-1.33	0.1852
	SD	10.73	13.78	1,129	6.49	0.0120*	2.52	2.38	1,129	0.16	0.6923
20-39											
	n	3	20	-	-	-	3	20	-	-	-
	Mean	58.83	58.15	21	0.10	0.9039	7.33	7.30	21	0.03	0.9731
	SD	7.11	9.21	1,21	0.14	0.7078	1.26	1.61	1,21	0.57	0.4580
40-59											
	n	7	14	-	-	-	7	14	-	-	-
	Mean	49.29	53.29	19	-1.08	0.2929	4.93	5.61	19	-0.79	0.4423
	SD	7.04	8.39	1,19	0.64	0.4330	1.92	1.84	1,19	0.00	0.9585
60-69											
	n	12	16	-	-	-	12	16	-	-	-
	Mean	39.50	46.59	26	-1.76	0.0901	6.00	6.28	26	0.33	0.7386
	SD	6.29	12.81	1,26	6.89	0.0144*	2.59	1.83	1,26	1.10	0.3034
70-79											
	n	11	20	-	-	-	11	20	-	-	-
	Mean	34.00	36.98	29	-0.81	0.4282	3.86	3.30	29	1.02	0.3177
	SD	8.14	10.66	1,29	0.52	0.4771	1.10	1.64	1,29	3.49	0.0718
80-89											
	n	8	20	-	-	-	8	20	-	-	-
	Mean	29.13	32.28	26	-1.27	0.2143	1.88	3.60	26	-2.04	0.0521
	SD	4.30	6.41	1,26	0.80	0.3800	2.08	2.00	1,26	0.09	0.7645

Significant difference (*$p < 0.05$)

Table 8.12b. Digit Symbol and Incidental Recall: Comparisons of Mean Scores and of Standard Deviations for two levels of Education (10-15 and 16+ years).

EDUCATION		10-15	16+	df	t/F	p	10-15	16+	df	t/F	p
		Digit Symbol					*Inc Recall*				
AGE GROUP											
20-89											
	n	71	60	-	-	-	71	60	-	-	-
	Mean	41.59	44.83	129	-1.41	0.1606	4.88	5.04	129	-0.37	0.7064
	SD	12.93	13.29	1,129	0.61	0.4356	2.54	2.32	1,129	1.31	0.2547
20-39											
	n	8	15	-	-	-	8	15	-	-	-
	Mean	60.44	57.07	21	0.87	0.3953	7.69	7.10	21	0.87	0.3969
	SD	5.34	10.19	1,21	2.33	0.1419	1.25	1.68	1,21	0.41	0.5288
40-59											
	n	11	10	-	-	-	11	10	-	-	-
	Mean	50.64	53.40	19	-0.78	0.4446	5.18	5.60	19	-0.51	0.6175
	SD	9.07	6.87	1,19	1.18	0.2913	2.03	1.71	1,19	1.07	0.3133
60-69											
	n	22	6	-	-	-	21	6	-	-	-
	Mean	42.07	49.00	26	-1.40	0.1740	6.35	4.55	25	3.70	0.0011*
	SD	11.02	9.61	1,26	0.01	0.9084	1.08	0.89	1,25	1.35	0.2555
70-79											
	n	17	14	-	-	-	17	14	-	-	-
	Mean	35.00	37.04	29	-0.57	0.5739	3.47	3.54	29	-0.10	0.9052
	SD	9.38	10.54	1,29	0.15	0.6995	1.22	1.79	1,29	2.69	0.1115
80-89											
	n	13	15	-	-	-	13	15	-	-	-
	Mean	30.12	32.47	26	-1.04	0.3085	2.54	3.60	26	-1.33	0.1960
	SD	6.61	5.37	1,26	0.23	0.6368	2.21	2.02	1,26	0.30	0.5890

Significant difference (*$p<0.05$)

Table 8.13a. Finger Tapping (Preferred and Non-Preferred Hands): Comparisons of Mean Scores and of Standard Deviations for two levels of Education (10-13 and 14+ years).

EDUCATION		10-13	14+	df	t/F	p	10-13	14+	df	t/F	p
		Finger Tapping Preferred					*Finger Tapping Non-Preferred*				
AGE GROUP											
20-89											
	n	39	84	-	-	-	39	84	-	-	-
	Mean	6.28	5.88	121	1.59	0.1153	6.38	5.93	121	1.71	0.0890
	SD	1.19	1.35	1,121	0.00	0.4393	1.11	1.43	1,121	1.66	0.2002
20-39											
	n	3	20	-	-	-	3	20	-	-	-
	Mean	5.41	5.26	21	0.28	0.7843	5.51	5.23	21	0.59	0.5629
	SD	0.73	0.90	1,21	0.24	0.6286	0.62	0.80	1,21	0.25	0.6240
40-59											
	n	7	14	-	-	-	7	14	-	-	-
	Mean	5.41	5.50	19	-0.20	0.8404	5.56	5.26	19	1.24	0.4739
	SD	0.88	1.04	1,19	0.42	0.5251	0.64	0.96	1,19	1.35	0.2604
60-69											
	n	11	16	-	-	-	11	16	-	-	-
	Mean	6.09	5.82	25	0.54	0.5951	6.14	5.81	25	0.65	0.5236
	SD	1.14	1.42	1,25	0.79	0.3823	1.08	1.42	1,25	0.98	0.3325
70-79											
	n	10	16	-	-	-	10	16	-	-	-
	Mean	7.01	6.16	24	1.54	0.1373	7.12	6.31	24	1.52	0.1406
	SD	1.02	1.53	1,24	2.70	0.1136	0.98	1.48	1,24	2.14	0.1564
80-89											
	n	8	18	-	-	-	8	18	-	-	-
	Mean	6.74	6.69	24	0.10	0.9374	6.82	7.01	24	-0.30	0.7616
	SD	1.22	1.40	1,24	0.00	0.9674	1.11	1.60	1,24	0.69	0.4157

Table 8.13b. Finger Tapping (Preferred and Non-Preferred Hands): Comparisons of Mean Scores and of Standard Deviations for two levels of Education (10-15 and 16+ years).

EDUCATION		10-15	16+	df	t/F	p	10-15	16+	df	t/F	p
		Finger Tapping Preferred					*Finger Tapping Non-Preferred*				
AGE GROUP											
20-89											
	n	67	56	-	-	-	67	56	-	-	-
	Mean	6.10	5.91	121	0.80	0.4269	6.22	5.90	121	1.32	0.1901
	SD	1.23	1.40	1,121	0.95	0.3327	1.27	1.43	1,121	0.48	0.4916
20-39											
	n	8	15	-	-	-	8	15	-	-	-
	Mean	5.35	5.23	21	0.30	0.7610	5.47	5.15	21	0.93	0.3602
	SD	0.52	1.02	1,21	2.90	0.1031	0.36	0.91	1,21	6.45	0.0190*
40-59											
	n	11	10	-	-	-	11	10	-	-	-
	Mean	5.23	5.74	19	-1.23	0.2349	5.18	5.57	19	-1.03	0.3154
	SD	0.91	1.00	1,19	0.02	0.8983	0.92	0.79	1,19	0.90	0.3546
60-69											
	n	21	6	-	-	-	21	6	-	-	-
	Mean	6.24	4.82	25	2.63	0.0144*	6.35	4.55	25	3.70	0.0011*
	SD	1.15	1.25	1,25	0.03	0.8750	1.08	0.89	1,25	1.35	0.2555
70-79											
	n	14	12	-	-	-	14	12	-	-	-
	Mean	6.61	6.35	24	0.47	0.6456	6.70	6.53	24	0.32	0.7595
	SD	1.39	1.46	1,24	0.58	0.4547	1.37	1.38	1,24	0.33	0.5707
80-89											
	n	13	13	-	-	-	13	13	-	-	-
	Mean	6.51	6.91	24	-0.78	0.4434	6.85	7.06	24	-0.36	0.7207
	SD	1.24	1.42	1,24	0.00	0.9753	1.40	1.54	1,24	0.01	0.9114

Significant difference (*$p < 0.05$).

Table 8.14. Trail Making Test: Comparisions of Means and of Standard Deviations for Gender.

GENDER		Female	Male	df	t/F	p	Female	Male	df	t/F	p
		TMT A					*TMT B*				
AGE GROUP											
60-89											
	n	70	35	-	-	-	70	35	-	-	-
	Mean	35.95	40.40	103	-1.74	0.0840	91.75	101.46	103	-1.06	0.2898
	SD	13.65	9.04	1,103	2.09	0.1509	46.86	37.81	1,103	0.59	0.4438
60-69											
	n	26	7	-	-	-	26	15	-	-	-
	Mean	27.27	34.34	31	-2.27	0.0304	67.63	75.61	31	-0.89	0.3785
	SD	6.81	9.16	1,31	2.75	0.1073	22.31	14.07	1,31	0.56	0.4586
70-79											
	n	24	14	-	-	-	24	14	-	-	-
	Mean	34.29	37.96	24	-1.45	0.1554	83.80	87.04	36	-0.25	0.8072
	SD	7.08	8.24	1,24	0.32	0.5763	39.84	38.09	1,36	0.06	0.8090
80-89											
	n	20	14	-	-	-	20	14	-	-	-
	Mean	49.24	45.88	32	0.73	0.4717	132.64	128.79	32	0.25	0.8054
	SD	16.19	6.96	1,32	7.50	<0.01**	52.57	28.42	1,32	3.52	0.0698

Significant difference (**p<0.01).

Table 8.15. Digit Symbol and Incidental Recall: Comparisons of Mean Scores and of Standard Deviations for Gender.

GENDER		Female	Male	df	t/F	p	Female	Male	df	t/F	p
		Digit Symbol					*Incidental Recall*				
AGE GROUP											
20-89											
	n	93	38	-	-	-	93	38	-	-	-
	Mean	43.70	41.51	129	0.87	0.3887	5.09	4.63	129	0.97	0.3332
	SD	13.29	12.82	1,129	0.95	0.3308	2.53	2.17	1,129	1.84	0.1778
20-39											
	n	15	8	-	-	-	15	8	-	-	-
	Mean	57.93	58.81	21	-0.22	0.8260	7.53	6.88	21	0.97	0.3414
	SD	8.62	9.77	1,21	0.53	0.4731	1.36	1.87	1,21	0.09	0.7673
40-59											
	n	15	6	-	-	-	15	6	-	-	-
	Mean	53.20	48.83	19	1.14	0.2705	5.67	4.67	19	1.13	0.2740
	SD	8.34	6.80	1,19	1.24	0.2801	2.08	0.88	1,19	5.87	0.0255*
60-69											
	n	24	4	-	-	-	24	4	-	-	-
	Mean	44.65	37.00	26	1.31	0.2019	6.10	6.50	26	-0.33	0.7398
	SD	11.17	7.57	1,26	0.68	0.4173	2.24	1.73	1,26	0.76	0.3905
70-79											
	n	21	10	-	-	-	21	10	-	-	-
	Mean	37.79	32.00	29	1.57	0.1265	3.71	3.05	29	1.18	0.2485
	SD	10.20	8.00	1,29	0.47	0.4980	1.38	1.64	1,29	0.38	0.5447
80-89											
	n	18	10	-	-	-	18	10	-	-	-
	Mean	29.58	34.60	26	-2.29	0.0306*	2.81	3.65	26	-1.00	0.3263
	SD	5.71	5.28	1,26	0.38	0.5424	2.36	1.65	1,26	1.54	0.2250

Significant difference ($*p < 0.05$).

Table 8.16. Finger Tapping Preferred and Non-Preferred Hands: Comparisons of Mean Scores and of Standard Deviations for Gender.

Gender		Female	Male	df	t/F	p	Female	Male	df	t/F	p
		Finger Tapping Preferred					*Finger Tapping Non-Preferred*				
AGE GROUP											
20-89											
	n	88	35	-	-	-	88	35	-	-	-
	Mean	5.93	6.22	121	-1.12	0.2643	6.02	6.22	121	-0.77	0.4450
	SD	1.31	1.31	1,121	0.05	0.8322	1.34	1.38	1,121	0.04	0.8339
20-39											
	n	15	8	-	-	-	15	8	-	-	-
	Mean	5.27	5.29	21	-0.05	0.9603	5.30	5.20	21	0.28	0.7808
	SD	1.01	0.56	1,21	1.81	0.1925	0.84	0.68	1,21	0.57	0.4598
40-59											
	n	15	6	-	-	-	15	6	-	-	-
	Mean	5.31	5.87	19	-1.21	0.2413	5.26	5.62	19	-0.86	0.3993
	SD	0.85	1.21	1,19	0.86	0.3646	0.80	1.04	1,19	0.60	0.4500
60-69											
	n	23	4	-	-	-	23	4	-	-	-
	Mean	5.80	6.64	25	-1.20	0.2434	5.86	6.42	25	-0.80	0.4305
	SD	1.27	1.35	1,25	0.00	0.9545	1.31	1.14	1,25	0.26	0.6160
70-79											
	n	19	7	-	-	-	19	7	-	-	-
	Mean	6.18	7.32	24	-1.93	0.0655	6.35	7.38	24	-1.82	0.0812
	SD	1.45	0.85	1,24	2.54	0.1238	1.37	1.01	1,24	1.11	0.3036
80-89											
	n	16	10	-	-	-	16	10	-	-	-
	Mean	6.99	6.25	24	1.44	0.1622	7.23	6.51	24	1.24	0.2254
	SD	1.11	1.56	1,24	1.94	0.1766	1.27	1.66	1,24	0.96	0.3369

8.3.2 Analyses of outliers

The focus of these within-groups analyses of outliers was to examine the nature of variability changes and to identify the extent to which the factors of education, IQ and gender have contributed to the increased variability which was noted above (section 8.2) on the overall analyses between groups. Thus whilst for completeness the characteristics of outliers were tabulated for all the tests and all age groups (see Tables 8.17, 8.18 and 8.19, pp.363-5), they were only discussed with respect to the age groups for which there were significant variability effects. From the within-samples analyses in the previous section it was apparent that levels of education, IQ and gender contributed to the patterns of significant alterations in variability more robustly for the Trail Making Test than for any of the other tests, and hence the characteristics of the outliers for this test were illustrated graphically and examined in particular detail.

8.3.2.1 Trail Making Test For the *Trail Making Test* it was apparent that for both parts A and B (see Histograms 8.2a and 8.2b, pp.366-7) there was significant overlap in the scores between the youngest 60's groups and the oldest 80's groups. The significantly widened distribution in the 80's for TMT A, and in the 70's and the 80's for TMT B, was due to a widening of the distribution which included a number of disproportionately good scorers in relation to the mean, and in particular a number of disproportionately very *poor* scorers.

Reference to Table 8.17 indicates the specific characteristics of these good and poor outliers with respect to education, NART IQ and gender. Following this table horizontally for TMT A, taking particular note of those more heavily weighted proportions of individuals noted in parenthesis which lie beyond 2SD's, it can be seen that the inflated variability in the 80's age group was a consequence of a high proportion of extremely *poor* scorers who were characterized by low education, low IQ and female gender, although there was a small proportion of individuals with high education who were not protected from being poor scorers. In addition the widening variability for TMT A in the 80's was a consequence of *good* scorers with high education who were relatively equally distributed for IQ level and gender, although high IQ and female gender showed a slight advantage. For TMT B, again reading Table 8.17 horizontally, the widening variability in the 70's age group was associated with a high proportion of extremely *poor* scorers with both low and high education, both high and low IQ, and both levels of gender, with high education, low IQ and female gender being somewhat more prevalent. In addition for TMT B in the 70's the widening variability was associated with *good* scorers with high education, high IQ and who were of female gender. The high variability for TMT B in the 80's age group was made up of a high proportion of extremely *poor* scorers who largely had low education, low IQ and were of female gender. In addition the high variability for TMT B in the 80's was constituted of *good* scorers who were equally balanced for education and IQ, and who were *exclusively* of female gender.

Overall this analysis of percentages of outliers reflects the findings from the statistical comparisons of standard deviations within groups for levels of education, NART IQ and gender discussed in the previous subsection. With respect to educational level and percentages of *poor* scorers for TMT A, and particularly for TMT B, the step-wise progression in Figure 8.4 (second level of graphs, left and right) is of particular note. It was evident that the extent to which high education acted as a protective factor which served to *diminish* the attainment of disproportionately poor scores became more evident with the advance of age (see downwards step-wise progression of percentages of subjects with each increasing age stage, Figure 8.4, second level, left graph). Conversely, the extent to which low education operated as a vulnerability factor contributing to very poor scores *increased* markedly with the progression of older age (see upwards step-wise progression of percentages of subjects with each increasing age stage, Figure 8.4, second level, right graph). These observations are consistent with the finding in the previous analysis for the 80's age group alone (section 8.3.1) of a significantly lower standard deviation for the high (14+ years) educational condition compared with the low (10-13 years) educational condition for both TMT A and TMT B ($p=0.0108$ and $p<0.01$, respectively). The analysis of the outliers reveals that the low standard deviation for the high IQ group was due to the marked reduction in the numbers of disproportionately poor scorers because of the protection factor of high education by the 80's age stage. Hence the overall high variability in the 80's age stage can be attributed to disproportionately low scorers with low education. Finally consideration of the *good* scorers (see Figure 8.4, lower two levels, all graphs) reveals that *generally* for both TMT A and B, the chances of being a disproportionately good scorer were greater with high education (left lower two graphs) than with low education (right lower two graphs). This was an effect which occurred in a somewhat more variable fashion across the

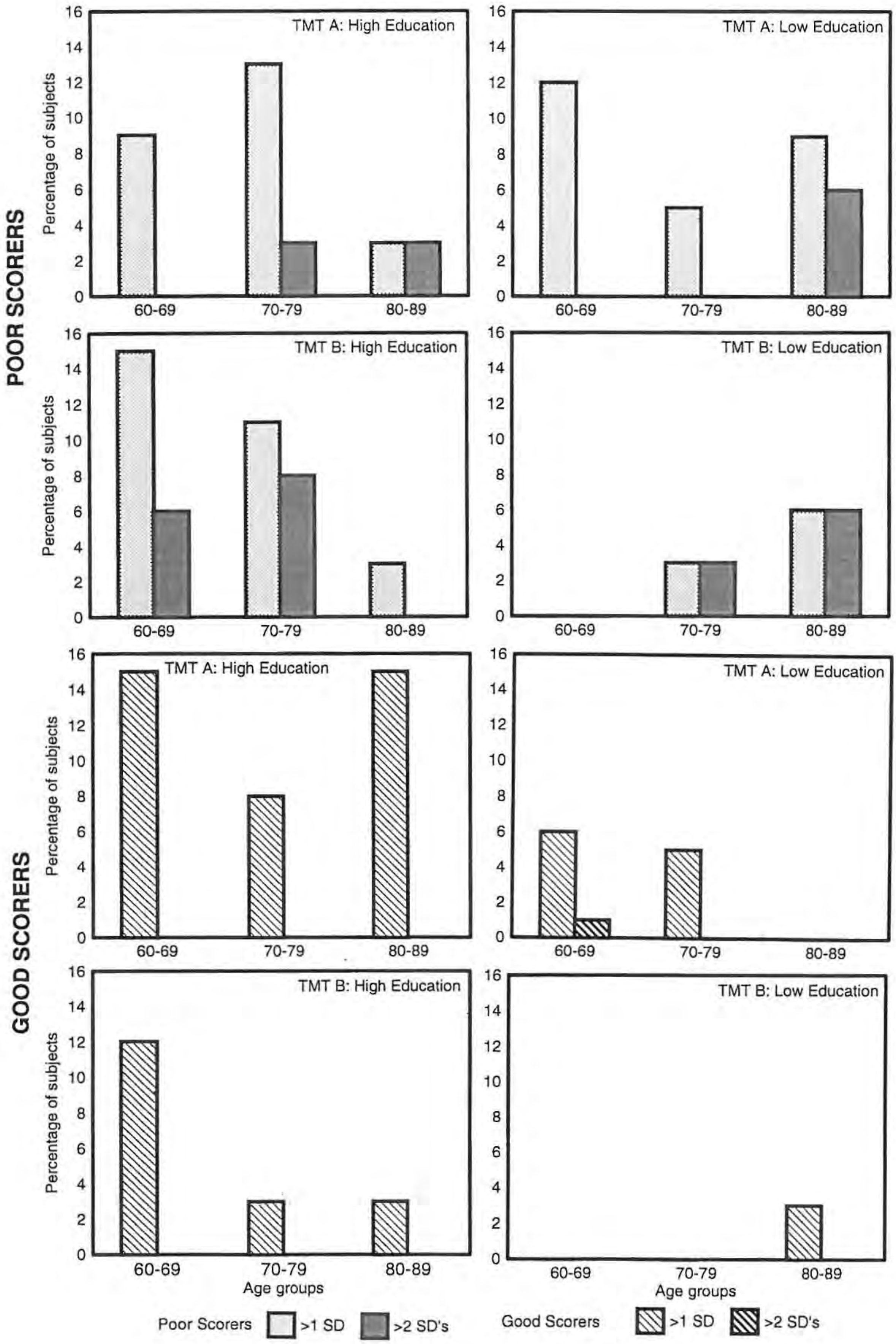


Figure 8.4. Trail Making Test Parts A and B: Percentage of extreme poor and good scorers per age group for low education (10-13 years) and high education (14+ years).

age stages than was evident for poor scorers. (As noted earlier, for *poor* scorers on TMT B - see graphs on the second level - a clear progression of escalating protective effects of high education occurred in association with age).

Commensurate with the findings for education, there were no clear-cut protective or vulnerability effects in operation for NART IQ until the 80's age stage for both TMT A and TMT B at which point high IQ served clearly as a protective effect against being a poor scorer. This is illustrated (see Figure 8.5, upper two graphs on the left), by the downwards step-wise progression of poor scorers which implies the increasing protective effect of high IQ with increasing age. This is consistent with the prior within-groups analysis for the 80's age group (section 8.3.1 above), where in the absence of any other significant effects for NART IQ, there was a significantly lower standard deviation for the high (130+ IQ condition) compared with the low (< 130 IQ condition) for TMT A ($p=0.0212$) which was present also as a strong directional trend in the 80's age group for TMT B.

For gender what was clearly apparent (see Figure 8.6, upper two levels of graphs on the left) is that whereas female gender appears to be a predominant vulnerability factor with respect to being a disproportionately low scorer across all age stages for both TMT A and TMT B, at the same time female gender provides the necessary protective factor to be a disproportionately good scorer (see Figure 8.6, lower two levels on the left). The striking finding for male gender is that whilst it does not entirely provide a protective factor against being a disproportionately poor scorer (Figure 8.6, upper two levels on the right), it is a catastrophic vulnerability factor in terms of failing to produce any good scorers, and this is most particularly so for the more challenging TMT B task for which there are no disproportionately good scorers for males at any of the age stages (Figure 8.6, bottom level, right graph). Thus it appears that a significant dedifferentiation effect has taken place for males in these late age stages, such that their scores are clustering closely around the mean. This observation is commensurate with and explains the analysis of the previous subsection (section 8.3.1) where in the 80's age group for TMT A the standard deviation for males was significantly *lower* than that for females ($p<0.01$), supported by a strong trend in the same direction in the 80's age group for TMT B ($p=0.0698$). As pointed out above, although caution needs to be adopted in the over-interpretation of this gender-related finding due to the disproportionately lower numbers of males relative to females in the samples, lower

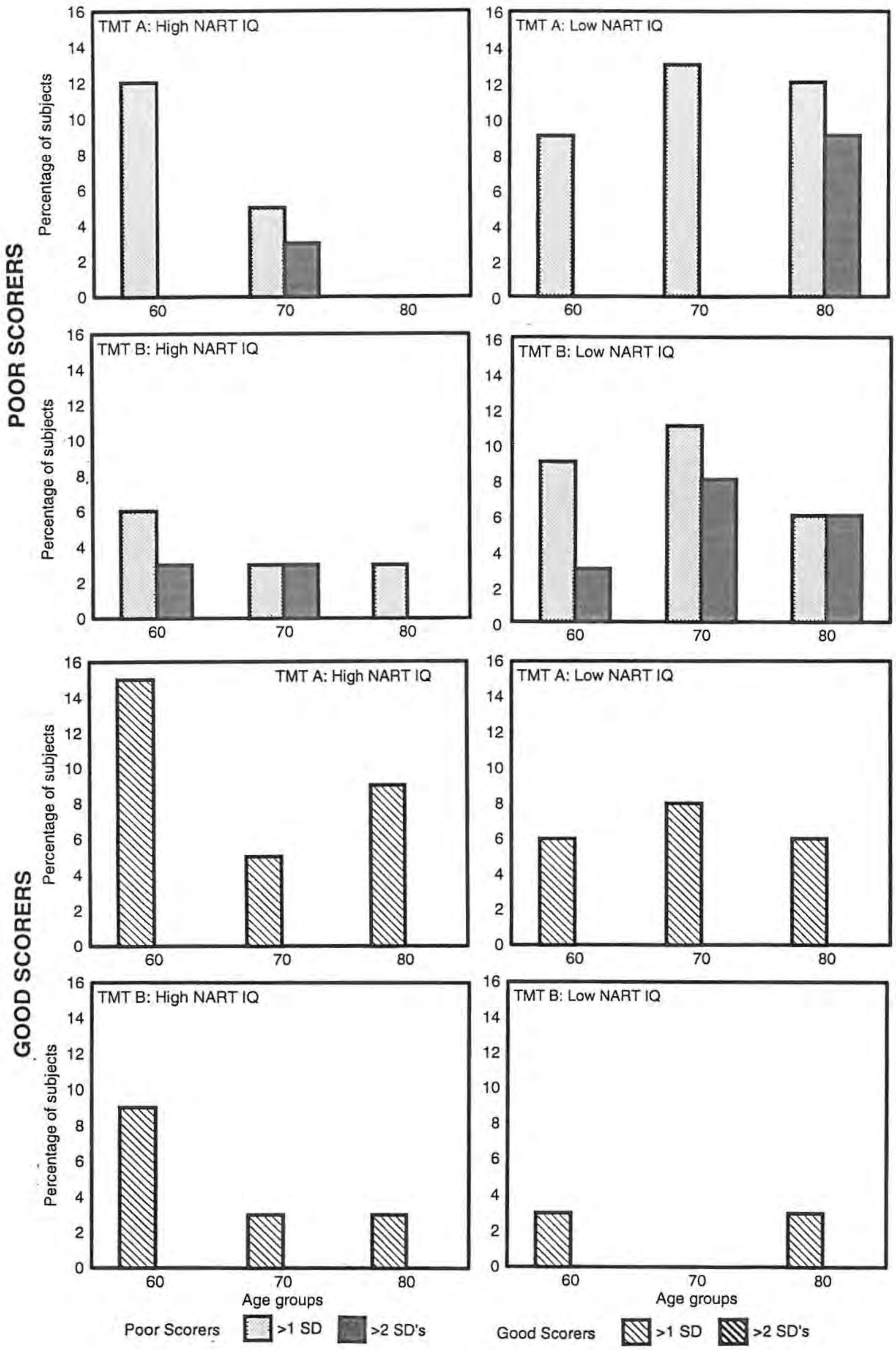


Figure 8.5. Trail Making Test Parts A and B: Percentage of extreme poor and good scorers per age group for low IQ (< 130) and high NART IQ (> 130).

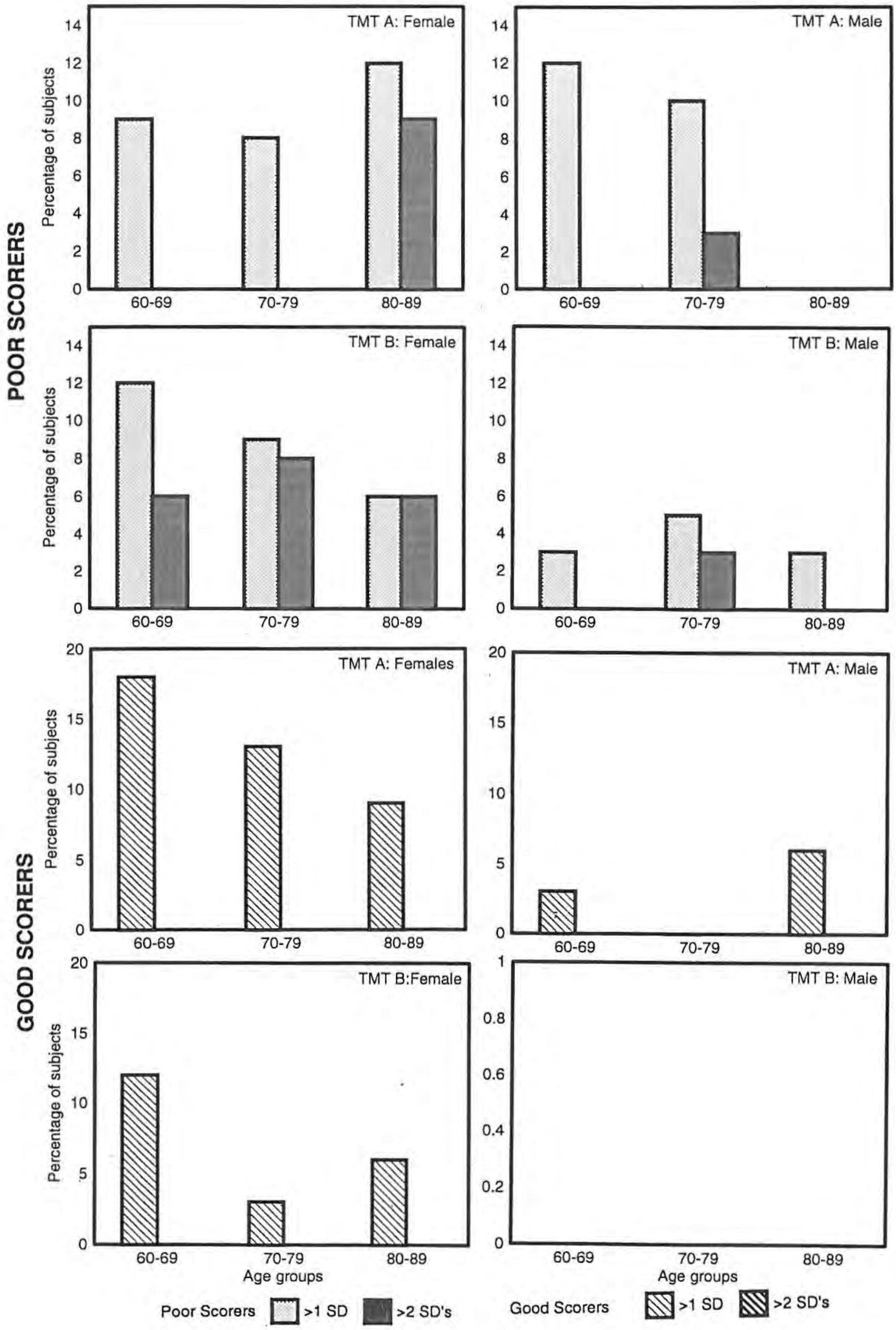


Figure 8.6. Trail Making Test Parts A and B: Percentage of extreme poor and good scorers per age group for Females and for Males.

numbers of males in comparative subsets in the prior within-groups analysis (section 8.3.1) did not effect the variability in the direction of low variability for males on a regular basis across age groups. Thus the male vulnerability factor identified here for the Trail Making Test appears to be sufficiently robust to suggest that it is probably a finding of relevance which goes beyond a mere sample artifact.

8.3.2.2 Digit Symbol and Incidental Recall

For *Digit Symbol* (see Histogram 8.3, p.368) it is apparent that there was significant overlap in the scores between the 60's age group and the younger 20-39 and 40-49 year-old groups, and that the significant widening of the distribution in the 60's was due to a general widening of the distribution including a number of disproportionately good scorers in relation to the mean, and in particular due to a number of disproportionately very poor scorers. In the 70's age group, for which there was still a significant widening of the variability, the overlapping of scores compared with the two younger age groups was still evident but there was a greater weighting of disproportionately poor scores rather than disproportionately good scores, although both were still in evidence. The significant reduction of variability in the 80's age stage was clearly apparent from the histogram which shows that there was minimal overlap with the scores of the youngest two age groups, and a close clustering of generally lowered scores.

Reading Table 8.18 horizontally for Digit Symbol, it is of note that the significant increase in variability in the 60's age group was a consequence of a proportion of disproportionately *poor* scorers who were characterized by both low and high education, low and high IQ and both levels of gender, and a proportion of particularly *good* scorers who were characterized by high education and female gender exclusively. In the 70's age group the identical pattern emerged with equal proportions of low and high education and levels of gender characterizing the disproportionately *poor* scorers, and a high proportion of extremely *good* scorers who were characterized by high education and female gender. This finding is consistent with the results of the statistical analysis for Digit Symbol in the previous subsection for which in the 60's age group there was a significantly higher standard deviation for the higher (14+years) educational condition compared with the lower (10-13 years) educational condition supported by a strong matching directional trend in the 70's age group. The overall indication was that the increased variability for Digit Symbol in the 60's and 70's age groups was strongly characterized by a spreading out of disproportionately good scorers who were of high education and female gender, indicating that these two factors in particular served a protective function in this test modality.

For *Incidental Recall* there was no significant variability effect, and nor were there any significant effects for education or gender on the within-groups statistical analyses of the previous subsection (8.3.1). The haphazard variability which occurred across all age groups was clearly evident from the distribution of scores (see Histogram 8.4, p.369). The analysis of outliers for Incidental Recall was included for completeness (see Table 8.18), but due to the lack of significant overall variability effects it was not further discussed.

8.3.2.3 <u>Finger Tapping Test</u>

For both Finger Tapping Preferred and Non-preferred hands (see Histograms 8.5a and 8.5b, pp.370-1) the progressive widening in the distribution from the 60's onwards was clearly apparent although it only approached significance in the overall analysis for the Preferred hand modality, and was only significant by the 80's age group for the Non-preferred hand modality. No doubt were the younger two age groups to be reduced in their age range to match the age groups from 60's onwards, in all probability there would be a narrowing effect on the variability in the younger years, and the widening in the later years would appear by contrast much more marked. Of particular note from the histogram was the overlapping of good scorers between the youngest and older age groups consistently for both hands *right up to the 80's age group*. In other words at every age level there were individuals capable of scoring at the peak level of the youngest age group for this timed hand motor test.

A horizontal reading of Table 8.19, purely in terms of extreme good scorers (those proportions of individuals falling >2SD's lower than the timed score mean), reveals that for the 60's, 70's and 80's age groups for both Finger Tapping Preferred and Non-preferred hands, there were no striking characterizations with respect to education or gender. In terms of education: For the Preferred hand, aside from a contribution to a small proportion of extremely poor scorers in the 60's age group, educational level did not appear to contribute specifically to the extreme scorers; for the Non-preferred hand high education did appear to be consistently associated with the extremely good scorers in each of the 60's, 70's and 80's, but it was also associated with a proportion of particularly poor scorers in the 80's age group. With respect to gender: In the 60's and 70's females appeared to characterize the good scorers, whereas in the 80's males had the advantage for both the Preferred and Non-preferred hand modalities. Thus overall for Finger Tapping it appeared that there was an exceptionally high degree of overlap of scores between younger and older age groups with the highest level of performance regularly in evidence across all age groups. However there was no clear indication that education or gender were making a specific contribution to the extremes of performance on either side of the distribution. There was a marginal indication that female gender and high education more strongly characterized the high scorers, although these factors did not entirely protect against the possibility of being a disproportionately low scorer. This lack of any clear characterization of outliers with respect to education and gender for Finger Tapping, is commensurate with the within-groups statistical analyses for Finger Tapping in the previous subsection, which revealed only negative results for within-groups comparisons between standard deviations for subsets of these variables. This was caused, it appears, by the neutralizing factor of *both* disproportionately high and disproportionately low scorers for each level of education and gender across age groups for the Finger Tapping test.

Table 8.17. Trail Making Test: Analysis of Extreme Poor and Good Scorers in terms of the Percentage of Subjects per Age Group that lie beyond ± 1SD, and beyond ± 2SD's (in parenthesis), for two levels of Education, NART IQ and Gender.

	POOR SCORERS						GOOD SCORERS					
	EDUCATION		NART IQ		GENDER		EDUCATION		NART IQ		GENDER	
	10-13yr	14+ yr	<130	130+	Fem	Male	10-13yr	14+yr	<130	130+	Fem	Male
TMT A												
60-69	12 (0)	9 (0)	9 (0)	12 (0)	9 (0)	12 (0)	6 (0)	15 (0)	6 (0)	15 (0)	18 (0)	3 (0)
70-79	5 (0)	13 (3)	13 (0)	5 (3)	8 (0)	10 (3)	5 (0)	8 (0)	8 (0)	5 (0)	13 (0)	0 (0)
80-89	9 (6)	3 (3)	12 (9)	0 (0)	12 (9)	0 (0)	0 (0)	15 (0)	6 (0)	9 (0)	9 (0)	6 (0)
TMTB												
60-69	0 (0)	15 (6)	9 (3)	6 (3)	12 (6)	3 (0)	0 (0)	12 (0)	3 (0)	9 (0)	12 (0)	0 (0)
70-79	3 (3)	11 (8)	11 (8)	3 (3)	9 (8)	5 (3)	0 (0)	3 (0)	0 (0)	3 (0)	3 (0)	0 (0)
80-89	6 (6)	3 (0)	6 (6)	3 (0)	6 (6)	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)	6 (0)	0 (0)

Table 8.18. Digit Symbol and Incidentall Recall: Analysis of Extreme Poor and Good Scorers in terms of the Percentage of Subjects per Age Group that lie beyond ± 1SD, and beyond ± 2SD's (in parenthesis), for two levels of Education and Gender.

	POOR SCORERS				GOOD SCORERS			
	EDUCATION		GENDER		EDUCATION		GENDER	
	10-13yr	14+yr	Fem	Male	10-13yr	14+yr	Fem	Male
Digit Symbol								
20-39	0 (0)	13 (4)	4 (4)	9 (0)	0 (0)	0 (0)	0 (0)	0 (0)
40-49	10 (0)	5 (0)	10 (0)	5 (0)	0 (0)	19 (0)	14 (0)	5 (0)
60-69	4 (0)	7 (0)	7 (0)	4 (0)	0 (0)	14 (7)	14 (7)	0 (0)
70-79	10 (0)	10 (0)	10 (0)	10 (0)	0 (0)	13 (7)	13 (7)	0 (0)
80-89	7 (0)	7 (0)	14 (0)	0 (0)	0 (0)	18 (4)	7 (0)	11 (4)
Inc Rec								
20-39	0 (0)	8 (4)	4 (0)	4 (4)	0 (0)	22 (0)	17 (0)	5 (0)
40-49	10 (0)	10 (0)	20 (0)	0 (0)	5 (0)	5 (0)	10 (0)	0 (0)
60-69	11 (0)	14 (0)	25 (0)	0 (0)	11 (0)	7 (0)	14 (0)	4 (0)
70-79	3 (0)	19 (0)	9 (0)	13 (0)	7 (0)	13 (0)	13 (0)	7 (0)
80-89	7 (7)	7 (7)	14 (14)	0 (0)	4 (0)	10 (0)	7 (4)	7 (0)

Table 8.19. Finger Tapping Preferred and Non-Preferred Hands: Analysis of Extreme Poor and Good Scorers in terms of the Percentage of Subjects per Age Group that lie beyond ± 1SD, and beyond ± 2SD's (in parenthesis), for two levels of Education and Gender.

	POOR SCORERS				GOOD SCORERS			
	EDUCATION		GENDER		EDUCATION		GENDER	
	10-13yr	14+yr	Fem	Male	10-13yr	14+yr	Fem	Male
Fing Tap Pref								
20-39	0 (0)	9 (4)	9 (4)	0 (0)	0 (0)	17 (0)	17 (0)	0 (0)
40-49	5 (0)	5 (5)	5 (0)	5 (5)	10 (0)	15 (0)	20 (0)	5 (0)
60-69	7 (4)	11 (0)	14 (0)	4 (4)	4 (0)	11 (0)	15 (4)	0 (0)
70-79	8 (0)	8 (0)	8 (0)	8 (0)	4 (0)	15 (0)	19 (4)	0 (0)
80-89	4 (0)	12 (0)	12 (0)	4 (0)	4 (0)	12 (0)	4 (0)	12 (4)
Fing Tap Non Pref								
20-39	0 (0)	13 (4)	9 (4)	4 (0)	0 (0)	22 (0)	13 (0)	9 (0)
40-49	5 (0)	10 (5)	5 (0)	10 (5)	0 (0)	19 (0)	14 (0)	5 (0)
60-69	11 (0)	15 (0)	22 (0)	4 (0)	0 (0)	11 (4)	11 (4)	0 (0)
70-79	4 (0)	12 (0)	8 (0)	8 (0)	4 (0)	12 (4)	16 (4)	0 (0)
80-89	0 (0)	16 (4)	12 (4)	4 (0)	4 (0)	8 (4)	4 (0)	8 (4)

Histogram 8.2a. Trail Making Test Part A: Distribution of Test Scores by Age.

```
         SIXTIES                       SEVENTIE                     EIGHTIES
MIDPOINTS.............................+............................+............................+
  84.000)
  80.500)                                                           *
  77.000)                                                           **
  73.500)                                                           *
  70.000)
  66.500)
  63.000)
  59.500)                                                           *
  56.000)                              *                            *
  52.500)                                                           *****
  49.000)                              ***                          M**
  45.500)                              ***                          ********
  42.000)****                          **                           ***
  38.500)***                           ******                       ***
  35.000)***                           M***                         ***
  31.500)*                             ***********                  **
  28.000)M*******                      *****                        *
  24.500)******                        ***                          -
  21.000)****
  17.500)****
         LEGEND FOR GROUP MEANS:     M - MEAN COINCIDES WITH AN ASTERISK
                                     N - MEAN DOES NOT COINCIDE WITH ANY ASTERISK
```

Histogram 8.2b. Trail Making Test Part B: Distribution of Test Scores by Age.

```
         SIXTIES                       SEVENTIE                      EIGHTIES
MIDPOINTS.............................+.............................+.............................+
 300.000)
 285.000)
 270.000)                                                            *
 255.000)                                                            *
 240.000)
 225.000)
 210.000)
 195.000)                                                            *
 180.000)                              **                            *
 165.000)                              **                            **
 150.000)                              *                             ******
 135.000)                                                            M
 120.000)**                            *                             *****
 105.000)***                           ******                        **********
  90.000)**                            M*****                        ****
  75.000)M*****                        **                            **
  60.000)****************              ************
  45.000)****                          ******
  30.000)
  15.000)
         LEGEND FOR GROUP MEANS:     M - MEAN COINCIDES WITH AN ASTERISK
                                     N - MEAN DOES NOT COINCIDE WITH ANY ASTERISK
```

Histogram 8.3. Digit Symbol: Distribution of Test Scores by Age.

```
         TWENTHIR            FORFIFT             SIXTIES             SEVENTIE            EIGHTIES
MIDPOINTS....................+...................+...................+...................+...................+
  70.000)
  67.500)****                *                   **
  65.000)***                 *                   *
  62.500)****                *
  60.000)                    **                  *
  57.500)M***                **                                      *
  55.000)****                **                                      *
  52.500)*                   M                   *
  50.000)                    ***                 **                  *
  47.500)                    *                                       *
  45.000)*                   ****                ****                *                   *
  42.500)*                   *                   M****               **
  40.000)                    **                  ***                 *****               **
  37.500)                                        **                  ****                **
  35.000)                                        **                  M**                 *****
  32.500)*                                       **                  *                   M*****
  30.000)                                        **                  **                  ***
  27.500)                                                            ***                 ****
  25.000)                                        *                   **                  **
  22.500)                                                            *                   **
  20.000)                                                            ***                 *
         LEGEND FOR GROUP MEANS:     M - MEAN COINCIDES WITH AN ASTERISK
                                     N - MEAN DOES NOT COINCIDE WITH ANY ASTERISK
```

Histogram 8.4. Incidental Recall: Distribution of Test Scores by Age.

```
          TWENTHIR            FORFIFT             SIXTIES             SEVENTIE            EIGHTIES
MIDPOINTS....................+...................+...................+...................+...................+
   10.000)
    9.500)
    9.000)*****               *                   *****
    8.500)***
    8.000)***                 *                   ****                                    *
    7.500)M
    7.000)***                 *****               ***
    6.500)**                  *                                                           *
    6.000)****                *                   M******             ***                 **
    5.500)                    M*                                      *
    5.000)*                   *                   **                  **                  ***
    4.500)                    **                                      *
    4.000)                    ***                 ***                 **********          *****
    3.500)                    *                   *                   M
    3.000)*                   **                  *                   ****                M***
    2.500)                                                            **                  *
    2.000)                    *                   **                  ***                 ***
    1.500)                                                                                **
    1.000)                                                            ****                **
    0.500)
    0.000)                                                                                ****
          LEGEND FOR GROUP MEANS:     M - MEAN COINCIDES WITH AN ASTERISK
                                      N - MEAN DOES NOT COINCIDE WITH ANY ASTERISK
```

Histogram 8.5a. Finger Tapping Preferred Hand: Distribution of Test Scores by Age.

```
         TWENTHIR             FORFIFT             SIXTIES             SEVENTIE            EIGHTIES
MIDPOINTS....................+...................+...................+...................+...................+
   9.000)                                                                                 *
   8.700)                                         *                                       *
   8.400)                                         *                   *                   **
   8.100)                                                             ***
   7.800)*                    *                   *                   ***                 **
   7.500)                                         *                   ***                 **
   7.200)                                         **                                      ***
   6.900)                     *                                       **                  ***
   6.600)                                         ***                 M*                  M*
   6.300)*                    ****                *                   ***                 *
   6.000)**                   **                  M****               *                   ***
   5.700)*****                *                   *                   *                   *
   5.400)M***                 M*                  ****                **                  **
   5.100)*                    ***                 *                                       *
   4.800)*                    **                  **                  **
   4.500)******               ****                **                  *
   4.200)*                                                                                *
   3.900)*                    *                                       *
   3.600)                                         *                                       *
   3.300)                                                             *
   3.000)                                         *
         LEGEND FOR GROUP MEANS:      M - MEAN COINCIDES WITH AN ASTERISK
                                      N - MEAN DOES NOT COINCIDE WITH ANY ASTERISK
```

Histogram 8.5b. Finger Tapping Non-Preferred Hand: Distribution of Test Scores by Age.

```
         TWENTHIR             FORFIFT             SIXTIES             SEVENTIE            EIGHTIES
MIDPOINTS....................+...................+...................+...................+...................+
  10.150)
   9.800)                                                                                 *
   9.450)
   9.100)                                                                                 **
   8.750)                                                             **
   8.400)                                         *                                       *
   8.050)                                         *                   **                  ****
   7.700)                                         **                  **                  *
   7.350)                     *                   ***                 **                  ***
   7.000)*                                                            *****               M*
   6.650)                     *                   **                  M***                ***
   6.300)**                   *                                       *                   *
   5.950)*                    ***                 M****               *                   ***
   5.600)*******              *****               *****               **                  ***
   5.250)M***                 M***                                    **
   4.900)**                   *                   *****               *
   4.550)***                  **                  *
   4.200)*                    *                                       *                   *
   3.850)**                   *                   *                                       *
   3.500)                     *                   *
   3.150)                                                             *
         LEGEND FOR GROUP MEANS:      M - MEAN COINCIDES WITH AN ASTERISK
                                      N - MEAN DOES NOT COINCIDE WITH ANY ASTERISK
```

8.3.3 Correlation Coefficients

For the *Trail Making Test* (see Table 8.20) as to be expected there was a significant positive correlation consistently between the two TMT modalities (TMT A and TMT B), and between education and NART IQ. Aside from this there was only a significant negative correlation between education and TMT A and TMT B for the 80's age group which was not present for any other age group. A similar trend was evident for NART IQ. This implies a more powerful effect for education and IQ in the oldest age group which is consistent with the finding from the within-groups analyses in the previous subsections (8.1 and 8.3.2), of an effect for education and IQ for the TMT which was most marked in the 80's age group. In contrast, for the rest of the tests (*Digit Symbol, Incidental Recall,* and *Finger Tapping*) (see Table 8.21) there was no evidence of a significant correlation with education for any of the age groups. This is consistent with the finding for the within-groups analyses in the previous two subsections which revealed the strongest effect for education with the Trail Making test.

As to be expected there was a positive correlation between the Preferred and Non-preferred hands submodalities of the Finger Tapping test. An additional consistent finding was a significant negative correlation for Digit Symbol and the Finger Tapping test which, for the Preferred hand, occurred consistently across all age groups until the 80's age group at which point it was not in evidence. For the Non-preferred hand the identical picture was evident, although a strong trend in the 20-39 year-old group did not reach significance. This effect is of particular interest considering the significant *increase* in variability of performance which was in evidence for the Finger Tapping test, which was due in part to a significant proportion of individuals with minimal loss of hand motor function by the 80's age group. This stands in contrast to the significant *reduction* in variability for the Digit Symbol test in the 80's indicating that there were no individuals who were spared an aging effect by this late age stage for this visuo-perceptual task. Thus overall the finding suggests that good hand motor speed contributed to variability in Digit Symbol performance by providing a protective factor until the 70's age stage, and hence significant correlations between the two tasks up to this stage. However in the 80's the factor of more pervasively impaired visuo-perceptual scanning due to aging for all individuals overrode any protective advantage afforded by sustained skill in hand motor function, hence resulting in the absence of a significant correlation between the two functions at this point.

Table 8.20. Correlation Coefficients per Age Group: Trail Making Test A and B, Education and NART IQ.

AGE GROUP	TMT A	TMT B	Education	NART IQ
60-79				
TMT A	1.000	-	-	-
TMT B	0.3423	1.000	-	-
Education	0.1162	0.0769	1.000	-
NART IQ	-0.0291	-0.0453	0.5735**	1.000
70-79				
TMT A	1.000	-	-	-
TMT B	0.3661*	1.000	-	-
Education	0.3042	0.0727	1.000	-
NART IQ	0.0555	-0.2394	0.4048*	1.000
80-89				
TMT A	1.000	-	-	-
TMT B	0.4234*	1.000	-	-
Education	-0.4002*	-0.3756*	1.000	-
NART IQ	-0.3330	-0.2284	0.5864*	1.000

Significant correlation (*$p < 0.05$; **$p < 0.01$).

Table. 8.21. Correlation Coefficients per Age Group: Digit Symbol, Incidental Recall, Finger Tapping Preferred and Non-Preferred Hands, and Education.

AGE GROUP	Digit Sym	Inc Rec	Fing Pref	Fing Non-Pref	Education
20-39					
Digit Sym	1.000	-	-	-	-
Inc Rec	0.3475	1.000	-	-	-
Fing Pref	-0.5372**	-0.0972	1.000	-	-
Fing Non-Pref	-0.3453	-0.0040	0.8106**	1.000	-
Education	-0.1529	-0.1220	0.1197	-0.1230	1.000
40-59					
Digit Sym	1.000	-	-	-	-
Inc Rec	0.3214	1.000	-	-	-
Fing Pref	-0.6429**	-0.2393	1.000	-	-
Fing Non-Pref	-0.5674**	-0.4311	0.8129**	1.000	-
Education	0.1342	0.0895	0.2943	0.168	1.000
60-69					
Digit Sym	1.000	-	-	-	-
Inc Rec	0.2298	1.000	-	-	-
Fing Pref	-0.4491*	-0.4486*	1.000	-	-
Fing Non-Pref	-0.4548*	-0.4066*	0.9296**	1.000	-
Education	0.2849	0.0433	-0.3082	-0.3408	1.000

Table 8.21 *continued*

Table. 8.21. Correlation Coefficients per Age Group: Digit Symbol, Incidental Recall, Finger Tapping Preferred and Non-Preferred Hands, and Education *continued.*

AGE GROUP	Digit Sym	Inc Rec	Fing Pref	Fing Non-Pref	Education
70-79					
Digit Symbol	1.000	-	-	-	-
Inc Rec	0.4107*	1.000	-	-	-
Fing Pref	-0.5852**	0.0114	1.000	-	-
Fing Non-Pref	-0.5820**	0.0603	0.9468**	1.000	-
Education	0.619	-0.1599	-0.2444	-0.2139	1.000
80-89					
Digit Sym	1.000	-	-	-	-
Inc Rec	0.4617*	1.000	-	-	-
Fing Pref	-0.2634	0.0132	1.000	-	-
Fing Non-Pref	-0.2775	-0.0680	0.9324**	1.000	-
Education	0.1648	0.1745	0.0251	0.0102	1.000

Significant correlation (*p<0.05; **p<0.01).

8.3.4 Summary of within-groups effects

From the small number of tests submitted to analysis for Phase 2 of the investigation it was clear that within-groups variability effects were highly task specific. Generally where variability increases in association with older age this was due to widening of the distribution at *both* the upper and lower scoring extremities, and was characterized by both disproportionately good and poor scorers. For the Trail Making Test and Digit Symbol, however, the increased variability was characterized in particular by a relatively large proportion of low scorers, whereas for the Finger Tapping test there was a larger proportion of exceptionally good scorers who retained ability at the level of the youngest age groups right into late old age.

Most generally over all the tests which revealed significant alterations in variability, *outlying good scorers tended most predominantly to be female and to have high education.* Thus without the protective factors of female gender and high education an individual was unlikely to be one of the exceptionally successful ages on this set of cognitive tasks in later old age. Furthermore there was evidence for the possible presence of a task-related protective function afforded by hand motor ability in association with the Digit Symbol task which no longer played a part by the later 80's age groups.

Importantly there were a number of indications from the analyses that the increase in distribution in association with aging was not merely a uniform effect across the adult life span. First, whilst there was the indication that high education, high IQ and female gender were *necessary* to retain a high level of performance into late old age, *none* of the factors of high education, high IQ or female gender precluded a significant proportion of individuals, and especially females, from being one of the extremely poor scorers. In particular there was the indication of a complex bimodal effect for female gender membership. For some individuals female gender membership provided a protective function, and for others it constituted a vulnerability factor. Further, the protective effect of high education was differentially present for tasks in that it was most marked for the Trail Making Test (for which high IQ also served a protective effect), was definitively present in addition for the Digit Symbol Test, but was barely apparent for the Finger Tapping Test. For the Trail Making Test high education in particular, but also female gender, revealed increasing protective effects as age advanced.

Thus the widening distributions in association with older age *did not appear simply to describe the widening out of the upper and lower percentiles of cognitive ability across the adult life span.* Rather the within-groups analyses conducted here cumulatively lend strong support to the presence of *threshold* effects in the presentation of cognitive decline due to differential combinations of levels of education, IQ and gender, and possibly also the contribution of particular contributory skills such as hand motor dexterity, in association with the advance of neural attrition due to age. Importantly, since the factors of education, IQ and gender did not invariably provide protective effects, the indication from these analyses is that there are additional powerful factors which came into play (which were not available for investigation in the present study) which overrode the protective effects of high education, high IQ and female gender, and contributed to the lowering of threshold effects.

8.4 OVERALL SUMMARY OF EFFECTS

Thus in sum, the findings from the Phase 2 between and within-groups analyses were as follows:

(i) From the Phase 2 *between-groups* analyses there was confirmation of the shuttle bulge effect in that there was the robust occurrence of a significant increase in variability, with or without a subsequent significant decrease depending on the task and functional domain, in association with the progression of older age, which occurred in parallel with a significant decline in average performance. Further there was support for the shuttle shift effect in that the higher task challenge of Trail Making Test Part B revealed the earlier onset of variability (left shuttle effect), and conversely the lower task challenge of Trail Making Test Part A revealed the later onset of variability (right shuttle effect).

(ii) From the Phase 2 *within-groups* analyses there was support for the fact that high education, high IQ and female gender provided a significant protective factor which raised the threshold for symptom onset due to aging, although this was not invariable indicating that other powerful factors probably came into play which overrode these protective effects. Conversely there was support for the fact that low education, low IQ and male gender were vulnerability factors which almost invariably lowered the threshold for the onset of cognitive deterioration due to aging. Both of these lowered and raised threshold effects were progressively more in evidence with encroaching age which served to explain the widening distribution of disproportionately good and poor scorers, and hence the significant increases in formal measures of variability in cognitive test performance which occurred in association with older age.

Altogether the above Phase 2 between and within-groups findings of a significant increase in variability in association with aging (and task dependent subsequent decrease), and a left versus right shuttle effect for high versus low levels of task challenge, were entirely commensurate with the cumulative evidence from the Phase 1 analyses. Further there was a striking consistency between the Phase 2 within-groups characterisation of outliers which revealed low education, low IQ and male gender as vulnerability factors which lowered the threshold for symptom presentation, and the Phase 1 indications of an *earlier* onset of variability in cognitive test performance (that is a *left* shuttle effect) for lower levels of education, and male gender. Conversely there was notable consistency between the Phase 2 characterisation of outliers which generally revealed high education, high IQ and female gender as protective factors, and the indications from the Phase 1 between-groups analyses of the *later* onset of variability (that is a *right* shuttle effect) for lower levels of task challenge, higher levels of education, and female gender.

9

CHAPTER 9: DISCUSSION

9.1 AIMS OF THE STUDY AND STATISTICAL HYPOTHESES

Broadly this research study set out to describe the nature of inter-individual variability in cognitive aging as it occurs across the adult life span. Prior to the present research there was no systematic formulation of cognitive heterogeneity in association with aging. As discussed in the initial chapters of the thesis (sections 1.2 and 4.1), uncertainty about the phenomenon prevailed to the extent that some researchers in the area strongly challenged the possibility of enhanced variability at all, as a feature of the aging process (Bornstein & Smircina, 1982; Montgomery and Borgatta, 1986; Salthouse, 1991); and such commentary occurred in spite of a significant body of literature which strongly emphasized the importance of taking heed of cognitive heterogeneity in older populations. In particular the work of Rabbitt and Schaie were noted in this regard (Nettelbeck & Rabbitt, 1992; Rabbitt, 1991; 1983; 1990; 1993a; 1993b; 1993c; Rabbitt et al., 1993; Schaie, 1983; 1985; 1988a; 1988b; 1990; 1994; Schaie & Willis, 1986). However, on the basis of a systematic review and collation of existing studies which have focused on variability in the cognitive aging literature (which had not previously been conducted), it was already possible to refute the view that the phenomenon of heterogeneity in older populations is just one of those additional mythical stereotypes that surrounds aging (see section 4.2.1). Furthermore, on the basis of a synthesis of the existing research, it was possible to pose a specific model - the shuttle model - of cognitive variability (sections 5.2 and 5.3). The model formally encapsulated the changing manner in which cognitive variability presents in relation to the adult age axis both in terms of its shape (the shuttle bulge) and its shifting mechanism (left versus right) depending upon a number of influences.

Whilst this bottom-up (see Table 5.1, p. 147) empirically-based shuttle model, served primarily to *describe* the pattern of variability with aging, an *explanatory* level of operation for the model was derived top-down (again see Table 5.1, p.147), from within the brain-behaviour concepts of a neuropsychological framework, and from causal mechanisms implied by brain reserve capacity (BRC) threshold theory as formulated by Satz (1993). In terms of the top-down theoretical standpoint, the implications for cognitive variability in association with aging were elaborated into a set of hypotheses for the purposes of this study (section 3.2.3.3). In terms of the bottom-up empirical data angle (as derived from the literature review), these hypotheses were developed into a set of more specific hypotheses for an investigation into patterns of cognitive variability, and for the testing of the shuttle model (section 5.3.4). These hypotheses in turn were operationalized into a set of statistical hypotheses pertinent to the present two-phase empirical investigation into variability (sections 6.3.1, and 6.3.2). The statistical hypotheses were as follows:

Phase 1 and Phase 2: Statistical hypotheses for the between-groups analyses of standard deviations and means.

1. It was hypothesized that there would be an overall significant difference in the standard deviations across adult age groups for the majority of neuropsychological tests included in both phases of the analysis, and that if this did not occur it would be for those tasks which have a relatively high crystallized verbal component and/ or present relatively low task challenge, and/ or have been shown on central trends analysis not to be particularly age-sensitive.

2. It was hypothesized that the significant differences in variability would be due to one of two distinctive patterns across the middle to late adult years in association with a decline in average cognitive performance: either there would be a non-linear (inverted-U) effect with an initial increase in variability followed by a subsequent decrease in variability in the later adult years; or there would be a linear effect of progressively increasing variability across the middle to late adult years. More specifically it was hypothesized that this would occur as follows:

(i) Should the Bartlett's test or Levene's F test reveal an overall significant difference in the standard deviations between all age groups for a data set with a non-linear inverted-U trend, backwards multiple pairwise comparisons of standard deviations from the oldest group with each younger age group, would reveal significantly lower standard deviations for the oldest age group (or groups) compared with the middle age group (or groups), and no significant difference compared with the youngest age group (or groups); forwards comparisons from the youngest age group with each older age group would reveal significantly lower standard deviations for the youngest age group (or groups) compared with the middle age group (or groups), and no significant difference compared with the oldest group (or groups). Simultaneously the pairwise comparison of the mean cognitive score between the oldest and youngest age group would reveal a significantly lower mean score (where a low score represents poor functioning) when compared with the older group, or a significantly higher mean score when compared with the older group (where a high score represents good functioning).

(ii) Should the Bartlett's statistic or Levene's F statistic reveal an overall significant difference in the standard deviations between age groups for a data set with a linear trend, backwards multiple pairwise comparisons of standard deviations from the oldest age group with each younger age group, would reveal significantly higher standard deviations for the oldest group (or groups) compared with the middle age group (or groups), and the youngest age group (or groups); forwards comparisons from the youngest age group with each older age group, would reveal significantly lower standard deviations for the youngest age group (or groups) compared with the middle age group (or groups) and/or the oldest age group (or groups). Simultaneously the pairwise comparison of the mean cognitive score between the oldest and youngest age group would reveal a significantly lower mean score (where a low score represents poor functioning) when compared with the older group, or a significantly higher mean score when compared to the older group (where a high score represents poor functioning).

3. It was hypothesized that the linear (rather than non-linear) variability trend would be more likely to occur for samples with a high educational/IQ level than for samples with a low educational/ IQ level, and for those tasks which have a high crystallized/verbal component, and/or present low challenge, and/or are not particularly age-sensitive, than for tasks which have a high fluid/ perceptual speed component, and/or present high challenge and/or are particularly age-sensitive.

4. It was hypothesized that for the same neuropsychological task, a significant increase in standard deviations would occur at a relatively earlier age stage, and a subsequent significant decrease in standard deviations would occur at a relatively earlier age stage for samples with lower levels of education, for tasks which present greater challenge, for tasks which are more age-sensitive and for male gender. Conversely it was hypothesized that a significant increase in standard deviations would occur at a relatively later age stage for samples with higher levels of education, lower levels of task challenge, tasks which are less age-sensitive, and female gender such that a subsequent significant decrease may be apparent at a later age stage, or may not appear at all even for the latest age stage.

Phase 2: Statistical hypotheses for the within age-groups comparisons of standard deviations, the characteristics of outliers, and the correlation coefficients.

1. It was hypothesized that there would be an effect for education and IQ (as revealed in the comparisons of standard deviations for different levels of education and IQ within age groups, and in the positive correlation coefficients for education/IQ and cognitive test performance across age groups), that would be more pronounced for the age groups with significantly increased variability across the adult age range than for those age groups which did not show increased variability, that would be more pronounced for those tasks which showed a prolonged progressive increase in variability in later old age compared with those which showed an initial increase followed by a subsequent decrease in variability in later old age, and that would be more pronounced for those tasks which call upon educationally-based/ crystallized/ reasoning skills to a greater extent than speeded perceptual/ fluid type skills.

2. It was hypothesized that where levels of education/IQ have a significant effect, the effect could be revealed in any one of the following forms for a particular age group:

(i) The standard deviation for low education/IQ may be significantly greater than that for high education/IQ due to a high proportion of individuals with low education who achieve disproportionately low scores which account for the significant increase in the overall distribution, and this effect would be reflected in a significant positive correlation between education/IQ and performance for that age group;

(ii) The standard deviation for high education/IQ may be significantly greater than that for low education/IQ due to a relatively large proportion of individuals with high education who achieve disproportionately high scores which account for the significant increase in the distribution, and this effect will be reflected in a significant positive correlation between education/IQ and performance for that age group;

(iii) There may be no difference between the standard deviations for low and high education because a relatively large number of disproportionately low scorers with low education/IQ, together with a relatively large number of disproportionately high scorers with high education/IQ, account for the significant increase in the distribution, but the effect will be revealed in a significant positive correlation between education/IQ and performance.

3. It was hypothesized that there might be differences between the standard deviations within age-groups for gender, and associated gender differences in the characterization of disproportionately good and poor scorers that account for significant increases in variability in the late adult years, and that the effect could be revealed in any one of the following forms for a particular age group:

(i) The standard deviation for males may be significantly greater than that for females due to a high proportion of males who achieve disproportionately low scores which account for the significant increase in the overall distribution;

(ii) The standard deviation for females may be significantly greater than that for males due to a relatively large proportion of females who achieve disproportionately high scores which account for the significant increase in the distribution;

(iii) There may be no difference between the standard deviations between males and females because a relatively large proportion of disproportionately low scoring males, together with a relatively large proportion of disproportionately high scoring females, contribute to the significant increase in the distribution.

9.2 CONFIRMATION OF THE STATISTICAL HYPOTHESES

9.2.1 Phase 1 and Phase 2 between-groups analyses. Taken together, the results from both Phase 1 and Phase 2 of the present study provided substantial support for the four statistical hypotheses of the between-groups analyses in that (i) the majority of neuropsychological tests included in both phases of the study did reveal a significant effect between standard deviations across all groups for each of the domains of attention, verbal and visual memory, language, visual and hand motor skills; (ii) significant effects for variability did occur as a consequence of two highly distinctive patterns (of either a non-linear effect of a significant initial increase followed by a subsequent significant decrease in variability, or a linear effect of a progressive and sustained significant increase in variability) across the middle to late adult years in association with a significant decline in average ability, as confirmed by the multiple pairwise comparisons which repeatedly revealed either a peak in standard deviations in the middle years followed by a significant decline in standard deviations in the later years, or merely a sustained increase in standard deviations between the early and later adult years, in each case together with a significantly lower or higher mean between the youngest and oldest age groups in the direction which indicated decrements in cognitive ability with age; (iii) there were a number of instances where there was a significant non-linear inverted-U pattern present for a task of greater challenge together with a linear pattern of sustained increasing variability present for a comparable task of less challenge, and similarly there were a number of instances where there was a significant non-linear inverted-U pattern present for more age-sensitive task together with a linear pattern of sustained increasing variability present for a comparable less age-sensitive task; (iv) the initial significant increase and subsequent significant decrease in standard deviations did consistently occur at a relatively *earlier* age stage for samples with lower levels of education, for tasks which present greater challenge, for more age-sensitive tasks and for male gender, and conversely the initial significant increase and subsequent decrease in standard deviations did occur at a relatively *later* age stage for samples with higher levels of education, for tasks which present less challenge, for less age-sensitive tasks and for female gender. Thus none of these four hypotheses were fundamentally refuted. Some additional unanticipated occurrences, however, did became apparent.

With respect to the shuttle bulge effect, there were occasions where there was a significant effect for variability across all groups which revealed neither a non-linear inverted-U pattern, nor a clearly linear pattern, but rather a haphazard pattern across the adult years. Consistently, however, this occurred in association with the lack of clearly declining average ability in association with age. This in itself, however, provided support for the model which poses that the variability pattern will occur in association with declining average ability (see as schematized in Figures 5.2 and 5.6, pp.153 and 163, respectively). It was also apparent that where there was a great deal of inherent variability within a particular task regardless of age, this resulted in the absence of clear-cut variability effects.

Thus overall the results indicated that variability effects occurred as lawfully and robustly as the well-documented effect of declining average ability in association with age, and were similarly task and domain specific. The indication was that attention and short term memory tasks (both verbal and visual, immediate and delayed) tended consistently to produce the non-linear pattern of an initial increase followed by a subsequent significant decrease in variability, whereas linear patterns occurred in the main for recognition tasks, language and problem solving tasks. As predicted, the classically more age-sensitive attention and short term memory tasks (both verbal and visual) and a relatively more age sensitive fluid visuo-perceptual task (specifically the timed visuo-perceptual Digit Symbol task), presented the full inverted-U pattern including an initial significant increase and subsequent significant decrease in variability, whereas the more crystallized/educationally based tasks (all the language tasks) consistently produced the linear effect of progressively increasing variability. For relatively pure visuo-spatial tasks *without* a speed component which are not particularly age-sensitive, there was a consistent absence of clear variability effects. Similarly, for recognition tasks, which compared with short term memory tasks are not particularly age-sensitive, there were either linear effects of progressively increasing variability (as in the verbal recognition tasks) or the absence of any clear effects (specifically the RVDLT visual recognition task).

Of additional note was the consistent linear pattern of significantly increased variability with no subsequent decrease in association with older age for problem solving ability which was clearly apparent for all the relatively pure problem solving tasks. Moreover problem solving appeared to produce a particularly potent protective effect in that where a visuo-perceptual scanning task called upon problem solving (in the form of rational ordering) to a greater extent than a visuo perceptual task that did not (that is the Trail Making test compared with the Digit Symbol task, respectively), the visuo-perceptual task *with* the problem solving component (the Trail Making Test) appeared consistently across a number of studies to present a linear pattern of increasing variability without a subsequent decrease into the 80's age stage (see results sections 7.5 and 8.2). In contrast, the Digit Symbol task (a visuo-perceptual task *without* a rational ordering component) revealed a non-linear pattern of variability consistently across different studies (see Digit Symbol data sets reported in Salthouse, 1991, Figure 4.4, p.100; Shuttleworth-Jordan & Bode data set, Figure 8b, p.115), with a significant increase in variability followed by a subsequent significant decrease in variability in the later adult years. Although compared with the Digit Symbol test, the virtually identical Symbol Digit Modalities task revealed an apparently contrasting linear pattern of increasing variability (see Figure 7.25, p.285), the latter data were only presented up to the 70's age stage. Hence there is no fundamental contradiction here, but rather the indication that were older age groups tapped the full non-linear inverted-U pattern for Symbol Digit Modalities (as was apparent for the Digit Symbol test) would most likely become apparent.

Thus overall for the between-groups empirical investigation there was a high level of consistency for the same task across different data sets, and within modalities across different tasks, which in turn consistently supported the statistical hypotheses. Inconsistencies for particular tasks between studies were probably no more than would be revealed due to error variance and sample bias for a large collection of data sets depicting *average* effects in association with aging. Moreover, in most instances it was possible to explain inconsistencies in variability in terms of probable sample effects such as for example unmatched groups for IQ or education. The presence of a number of significant *erratic* patterns of variability were consistent in their occurrence together with the *absence* of clearly declining average cognitive ability. The only set of results which in any way challenged the concept of a fundamental single bulge in variability in association with older age across the comprehensive set of functional modalities tapped, were a cluster of tests for verbal memory which revealed a prior mini-bulge in variability around the 30's age stage. However, this also did not occur in association with the steady indication of declining average ability at that stage, and could possibly be understood as the release of a ceiling effect for these tests. That an unusual effect may be present specifically for verbal memory at this early adult age stage is supported by norms for the AVLT reported by Query and Megran (1983). These authors note an idiosyncratic significant *increase* for learning in the 35-39 year old age group.

When comparing the overall between-groups outcome of the present empirical investigation with the synthesis of variability trends extracted from the literature review and as catalogued in Table 4.1 (p.121), the variability trends for attention and memory (both verbal and visual) are highly commensurate. For both the review and the empirical analyses, an overall non-linear pattern was in evidence with an initial increase and subsequent significant decrease in variability for *most* data sets reported or examined in these two modalities. For the modalities of Language and Problem Solving, however, this was not the case. For the empirical investigation *all* tasks in these latter two modalities revealed a sustained pattern of increasing variability with age. From the review it was noted that Word Naming and the

Wisconsin Card Sorting Test (for the Language and Problem Solving modalities, respectively) reportedly revealed the linear effect of a sustained increased variability consistent with the present investigation; on the other hand, Verbal Fluency and Similarities (also for the Language and Problem Solving modalities, respectively) revealed an initial increase and subsequent decrease in variability for tasks in these modalities.

The explanation for these discrepancies for problem solving, between the review and the present study, may be that the data included in the empirical investigation did not tap old enough age groups for the tests in question, thus not revealing a decrease which would occur for older groups. Alternatively, the data sets in the present investigation may all have been constituted from higher functioning samples than those gleaned from the review, and hence, due to a right shift in the variability curve, only a sustained linear increase in variability was in evidence. In any event, although the dedifferentiation effect was not in evidence at all for the Language and Problem Solving modalities in the empirical investigation, evidence for the possibility of the more severe full non-linear inverted-U effect in these modalities from the review, indicates that given sufficient vulnerability of the cohort due either to very old age and/or lower functioning cohorts, most language and problem solving tasks *would* reveal a subsequent dedifferentiation effect. That this is not invariably in evidence from all the studies reported makes sense in terms of the relatively less age-sensitive position given to reasoning and language-related tasks in the cascade model as described in Birren and Cunningham (1985) (see discussion in section 3.2.3.1). In this model the three cognitive constructs of 'perceptual speed', 'reasoning', and 'verbal comprehension' are posed, in that order, to be progressively less sensitive to aging.

Thus another source of *apparent* inconsistencies when reviewing all the findings in *absolute* terms across both the literature review and the empirical investigation, were differences which could be accounted for by shuttle effects consequent on task-related and demographic influences, or that could be explained on the basis of overly narrow study-window effects. In view of this shifting nature of the variability patterns due to multiple demographic, task-related and study-window influences, such that a full non-linear inverted-U pattern did or did not appear, it was only possible to make categorical statements about the particular shape of variability patterns across tasks in fairly gross terms. As is very apparent from the current discussion, the attempt to integrate apparent differences or dissociations in effects in absolute terms becomes extremely contorted. Thus the outcome from the between-groups analyses indicates, as already noted on the basis of the literature review in Chapter 5 (section 5.1), that to tabulate variability effects as attempted in Table 4.1, (p.121), is not a satisfactory means of depicting variability effects. Rather a more flexible *process* model is required for an adequate description of the regularly occurring variability effects which were confirmed on the basis of the Phase 1 between-groups study.

9.2.2 Phase 2 within-groups analyses. Broadly, the first statistical hypothesis for the Phase 2 within-groups study was supported by the results as follows. There was an effect for education and IQ (as revealed in the comparisons of standard deviations for different levels of education and IQ within age groups, and in the positive correlation coefficients for education and cognitive test performance across age groups), which was more pronounced for age groups with significantly increased variability across the adult age range than for those age groups which did not show increased variability, in that generally where education and IQ effects occurred (specifically for the Trail Making Test and also for the Digit Symbol task) this was in the oldest age group or groups for which there was enhanced variability. Furthermore the education and IQ effects were more pronounced for tasks which call upon educationally based skills to a greater extent than pure speeded fluid/visuo-perceptual type skills, in that again the effect occurred more frequently and was more obviously present for the Trail Making Test. (This is a task which arguably calls upon educationally-based rational ordering more than is called upon by the other tasks included in the study of Incidental Recall, Digit Symbol and Hand Motor function).

In very broad terms the second and third statistical hypotheses were also supported in that the significant effects that occurred for education and IQ, and in addition for gender, were explicable in terms of proportions of outliers with specific characteristics in terms of these variables, and were reflected in some instances in the correlation coefficients. However, whilst it was anticipated in these hypotheses that the effects could present in a variety of forms, the manner in which these effects actually occurred was even more variable for different influences between the tests, and even more subtle, than could have been anticipated from previous research and reflected in the original hypotheses. Generally, where variability increased in association with older age, this was due to the widening of the distribution at *both* the upper and lower scoring extremities, although the extent to which the distribution was inflated due to a greater number of extremely poor scorers compared with extremely good scorers and vice versa, varied between tests. Thus for example, the inflated distributions of the Trail Making Test and Digit Symbol Test in the later years was because of the presence of both disproportionately poor and good scorers, but was in large part due to numbers of particularly disproportionately poor scorers. For Finger Tapping the distribution increased due to a more balanced set of extreme good and poor scorers, and what was notable in particular for the Finger Tapping test were the numbers of good scorers which overlapped with the outside limits of the distribution of the youngest age groups.

Generally over all the tests which revealed significant alterations in variability, outlying good scorers were predominantly female and had high education, although there were also females and individuals with high education who were amongst the disproportionately low scorers. Thus a significant negative correlation between education and Trail Making Scores that occurred *exclusively for the 80's age group*, indicated the *especially* strong influence of high education on the ability to be a high scorer, and for low education to be associated with being a poor scorer, for this particular test and age group. However, the absence of significant correlations between the Trail Making Test and IQ, and between all the other tests and education, did not mean there were no important effects with respect to these two variables in terms of unravelling the nature of inter-individual variability in association with aging. Rather the lack of significant correlations was consistent with a strongly prevailing feature across all the tasks of *'opposite' effects*. (The term opposite effect is used here to describe those occasions when the generally expected high performance in association with high levels of education/and or IQ does not occur, and an opposite effect occurs of a low level of performance in association with high levels of education and/or IQ. In the case of longitudinal research 'opposite' effects might be demonstrated in the form of 'cross-over' effects for particular individuals in association with aging. This would be the case for individuals who initially might show the expected high level of performance in association with a high level of education and/or IQ, but at later age stages cross over to demonstrate the opposite effect of poor performance in association with high education and/or IQ). Thus contributing to the overall significant increases in variability with older age, there were numbers of individuals with high levels of IQ and education who were amongst the extreme scorers on *both* the good and poor scoring sides of the distribution. This would constitute a neutralizing effect in the correlation analyses and serve to produce an absence of any significant correlation coefficients between levels of IQ and test scores such as was seen in most instances for IQ and education in the Phase 2 analyses.

In essence, therefore, taken together, the results of the Phase 1 and Phase 2 between-groups analyses, and the Phase 2 within-groups analyses, have verified each one of the statistical hypotheses, and there was no evidence which seriously stood in contradiction of these hypotheses. The setting of the overall level of significance for the multiple comparisons at $p=0.10$ is considered to have provided a good balance between the danger of making either Type 1 or Type II errors (see Miller, 1981). Whenever the all groups analyses revealed a significant effect, consistently the nature of significant effects were identified by the multiple comparisons and highlighted those significant overall patterns which were conceptually consistent with theoretical indications. Frequently the differences for the pairwise comparisons were highly significant at the 1% level of significance following adjustments for the number of group comparisons, and thus would still have been significant even if the overall level of significance for multiple comparisons had been set more stringently at $p=0.05$. However, due to the possibility of making Type II errors that degree of stringency may have failed to crystallize out the nature of the patterns as effectively. The appropriate choice of significance level is confirmed to a degree by the fact that when the all-groups analyses did not reveal a significant effect, but approached significance with p-values between 0.05 and 0.10, frequently the lack of a significant effect was confirmed by a lack of significant effects for any of the pairwise comparisons.

9.3 VALIDATION OF THE SHUTTLE MODEL

As noted above (section 9.1), the statistical hypotheses were operationalized originally from the experimental hypotheses posed in terms of the shuttle model of variability with the purposes of formal testing of the model. As further noted above in (section 9.2), there was substantial support for all the statistical hypotheses, thus implying that in turn the experimental hypotheses posed directly from the shuttle model (section 5.3.4) were also fundamentally confirmed. These were as follows:

> 1. There is likely to be increased variability of scores on cognitive tasks in association with the initial stages of the cognitive aging process as identified on central trends analysis, followed by a subsequent decrease in variability as the aging process progresses, producing an overall *non-linear pattern* of an initial increase and subsequent decrease in variability in association with aging;
>
> 2. The presentation of the non-linear pattern of an initial increase and subsequent decrease in inter-individual variability in cognitive test performance in association with aging, will be differentially affected by the cohort effects of lower versus higher

levels of *education/IQ* and *male versus female gender*, and the task type effects of lower versus higher levels of *task challenge,* and lower versus higher levels of *age-sensitivity of tasks*. Specifically lower levels of education/IQ, male gender, higher levels of task challenge, and greater age-sensitivity of tasks, will cause the variability curve to shuttle to the left, that is the onset of increased variability and subsequent decrease in variability will occur *earlier* in the adult age range; conversely higher levels of education/IQ, female gender, lower levels of task challenge, and lesser age-sensitivity of tasks will cause the variability curve to shuttle to the right, that is the onset of increased variability and subsequent decrease in variability will occur *later* across the adult age range. (Although it was not possible on the basis of Satz's review to indicate definitive directions for gender effects, this refinement was made possible on the basis of the subsequent review on the origins of inter-individual variability - see section 4.2.2.3, and the arguments posed above under cohort-specific effects, section 5.2.1.1).

3. In some instances higher levels of education/IQ, and/or female gender, and/or lower levels of task challenge and lesser age-sensitivity of tasks may reveal a *progressive increase* in variability across the adult age range without any subsequent decrease being in evidence. This may imply that a subsequent decrease in variability would become apparent to reveal the full non-linear variability pattern of an initial increase followed by a subsequent decrease, were progressively older age groups assessed. On the other hand it may imply that a prolonged retention of increased variability would become apparent, without a subsequent decrease in variability, were much older age groups assessed.

4. In terms of hypotheses two and three above, it is hypothesised that when making comparisons of two variability effects (for the differential influences of education, IQ, task-type and gender), variability effects may occur in any one of the following combination of forms:

(i) Two full non-linear variability patterns of an initial increase and subsequent decrease may be in evidence across the adult age axis with the one for lower levels of education/IQ, male gender, higher levels of task challenge, and higher age-sensitivity of tasks shifted comparatively more to the left (that is the initial increase and subsequent decrease in variability will occur earlier), than the one for higher levels of education/IQ, female gender, lower levels of task challenge and lesser age-sensitivity of tasks which will be shifted comparatively more to the right (that is the initial increase and subsequent decrease in variability will occur later);

(ii) One full non-linear variability pattern of an initial increase and subsequent decrease may be in evidence for lower levels of education/IQ, male gender, higher levels of task challenge, or higher age-sensitivity of tasks which will be shifted to the left (that is occur earlier), together with only a progressive increase in variability in evidence for higher levels of education/IQ, female gender, lower levels of task challenge, or lesser age-sensitivity of tasks respectively, which will be shifted comparatively more to the right (that is the initial increase in variability will occur later and there will be no subsequent decrease in evidence);

(iii) No full non-linear pattern of an initial increase and subsequent decrease in variability may be in evidence at all across the adult age axis, but only two progressive increases in variability with the one for lower levels of education/IQ, higher levels of task challenge, higher age-sensitivity of tasks, or male gender shifted comparatively more to the left (that is the progressive increase in variability will occur earlier), and the one for higher levels of education/IQ, lower levels of task challenge, lesser age-sensitivity of tasks, or female gender, respectively, which will be shifted comparatively more to the right (that is the progressive increase in variability will occur later).

5. The above presentations of variability in association with the progression of adult age will occur in the presence of an *overall steady decline in average cognitive test performance* as reflected in mean scores.

6. Since differential levels of education, IQ and gender are likely to be intricately related to differential variability effects in association with aging depending upon the extent to which they serve a protective function, it is hypothesized that the effects of education, IQ and gender on the individual distribution of scores will not be uniform across age groups, and that this will be reflected in *differential characterization of the disproportionately low and high scorers (that is the outliers)* as follows:

(i) The effect of education, IQ and gender will be more pronounced for the age groups with significantly increased variability than for those age groups which do not show increased variability;

(ii) Where education, IQ and gender are having an effect higher education/IQ and female gender will characterize the disproportionately high scorers, and lower education/IQ and male gender will characterize the disproportionately low scorers.

The confirmation of the above hypotheses will be discussed under the two dimensions under which the shuttle model was originally conceived (see sections 5.3.1, and 5.3.2): the *shuttle shape* and *shuttle mechanism* of variability in association with aging. The mode of graphical illustration which follows in this subsection (Figures 9.1a to 9.4d, pp.395-6, 399-400, 402, 404-7), is commensurate with that in the model schema of Figure 5.6 (p.163), and serves to highlight both the shape and mechanical features of the shuttle effect in association with the central tendency. *In Figures 9.1a to 9.4d using data from Phase 1 and Phase 2 of the present study, variability effects are selectively illustrated by plotting average performance together with corresponding standard deviations against age.*

9.3.1 The shuttle shape of cognitive variability

First, in terms of the above hypotheses (specifically derived in terms of the shuttle variability model), there was clear support for the existence of a predictable significant bulge in cognitive variability in parallel with declining average performance in association with aging (the *shuttle bulge* effect). As hypothesized, this was revealed either in a full non-linear pattern of an initial increase and subsequent decrease in variability, or took the form of a sustained increase in variability with no subsequent decrease (hypotheses 1, 4 and 5). Specifically, the full differentiation/subsequent dedifferentiation pattern was apparent in most instances of tests of attention, visual and verbal memory; the sustained increase in variability with no subsequent decrease prevailed for memory recognition tasks, language, problem solving and hand motor tasks. For visual skills the pattern of variability effects (differentiation/subsequent dedifferentiation, or the sustained differentiation effect) varied, depending on the mix of abilities called upon for a particular visual task. Selectively, from the available Phase 1 and Phase 2 data, this shuttle bulge effect is graphically portrayed in each of Figures 9.1a to 9.4b, both upper and lower graphs (pp.395-6, 399-400, 402, 404-7), where it is revealed either in the form of only the sustained differentiation effect, or the initial differentiation/subsequent dedifferentiation effect, all in association with steadily declining average ability.

Overall, the implication is that for all the functional modalities tapped in the study, of attention, memory, language, problem solving and hand motor ability, the variability bulge (either comprising the initial differentiation effect, or the full differentiation/subsequent dedifferentiation effect) is a consistent feature in the presence of declining average ability. For the modalities where only a sustained increase in variability was apparent, it is not possible to know from the present study whether a subsequent decrease in variability would have become apparent to reveal the full shuttle bulge were older age groups tapped. A measure of support for this possibility was derived from the literature review in which there were reports of an initial increase followed by a subsequent decrease in variability for some language and problem solving tasks, which on the present empirical investigation tended not to reveal such an effect. However, since these reports were not derived from rigorous investigations into variability, further systematic research is needed, which includes much older age groups, in order to identify whether language, problem solving and hand motor tasks might continue to show the sustained bulge in variability right up to the point of demise. Thus taken together, the results of the literature review, and those of the present empirical investigation, could not remove the uncertainty in the shuttle model on this issue, which was noted in its original conception (see section 5.3.3), and is reflected in hypothesis number 3.

9.3.2 The shuttle mechanism of cognitive variability

Second, in terms of the above hypotheses (specifically derived in terms of the shuttle variability model), there was recurring support for the presence of predictable significant shifts in the overall variability curve in relation to the age axis, depending on differential cohort and task influences (the *shuttle shift* effect) (hypotheses 2 and 3). The confirmation of each of these particular effects in validation of the shuttle model are further discussed below for the cohort-specific effects of levels of *education* and *gender*, and for the domain/task-specific effects of levels of *task challenge* and *age-sensitivity of task*.

The shift effects are selectively illustrated in Figures 9.1a to 9.4d to follow. As noted before, these graphical representations are derived by plotting a selection of Phase 1 and Phase 2 average performance data together with the corresponding standard deviations against age. *In order to identify the shift effects, the upper and lower graphs in each instance need to be read as pairs. In each case the upper diagrams illustrate a left shuttle effect, and in relation to this, in each case the lower diagrams illustrate a right shuttle effect.* For each pair of graphs with respect to the age axis, the shuttle shift effect can be identified as follows. In the upper graph, the initial up-arrow (indicating a significant increase in variability) is in a *left-sided position* relative to the age axis, whereas in the lower graph the initial up-arrow (indicating a significant increase in variability) is in a *right-sided position* with respect to a comparable age axis. The up-arrows may or may not be followed by subsequent left and right-sided positioned down-arrows (indicating significant decreases in variability), depending on whether or not there is the presence of a subsequent dedifferentiation effect. Thus for all the figure-pairs, the up-arrows in the upper graphs delineate the relatively *early onset* of significantly increased variability with or without a subsequent down-arrow indicating the relatively *early onset* of a subsequent decrease in variability (*left shift effect*); in comparison with this, the up-arrows in the lower graphs indicate the relatively *late onset* of significantly increased variability with or without a subsequent down-arrow indicating the relatively *late onset* of a subsequent decrease in variability (*right shift effect*).

9.3.2.1 Cohort-specific effects

Generally with respect to cohort-specific shift effects, there were limited opportunities available in the Phase 1 and Phase 2 between-groups analyses for the investigation of the differential effects of education and gender, and none for levels of IQ. However in *all* instances of the data, where there *were* significant overall variability effects between age groups, and the subdivisions of samples were such that it was possible to make comparisons for subgroups of education and gender, shuttle shift effects *did* occur, and regularly were in the predicted direction.

With regard to *education*, as noted above (section 9.2.1) the results provided consistent confirmation of an earlier onset of variability (left shuttle effect) for low education compared with a later onset of variability (right shuttle effect) for high education. Specifically this phenomenon occurred in the Read (1987, in Spreen & Strauss, 1991) study for the FAS Controlled Word Association test, and for the Robertson-Tehabo and Arenberg (1989, in Spreen & Strauss, 1991) study on the Benton Visual Retention-Revised (BVRT-R) test, in both the male and female conditions. These shift effects for education are selectively illustrated in Figures 9.1a and 9.1b for the FAS Controlled Word Association test, and for the BVRT-R test in the female condition. In Figure 9.1a for the FAS Controlled Word Association Test, a non-linear pattern of an initial increase and subsequent decrease in variability was apparent for the low education (<12 years) condition and a linear pattern of an increase in variability only for the high education (>13years) condition, with the non-linear pattern for lower education shifted significantly more to the left (upper graph), and the linear pattern for higher education shifted significantly more to the right (lower graph). In Figure 9.1b for the BVRT-R test in the female condition, two progressive increases in variability were apparent, with the one for lower (No Degree) education shifted significantly more to the left (upper graph), and the one for higher education (College Degree) shifted significantly more to the right (lower graph). As noted above, although not illustrated in the present series of graphs, the same direction of shift effect for education was apparent for the BVRT-R test for males (see results section, Figure 7.14, p.242).

The highly consistent direction of the shift effects for education, for a verbal fluency task from Read (1987) and a visual retention task from Robertson-Tehabo and Arenberg (1989), are in turn consistent with the cohort-specific effects identified from the data presented in Salthouse (1991) (see discussion in section 4.2.1.1, and illustrated for Comprehension, Digit Span, Similarities and Vocabulary in Figure 4.4, p.100). Although Salthouse suggested there was no particular pattern of effects, it was apparent on further scrutiny of the data he presents, that the arguably lower functioning WAIS cohort consistently showed the earlier onset of increased variability and earlier subsequent decrease in variability (left shuttle

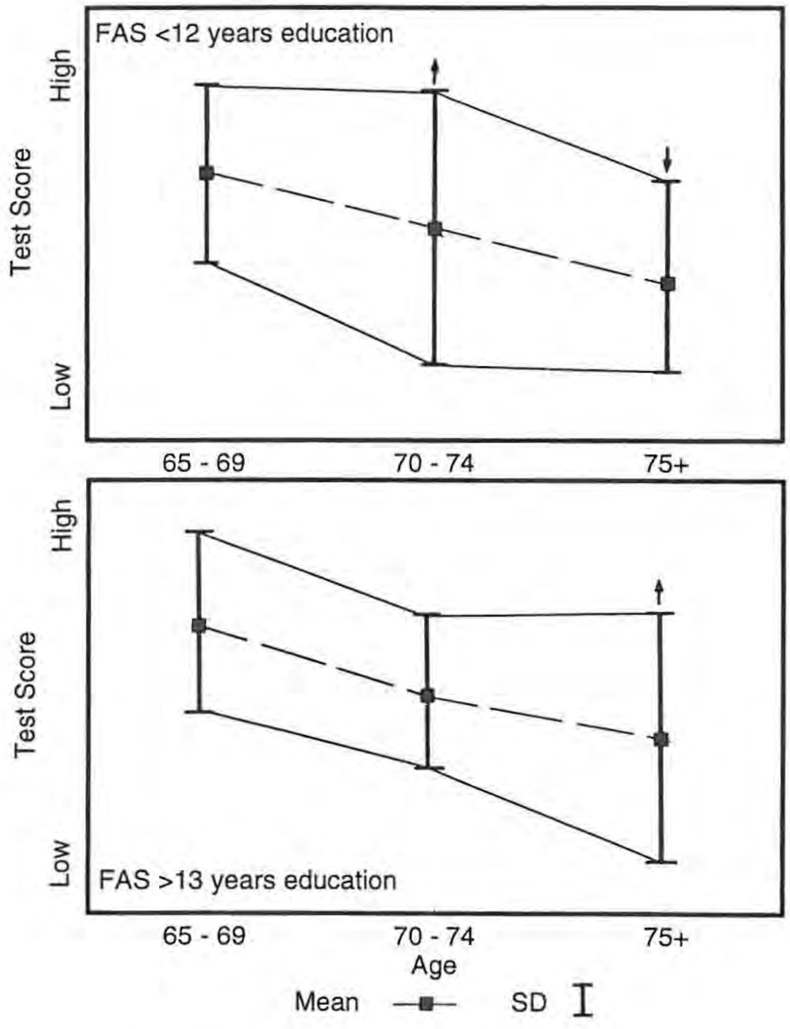


Figure 9.1a. The shuttle effect for low versus high levels of education: Illustration of the earlier onset of increased variability (up-arrow) and subsequent decrease in variability (down-arrow) for the FAS < 12 years of education condition (left shuttle effect, upper graph) compared with the later onset of increased variability and no subsequent decrease in variability for the FAS > 13 years of education condition (right shuttle effect, lower graph). (Data set for the FAS Word Association Test from Read, 1987, in Spreen & Strauss, 1991).

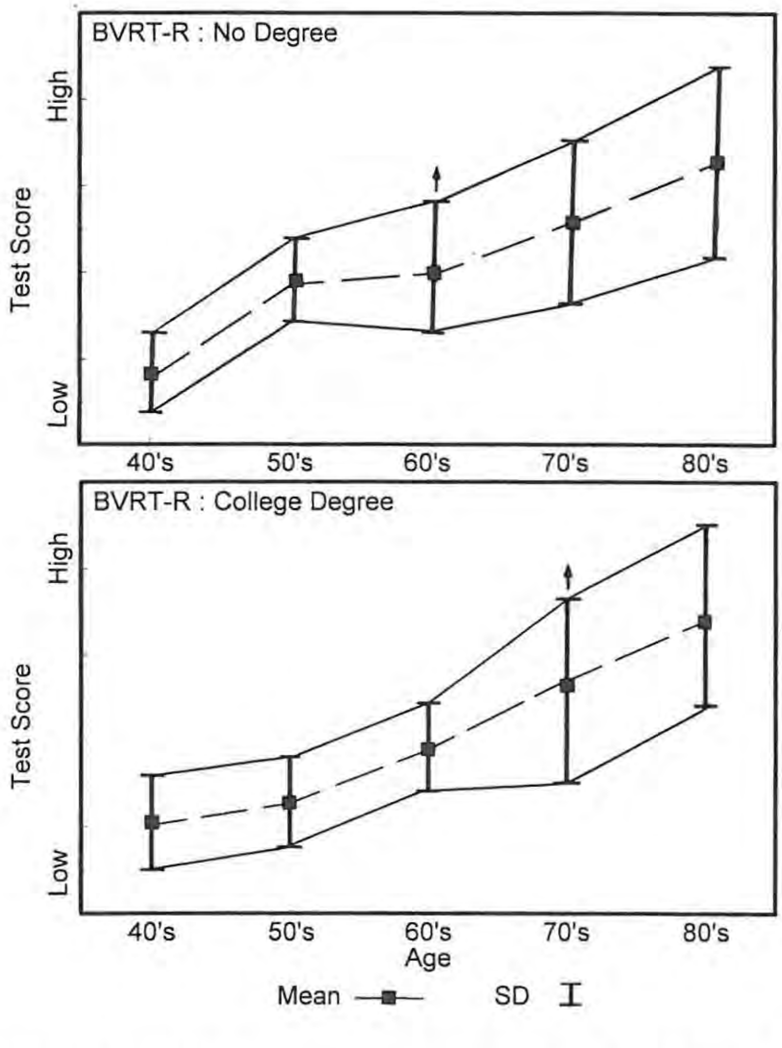


Figure 9.1b. The shuttle effect for low versus high levels of education: Illustration of the earlier onset of increased variability (up-arrow) for the BVRT-R No Degree condition (left shuttle effect, upper graph) compared with the later onset of increased variability for the BVRT-R College Degree condition (right shuttle effect, lower graph). (Data set for the Benton Visual Retention Test-Revised, BVRT-R, from Robertson-Tehabo & Arenberg, 1989, in Spreen & Strauss, 1991).

effect), compared with the arguably higher functioning WAIS-R cohort which consistently showed the later onset of increased variability and later subsequent decrease in variability (right shuttle effect). Although comparative subgroups for IQ were not directly available for the examination of this effect, due to the high correlation between education and IQ (see discussions in sections 4.2.2.2, and 5.2.1.1), the inference can be drawn that the same direction of shift effects would be true for IQ. This inference gains support from the Phase 2 within-groups analyses where it was possible generally to investigate the effects of education, but also for IQ specifically on the Trail Making Test. These within-groups analyses revealed a notable connection between enhanced variability in association with aging, and education and IQ. This was in that significant within-groups variability effects for both education and IQ were more consistently in evidence for the age groups with significantly increased variability in the between-groups comparisons, than for those age groups which did not show increased variability. Further from the Phase 2 within-groups analyses on the Trail Making Test, it was evident that high education and high IQ characterized the disproportionately high scorers with increasing potency with advancing years.

Thus this particular left versus right cohort-specific shift effect for higher versus lower functioning samples, respectively, (i) was in evidence from widely disparate sources including the WAIS versus WAIS-R subtest data illustrated by Salthouse, and two data sets in the present investigation including the Read (FAS) data, and the Robertson-Tehabo and Arenberg (BVRT-R) data; (ii) occurred across a spectrum of functional modalities including both verbal and non-verbal, crystallized and fluid functions; (iii) was apparent (albeit for only one study) for both male and female conditions of gender; and finally, (iv) was supported by the characterization of the outliers in terms of both education and IQ in the Phase 2 analyses. Together this all adds up to a relatively robust measure of validation for the predictable occurrence of left versus right shift effects, respectively, for low relative to high functioning samples, including in particular low versus high levels of education, and more tentatively IQ.

With respect to *gender*, as noted above (section 9.2.1) the Phase 1 between-groups results provided consistent confirmation of an earlier onset of variability (left shuttle effect) for male gender compared with a later onset of variability (right shuttle effect) for female gender. This direction of effects occurred in each of the four instances where it was possible to make comparisons between subgroups for gender and where there were significant overall variability effects. Specifically it was apparent in the Geffen et al., (1990, in Spreen & Strauss, 1991) study for the Rey Auditory Verbal Learning Test (RAVLT) Trial 1 and Recognition tasks, and for the Robertson-Tehabo and Arenberg (1989, in Spreen & Strauss, 1991) study on the Benton Visual Retention-Revised (BVRT-R) test, for both the low and high educational conditions. These comparative effects for subgroups of gender, are selectively illustrated for the RAVLT Trial 1 (Figure 9.2a), for which two full inverted-U variability curves were in evidence, with the one for male gender shifted significantly more to the left (upper graph), and the one for female gender shifted significantly more to the right (lower graph); and for the RAVLT Recognition trial (Figure 9.2b), for which two significant increases in variability with no subsequent decreases were in evidence, with the one for male gender shifted significantly more to the left (top graph), and the one for female gender shifted significantly more to the right (lower graph).

The noted shift effects for gender, of a earlier onset of variability for males relative to females, albeit consistently present, were nevertheless restricted to the limited number of studies available for such analysis. Moreover (unlike for educational/IQ levels), there were no specific variability patterns available from the review on the possible direction of between-groups gender effects which could provide support for these effects. However, some potency for the observations is gained by the fact that it was present in a data set which controlled for levels of education across subsets of gender (see results for the BVRT-R, Figure 7.14, p.242). Also, the direction of effects were consistent with gender-related indications from the Phase 2 within-groups analyses. In the 80's age group for TMT A, the standard deviation for males was significantly lower than that for females ($p<0.01$); and in the 80's age group for TMT B there was a strong trend in the same direction ($p=0.0698$). Together this implies a more severe aging effect (and associated earlier dedifferentiation effect) for males compared with females for this task. Further in the Phase 2 within-groups analyses, female gender characterized the disproportionately high scorers with increasing potency with advancing years. Thus the male vulnerability factor appeared to be a sufficiently robust occurrence across the various levels of analysis in both Phase 1 and Phase 2 to suggest that it was a finding of relevance. Furthermore, whilst the number of studies available for analysis of gender effects were limited, and thus the observations must remain tentative, there was no evidence which served to *invalidate* the presence of gender effects, or to contradict the particulár direction of the effects.

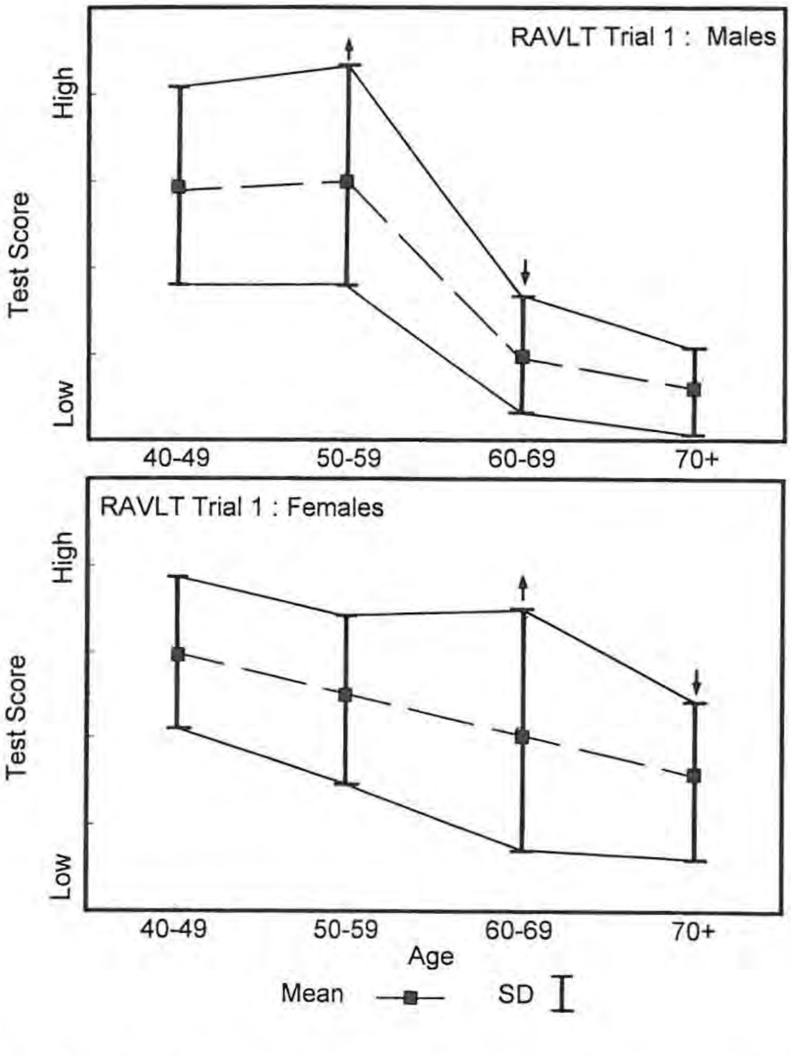


Figure 9.2a. The shuttle effect for male versus female gender: Illustration of the earlier onset of increased variability (up-arrow) and subsequent decrease in variability (down-arrow) on the RAVLT Trial 1 for males (left shuttle effect, upper graph) compared with the later onset of increased variability and later subsequent decrease in variability on the RAVLT Trial 1 for females (right shuttle effect, lower graph). (Data set for the Rey Auditory Verbal Learning Test, RAVLT, from Geffen et al. 1990, in Spreen and Strauss, 1991).

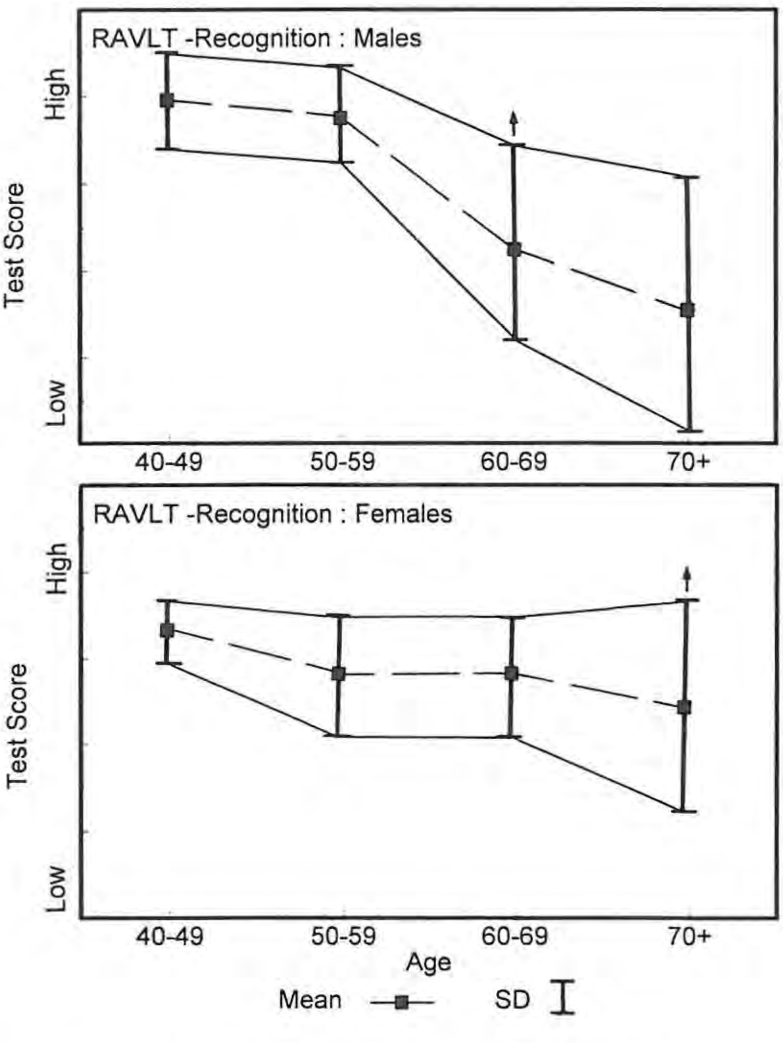


Figure 9.2b. The shuttle effect for male versus female gender: Illustration of the earlier onset of increased variability (up-arrow) on the RAVLT Recognition for males (left shuttle effect, upper graph) compared with the later onset of increased variability on the RAVLT Recognition for females (right shuttle effect, lower graph). The up-arrows represent the onset of significantly increased variability in each case. (Data set for the Rey Auditory Verbal Learning Test, RAVLT, from Geffen et al. 1990, in Spreen and Strauss, 1991).

Thus generally, with respect to the *cohort-specific effects* of both *educational/IQ levels* and *gender*, there was strong cross-validation from the Phase 2 within-groups analyses for the direction of the shuttle effects noted in the Phase 1 between-groups analyses (as illustrated selectively in Figures 9.1a and 9.1b, 9.2a and 9.2b). This was in that in the Phase 2 within-groups analyses (in full support of hypothesis 6), the effects of education, IQ and gender were more pronounced for the age groups with significantly increased variability than for those age groups which did not show increased variability; and that high education, high IQ and female gender characterized the disproportionately high scorers with increasing potency with advancing years. An important supplementary finding however, was of opposite effects (see definition above, p.388), in that high education, high IQ and female gender, whilst being the requisite characteristics for successful scorers, did not *necessarily* preclude the possibility of being a disproportionately unsuccessful scorer.

9.3.2.2 Domain/task-specific effects

Generally with respect to domain/task-specific shift effects, there were limited opportunities available in the Phase 1 and Phase 2 between-groups analyses for the investigation of the differential effects of level of *task challenge*. In addition, a shift effect was not always in evidence across two levels of the same task (for example the van Gorp et al. 1990, and Stuss et al. 1987 data for the Trail Making Test). However, whenever a shift effect *did* occur, there was consistent support for its occurrence in the predicted direction of an earlier onset of variability (left shuttle effect) for high task challenge compared with a later onset of variability (right shuttle effect) for low task challenge. This effect occurred specifically in *two* instances of the higher challenge PASAT 1.2 presentation versus the lower challenge 2.4 level of presentation (Stuss et al. 1987, and the Roman et al. 1991 studies); and for the higher challenge Trail Making Test B versus the lower challenge Trail Making Test A (Cornfield & Shuttleworth-Jordan, 1996 study, see Figures 7.1, p.195, and 8.1, p.328). The shift effect for high versus low level of task challenge is selectively illustrated again for the PASAT 1.2 and 2.4 second presentations (Roman et al. 1991 study) in Figure 9.3. A significant inverted-U effect is apparent for the high challenge 1.2 second task (upper graph), and only a significant progressive increase in variability with no subsequent significant decrease is apparent for the low challenge 2.4 second task (lower graph).

For *age-sensitivity of task* it has been possible to make a fair number of comparisons which provide consistent support for shift effects in the predicted direction of an earlier onset of

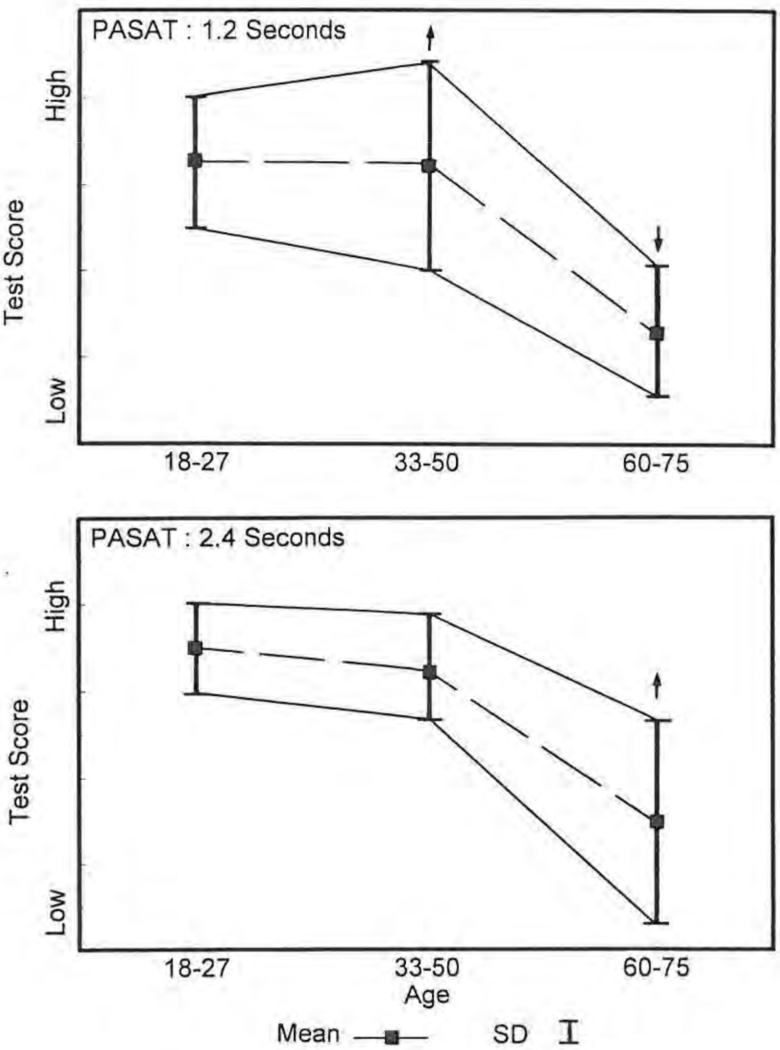


Figure 9.3. The shuttle effect for high versus low levels of task challenge: Illustration of the earlier onset of increased variability (up-arrow) and subsequent decrease in variability (down-arrow) for the PASAT 1.2 second more challenging task (left shuttle effect, upper graph) compared with the later onset of increased variability and no subsequent decrease in variability for the PASAT 2.4 second less challenging task (right shuttle effect, lower graph). (Data set for the Paced Auditory Attention Addition Test, PASAT, from Roman et al., 1991).

variability (left shuttle effect) for high age sensitivity of task compared with a later onset of variability (right shuttle effect) for lower age sensitivity of task. This effect was highlighted for the relatively more age-sensitive Buschke Selective Reminding Test BSRT-RLTR and CLTR learning trials compared with the arguably less age-sensitive Buschke Selective Reminding Test BSRT-Total and Delayed short-term recall trials (in that the former tasks are based more heavily on sustained learning) (see prior argument pp. 217-8 and Figure 7.8, p.217). Figure 7.8 shows an increase followed by the (presumably earlier) appearance of a subsequent decrease in variability for the RLTR and CLTR more age-sensitive tasks, (lower graph) compared with only a sustained linear increase in variability with no appearance of a subsequent decrease in variability for the less age-sensitive BRSR-Total and Delayed BSRT (upper graph). Further a task-related shift effect was noted for the more age-sensitive speeded Rey Visual Design Learning Test compared with the arguably less age-sensitive non-speeded WMS Visual Reproduction task (see previous argument pp.246-7, and Figure 7.16, p.248). Reference to the Figure 7.16 upper two graphs, reveals the earlier onset of the full variability curve for the more age-sensitive speeded RVDLT (middle graph, left shuttle effect), compared with the later onset of the full variability bulge for the relatively less age-sensitive non-speeded WMS Visual Reproduction task (top graph, right shuttle effect).

Several more of these effects are selectively illustrated in Figures 9.4a, 9.4b, 9.4c and 9.4d. In Figure 9.4a, the relatively more age-sensitive fluid visuo-perceptual Digit Symbol task reveals an earlier onset of variability (left shuttle effect) when compared with a relatively less age-sensitive language task; in Figure 9.4b the arguably more age-sensitive fluid visuo-perceptual Digit Symbol task reveals an earlier onset of variability (left shuttle effect) when compared with the arguably relatively less age-sensitive Trail Making B task (due to its rational ordering component); and in Figure 9.4c, the relatively more age-sensitive fluid visuo-perceptual Digit Symbol task reveals an earlier onset of variability (left shuttle effect) when compared with a relatively less age-sensitive problem solving task. In Figure 4d, the relatively more age-sensitive RAVLT short term verbal memory task similarly reveals an earlier onset (left shuttle effect) when compared with the less age-sensitive RAVLT recognition task. In all instances of these illustrations there was a significant inverted-U variability curve in evidence for the relatively more age-sensitive tasks which were shifted to the left (upper graphs in all instances), in conjunction with only a sustained significant increase in variability with no subsequent decrease in evidence for the less age-sensitive tasks, with the initial onset in variability shifted to the right (lower graphs in all instances).

Thus overall from the results of Phase 1 and Phase 2 between-groups analyses, and the Phase 2 within-groups analyses, there was a high level of verification for all six hypotheses of the shuttle model of variability, including the postulated shuttle bulge, and each of the proposed shuttle shift effects due to cohort and task-related influences. There was no evidence to suggest that any of the basic tenets of this model should be refuted.

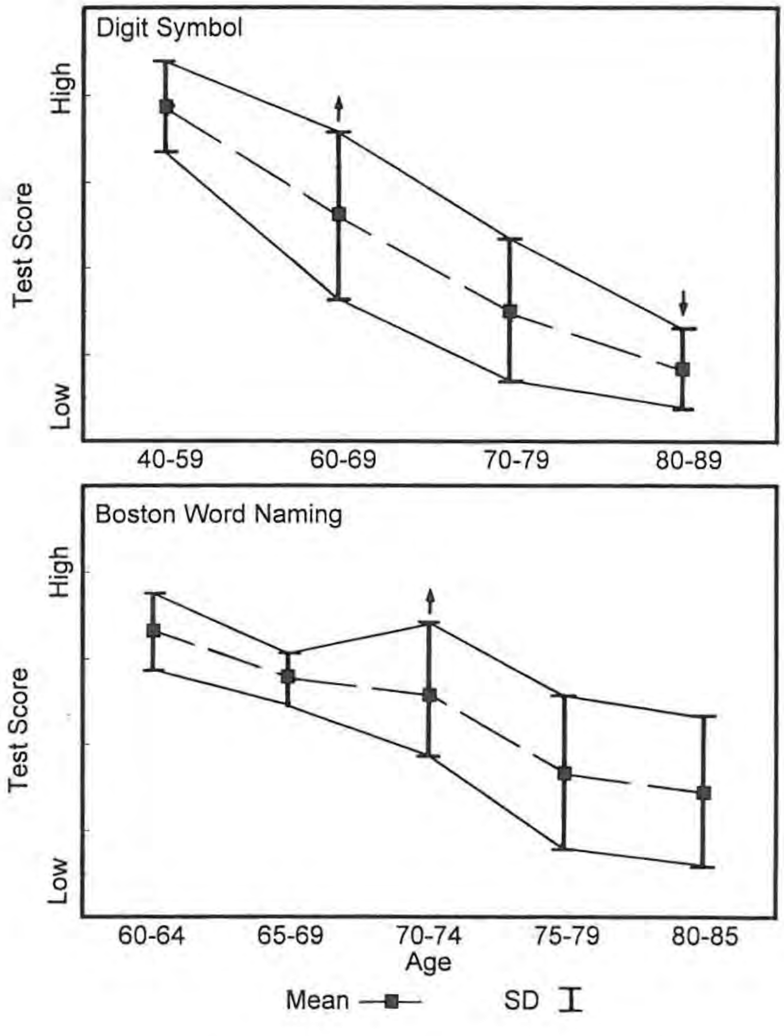


Figure 9.4a. The shuttle effect for high versus low levels of age sensitivity of task: Illustration of the earlier onset of increased variability (up-arrow) and subsequent decrease in variability (down-arrow) for the Digit Symbol task (left shuttle effect, upper graph) compared with the later onset of variability and no subsequent decrease in variability for the Boston Word Naming task (right shuttle effect, lower graph). (Data set for Digit Symbol from Shuttleworth-Jordan & Bode, 1993; Data set for Boston Word Naming from van Gorp et al., 1986, in Spreen & Strauss, 1991)

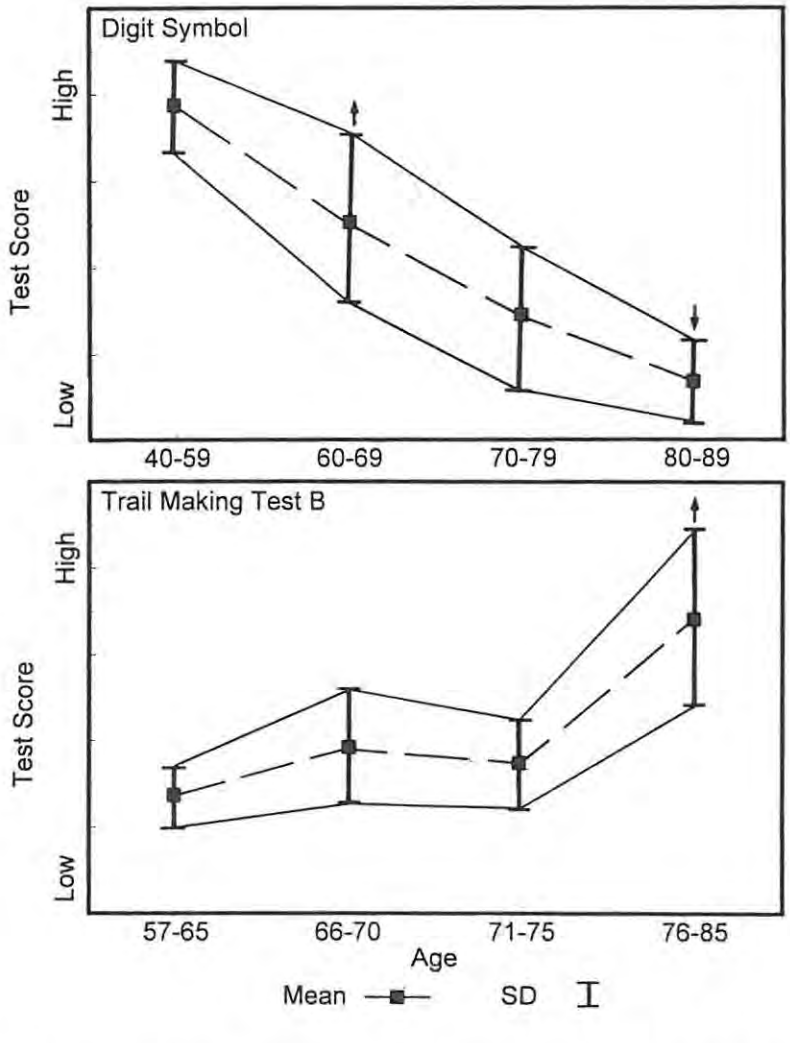


Figure 9.4b. The shuttle effect for high versus low levels of age sensitivity of task: Illustration of the earlier onset of increased variability (up-arrow) and subsequent decrease in variability (down-arrow) for the Digit Symbol task (left shuttle effect, upper graph) compared with the later onset of increased variability with no subsequent decrease in variability for the Trail Making Test B (right shuttle effect, lower graph). (Data set for Digit Symbol from Shuttleworth-Jordan & Bode, 1993; Data set for Trail Making Test B from van Gorp et al., 1990).

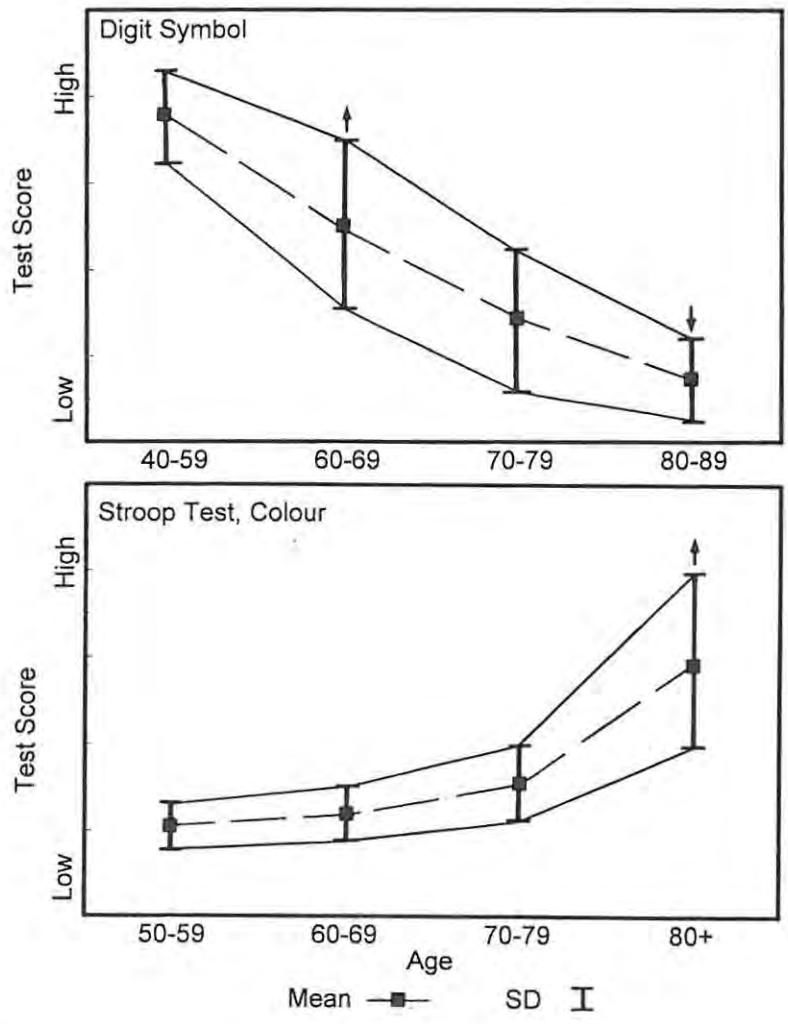


Figure 9.4c. The shuttle effect for high versus low levels of age sensitivity of task: Illustration of the earlier onset of increased variability (up-arrow) and subsequent decrease in variability (down-arrow) for the Digit Symbol task (left shuttle effect, upper graph) compared with the later onset of increased variability and no subsequent decrease in variability for the Stroop Test (right shuttle effect, lower graph). (Data set for Digit Symbol from Shuttleworth-Jordan & Bode, 1993; Data set for the Stroop Test from collated data in Spreen & Strauss, 1991).

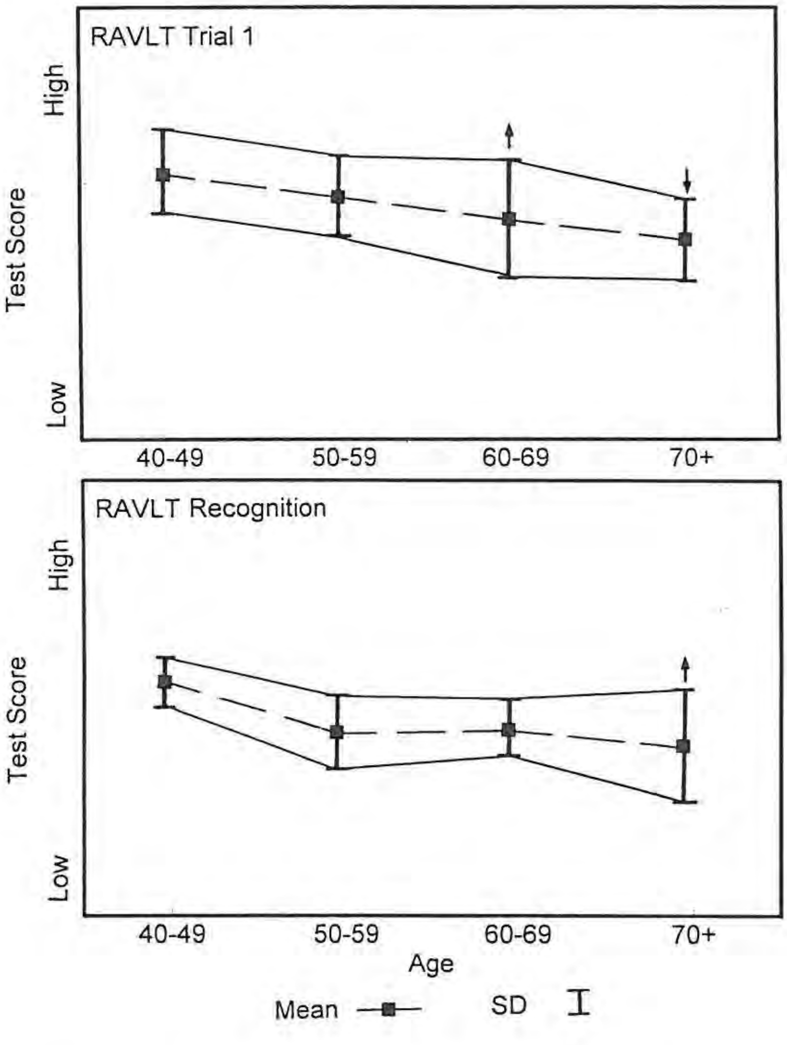


Figure 9.4d. The shuttle effect for high versus low levels of age sensitivity of task: Illustration of the earlier onset of increased variability (up-arrow) and subsequent decrease in variability (down-arrow) for RAVLT Trial 1 (left shuttle effect, upper graph) compared with the later onset of increased variability and no subsequent decrease in variability for RAVLT Recognition (right shuttle effect, lower graph). (Data set for the Rey Auditory Verbal Learning Test, RAVLT, from Geffen et al. 1990, in Spreen and Strauss, 1991).

9.4 SUPPORT FOR BRAIN RESERVE CAPACITY (BRC) THRESHOLD THEORY

As noted earlier, the shuttle model of variability is fundamentally an empirically-based model derived from patterns noted in the review of variability research. However, the model derived explanatory power from the central indications which emerged from the literature review on neural and psychometric perspectives on aging, including, in particular, the tenets of brain reserve capacity *threshold* theory as formulated by Satz (1993). The hypotheses for the shuttle model of variability were formulated as an elaboration of specific BRC postulates which were stipulated in Chapter 3 (section 3.2.3.3), and were necessarily entirely commensurate with these postulates. The verification of the shuttle model, therefore, provides strong support in turn for the validation of the threshold hypotheses extrapolated from BRC theory, which were as follows:

> 1. From BRC theory, it is hypothesized that the inter-individual variability in cognitive test performance in association with the neural aging process will be influenced by inter-individual differences in levels of brain reserve capacity which alter the threshold of vulnerability to, and protection against, the onset of functional symptom presentation. In particular, from the Satz review it is expected that low education and IQ, high task challenge, and by inference male gender, will serve as threshold-lowering factors and will thus be associated with earlier symptom onset with aging; and conversely it is expected that high education and IQ, low task challenge, and by inference female gender, will act as threshold-raising factors and will be associated with later symptom onset with aging. (The Satz review did not provide a definitive direction for gender effects, however these were inferred from the well-known fact of female longevity in conjunction with general BRC principles).
>
> 2. From the psychometric perspective in conjunction with BRC theory, it is hypothesized that differential sensitivity to aging processes <u>between</u> different cognitive tasks and functional modalities, as identified on central trends analyses, will have an associated differential effect on the patterns of variability in cognitive task performance. This is in that tasks which are less sensitive to aging on central trends analysis are evidently those that present less challenge to brain capacity thresholds in association with aging, and those tasks which have been demonstrated to be more sensitive to aging on central trends analysis, are evidently those that present more challenge to brain reserve capacity thresholds; hence differential effects between test modalities will be reflected similarly for both central trends and variability data in that for those tasks and modalities where central trends effects are more marked, variability effects will also be more marked and vice versa.

All aspects of the above two hypotheses (which were extrapolated on the basis of the Satz BRC theory) were supported by the significant shuttle effects identified in the between-groups analyses of Phase 1 and 2 of the study (as selectively illustrated in Figures 9.1a to 9.4d, pp. 395-6, 399-400, 402, 404-7), which clearly demonstrated a series of differential shift effects for the cohort-specific factors of levels of education and gender, and the task/domain-specific effects of task challenge and age-sensitivity of task. There was the consistent indication that in each case, low levels of education, male gender, high levels of task challenge, and greater age-sensitivity of task, served as threshold-lowering factors in that they were associated with an earlier onset of cognitive variability in conjunction with aging (left shuttle effect); and conversely there was the indication that in each case, high levels of education, female gender, low levels of task challenge, and less age-sensitivity of tasks served as threshold-raising factors in that they were associated with relatively late onset of cognitive variability with aging (right shuttle effect). Taken together each pair of graphs in Figures 9.1a to 9.4d, in effect illustrate a series of interaction effects: between levels of education and age, gender and age, level of task challenge and age, degree of task-sensitivity and age. For example, given the same task, there are different effects for high versus low education due to age, and for male versus female gender due to age; and given a comparable cohort, there are different effects for high versus low task challenge and high versus low age-sensitivity of task due to age. Thus overall, *including* the commensurate indications from the literature review on variability trends (as drawn together in the previous subsection), it was persuasively apparent that low levels of education, male gender, exposure to high task challenge and exposure to a highly age-sensitive task all served to increase individuals' vulnerability to symptom onset, and, in aggregation with the brain stressor of aging, resulted in relatively early symptom onset. Compared with this, high levels of education, female gender, exposure to low task challenge and exposure to tasks of low age-sensitivity, all served to protect individuals against symptom onset in aggregation with aging which occurred at a relatively later age stage.

Further from the Phase 2 within-groups analyses, particular support for BRC threshold theory was gained from the finding that for the Trail Making Test, the protective effects of high education in particular, but also IQ and female gender, appeared to *increase* with advancing age; and conversely the vulnerability effects of low education and male gender became more pronounced with older age. This implies that *as neural attrition takes on an increasing hold, factors which serve to increase protection against symptom onset, or conversely factors which serve to enhance vulnerability to symptom onset, elicit more powerful effects.* An excellent example of this particular progressive and cumulative unfolding of a BRC threshold effect has been explicated by Gronwall (1989) in relation to the vulnerability factor of mild head injury. She describes a classical experimental demonstration by Ewing et al. (1980) which revealed that whereas there were no cognitive deficits apparent for subjects with a history of mild head injury compared with non head-injured controls when tested at a *normal* altitude, deficits became clearly apparent for the head-injured subjects (but *not* for the controls) when subjects were tested at a *high* altitude. This experiment served to indicate that the vulnerability factor of a previous mild head injury came into effect *more potently* with the presence of the additional brain stress factor of high altitude. Similarly, in the present study the vulnerability factors of low education, low IQ and male gender (as well as the protective factors of high IQ, high education, and female gender) were seen to take on progressively *enhanced* effects in the presence of the strengthening brain stress factor of neural attrition in association with progressively older age.

Possibly the most subtle, and yet also most striking illustration of progressively marked effects on variability due to increasing brain stress, which as good as serves the purposes of a critical experiment in support of BRC theory, comes from the cumulative consequences of older age in aggregation with increased levels of *task challenge*. Viewed together, Figures 9.5a for the PASAT, and 9.5b for the Trail Making Test (which are replications of Figures 7.3 and 8.1 from the Phase 1 and Phase 2 results, respectively), serve to highlight the markedly different *interactive* effects of the *standard deviations in relation to mean scores* due to age that occur for low levels of task challenge compared with high levels of task challenge (upper versus lower graphs, respectively) in the case of both sets of tasks. In each instance of *low* task challenge, variability increases *closely in parallel* with declining average ability (PASAT 2.4 seconds) or *later* than decreasing average ability (TMT A). In other words, in the case of the two low challenge tasks, at the age stage when the mean (dotted line) starts to decline, the standard deviation (solid line) also starts to increase (PASAT 2.4 seconds), or subsequently increases (TMT A). In prominent contrast to this, for each instance of *high* task challenge (PASAT 1.2 seconds and TMT B) variability increases dramatically *before* the onset of markedly diminishing average tendencies. Thus it appears that for tasks of particularly high challenge, the effect of differential levels of BRC between individuals is potentiated in a manner which causes enhanced variability of scores to *signal* the onset of aging effects more strongly and earlier than is evident from the central tendency.

Of particular note from the Phase 2 within-groups analyses, with reference to BRC threshold theory, was the indication that whilst high education and female gender served invariably to characterize the individuals who were extremely good scorers in spite of advancing age, these

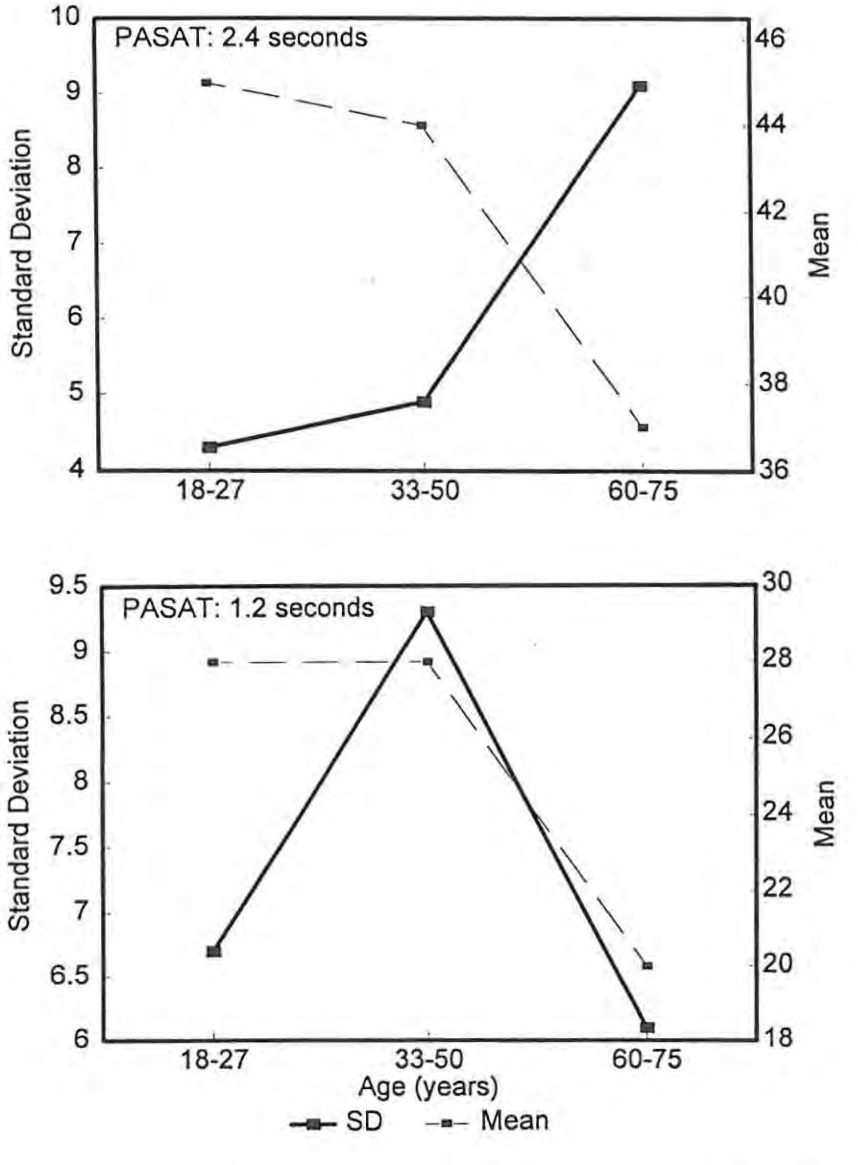


Figure 9.5a. PASAT (2.4 and 1.2 seconds): Illustration of the differential interactive effects of standard deviations in relation to mean scores across age groups due to low versus high task challenge (upper versus lower graphs, respectively).

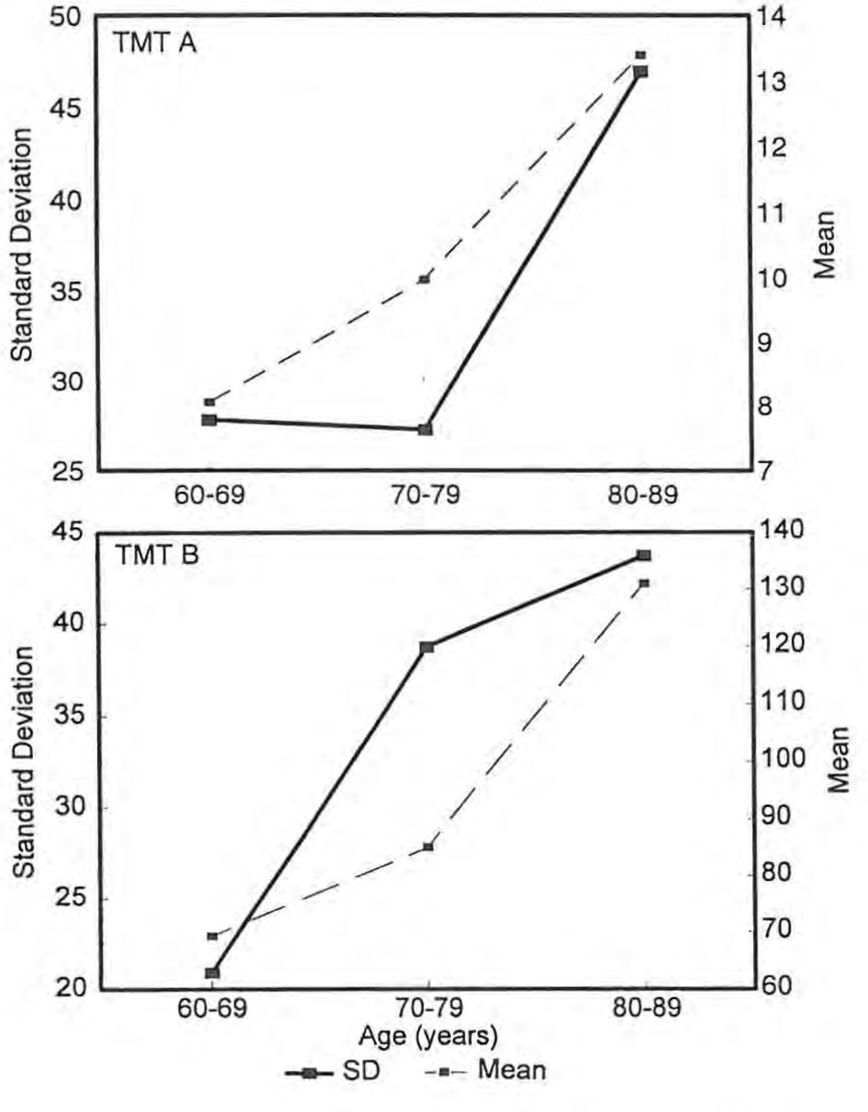


Figure 9.5b. Trail Making Test (Parts A and B): Illustration of the differential interactive effects of standard deviations in relation to mean scores across age groups due to low versus high task challenge (upper versus lower graphs, respectively).

factors did not invariably serve to protect against the possibility of individuals with high education or female gender being amongst the extremely poor scorers. Thus there were opposite effects for individuals with respect to the predominantly protective factors of high education and female gender. This implies that powerful vulnerability effects came into play which overrode the generally protective effects of high education and female gender, and served to lower BRC threshold differentially between individuals, in this way further complicating and contributing to the picture of the enhanced variability in association with aging. The many possible sources of such overriding vulnerability factors is an important route for further investigation (see discussion to follow below). However, the phase 2 within-groups correlational analyses provided support for one type of influence which might contribute to such apparently contradictory effects as follows.

There was a significant negative correlation between the Digit Symbol subtest and timed Finger Tapping scores for all age groups except the 80's age group, the age group for which there was also a significant dedifferentiation effect for the Digit Symbol task. This finding of a significant correlation between the two functional variables of visuo-perceptual graphic speed (Digit Symbol task) and hand motor dexterity (Finger Tapping test) across all age groups up to the 70's is commensurate with the argument of Salthouse (for example, Salthouse 1994), that individual differences in one functional area are strongly dependent on individual differences in other relevant functional areas. However, the fact that *selectively* there was no longer a significant correlation between Digit Symbol and hand motor function in the older 80's age group is highly commensurate with a similar dissociation of effects due to age noted by Rabbitt (1993b; 1993c). In Rabbitt's research he found that increasing numbers of outliers revealed a dissociation between IQ and memory functioning with advancing age. This caused him to challenge Salthouse's single factor hypothesis which implies the inter-dependency of functions as a universal observation across all functions and all age stages (see discussions of these debates in section 3.2.3.3). Rabbitt's findings suggest that the increased stress to the brain of late old age has overridden the predominantly protective effects of IQ for some individuals, which was in operation in relation to memory function in the earlier years, resulting in this dissociation of such functions for some individuals in late old age. Similarly the correlational analyses of the present research suggest that the increased stress to the brain of late old age has overridden the protective factor of hand motor skill for some individuals which was prevalent in relation to Digit

Symbol test performance in the earlier years, resulting in a dissociation of these generally related functions for a significant number of individuals in late old age.

Taken together, the cumulative evidence from the Phase 2 within-groups analyses indicates strongly that the widening distributions in association with older age in some measure appear to describe the widening out of the upper and lower percentiles of cognitive ability across the life span. Thus for example, high education, is a necessary attribute to have in order to be one of the extremely successful cognitive agers. However, on the basis of evidence for opposite (apparently contradictory) effects in the present study, as well as the evidence for differential sets of relationships between variables which change in relation to the progression of age, it is apparent that the situation is not as simple as that - a point that Rediess and Caine (1996) have also strongly emphasized. In other words the widening distributions in association with older age do not serve *entirely* to describe, or explain, the widening out of the upper and lower percentiles of cognitive ability with aging, and nor (as indicated above) does the single factor model of aging of Salthouse (1994). Rather BRC threshold theory as invoked here (which is more commensurate with Rabbitt's multifactor model of aging), serves to elucidate the complexity of these occurrences. The fundamental BRC postulate (as explicated in section 3.2.3.3) is that individual differences exist in levels of brain reserve capacity which alter the threshold of symptom onset (due to neural disease) differentially between individuals. This supposition can explain the expanding distributions in association with older age due to the widening of upper and lower limits in generally predictable directions. It also explains the contribution to the widening of the distribution due to less obviously predictable opposite effects, as follows.

First, as argued earlier, it appears that older age creates a neural stressor which *enhances* the lowering of the BRC threshold, and which in turn initiates the onset of symptom presentation in generally predictable directions. Thus a variable such as low education will most frequently serve as a vulnerability factor and high education as a protective factor. In association with the cumulative brain stress of neural aging, this in turn will serve to escalate the extent to which an individual with low education or high education would become disproportionately poor or good scorers, respectively. Second, however, as also indicated earlier, the cumulative effects of neural attrition with progressing age will lower individuals' BRC threshold levels *differentially*. Thus in some instances the additive effect of advancing age and other stress factors will serve to override individual protective factors such as those of high education and IQ, or, for example, high levels of hand motor skill, and such previous advantages will no longer protect a proportion of individuals against symptom onset in the face of the insidiously progressing stress factor of neural aging. These differential opposite effects will also contribute to the widening distribution. However, once *all* protective factors are overridden by *serious* advances in age with respect to a particular task, the consequence will be a *reduction* in the previously widened distribution - the phenomenon which is reflected in the frequently-noted older age dedifferentiation effect.

Thus in sum, the whole notion of the bulge in variability in association with aging, verified on the basis of the present research can be conceptualized purely in terms of BRC theory: Initially as individuals pass differentially below the BRC threshold of symptom presentation in association with the insidious onset of neural aging there will be a pattern of increased inter-individual variability due to the *enhancement* of differential vulnerability and protective factors in both *predictable* directions (spreading out of the normal upper and lower limits of the distribution), as well as in *unpredictable* opposite directions; subsequently, in interaction with the degree of challenge in a task, all individuals are likely to pass well beyond the BRC threshold and succumb more uniformly to the greater brain stress of further aging, such that in later old age there will be a subsequent pattern of *diminished* inter-individual variability.

Overall, therefore, the persistent bulge and shuttling mechanism of cognitive variability that has been demonstrated here, as a function of the risk factor of older age in aggregation with specific cohort and task-related threshold-altering factors, appears, in effect, to be nothing less than a powerful demonstration of fundamental tenets of BRC threshold theory as operationalized by Satz (1993). Importantly, whilst the validation of the shuttle variability model, provides in turn validation for Satz's proposed BRC threshold-altering factors, it does not contribute to clarification of the notion of what constitutes brain reserve capacity in itself. This is a concept which Satz concedes to have defined somewhat broadly, including its equivalence to factors such as brain size, dendritic branching, neuronal efficiency, adaptive behaviour, intelligence and education (see discussion in section 3.2.3.3). However, without entering into discussion about fine details of the brain reserve concept itself (which is beyond the scope of the present thesis), the central point is as follows. *Arising from the BRC notion, Satz has identified a number of influences which are likely to raise or lower the threshold of symptom onset in association with the progression of neuronal degeneration, and the outcome of the present study provides significant confirmation for the postulated direction of Satz's* *proposed effects*. In terms of available aging theories, it is difficult to explain the shuttle bulge and shuttle shift variability effects, *without* invoking the notion of brain reserve capacity (broadly defined as neuronal efficiency), which differs between individuals due to a variety of vulnerability and protective factors, and which, in aggregation with a superimposed brain stressor in the form of neural aging, leads to differential symptom onset between individuals.

9.5 CRITICAL EVALUATION OF THE STUDY

On the basis of a comprehensive literature review on inter-individual variability in cognitive aging recurring trends were identified, and on the basis of this a specific shuttle model of variability was developed. This model in turn has received substantial verification on the basis of the present two-pronged empirical investigation (as discussed in sections 9.1 to 9.4, above). However, two potentially problematic issues remain which need to be addressed. First there is the problem of whether or not the whole shuttle effect (including the bulge and the shifting mechanism) is simply some form of measurement or sample artifact rather than an effect which is explicable, as posed above, in terms of BRC threshold theory. The second problem to be addressed is the extent to which the type of research investigation conducted here was an effective means of answering the research question.

9.5.1 The Problem of Measurement and Sample Artifacts

9.5.1.1 Measurement artifacts

In the consideration of variability effects, artifactual measurement influences, and in particular the possibility of floor and ceiling effects, have been identified as important potential confounding factors to be taken into account (Christensen, et al., 1994; Salthouse, 1985; 1991; Rabbitt, 1993c). Salthouse especially has been dismissive about fallacious discussions which occur around the meaning of variability effects (although presumably with respect to brief reaction timed tasks) which he suggests are due to the influence of ceiling and floor effects of human potential. The argument is that a ceiling effect in the early adult years, and a floor effect in the late adult years, could account for narrow distributions at both of these age extremities; and the interim absence of either ceiling or floor effects in the intermediate years could arguably account for the variability bulge in the middle years.

In an evaluation of the literature review on variability the whole problem of artifacts was fully discussed (see section 4.2.1.3). It was concluded that most of the psychological tests examined for variability effects in the literature review were suitably wide-ranging enough not to be subject to truncated ceiling and floor effects (for example, Digit Symbol, Trail Making Test, long list learning tests, recall of passages, extended motor activity tests, are all relatively time-consuming tasks which call upon complex combinations of a number of functions). It was considered that in the main the tests reviewed were used for clinical purposes (as distinct from brief laboratory-type reaction time tests), and would have limited utility if they were subject to gross ceiling and floor effects. Thus it was concluded that measurement artifacts were not major contributing factors in the particular patterns of variability which were identified and upon which the shuttle model was subsequently built.

Identical arguments can be brought to bear in an evaluation of the empirical aspect of the present research, but even more strongly. Taken together, the tests included in the Phase 1 and Phase 2 analyses were wide-ranging, frequently employed clinical tests which would have limited utility if they were strongly prone to floor and ceiling effects. Specifically tests which did not incorporate a sufficiently wide range of potential levels of performance so as to preclude ceiling or floor effects (such as the Orientation and Mental Control subtests from the Wechsler Memory Scale) were not included in the analysis. Where ceiling effects appeared to be having an effect (such as for some of the verbal memory tests and recognition tasks) this was identified as an aberration compared with the regular variability effect, rather than the regular pattern in itself. Irregular variability effects were also present for tests with a large measure of inherent variability regardless of age stage (such as for the Digit Symbol Incidental Recall task, and to a lesser extent the Wisconsin Card Sorting test). Again these artifactual effects were clearly apparent as irregular effects in contradistinction to the more lawful linear or non-linear inverted-U patterns, and serve in fact to highlight clinical shortcomings of these particular tests.

When the raw scores available for all the tests included in Phase 2 of the study were subjected to the Kolmogorov-Smirnov test of normality procedure, each analysis resulted in a p-value >0.05 indicating that the assumption of normality of the distributions of the scores for those neuropsychological tests was valid. In addition an examination of the Phase 2 histograms (see Histograms 8.2a to 8.5b, pp.366-371) does not reveal the presence of the tight clustering of scores at ceiling and floor levels in the youngest and oldest groups respectively for any of the tests, which would be expected if there were significant ceiling and floor effects. Thus the more detailed examination of the distributions of the raw score data which was possible in the Phase 2 analyses (the findings of which were commensurate with the Phase 1 analyses) effectively rules out the possibility of attributing the bulge in the middle years to ceiling and floor effects at the younger and older extremities of age, respectively. Rather scrutiny of the histograms lends particularly strong support to the presence of genuinely enhanced variability effects (together in some instances with a genuine subsequent decrease in the heterogeneity of scores) *directly* due to advancing age rather than a set of mere measurement artifacts. Similarly in their studies of variability, Christensen et al. (1994) and Rabbitt (1993c) have both considered and discounted the overriding influence of artifactual ceiling or floor effects to explain the enhanced variability they have noted from their analyses of clinical test scores.

A final observation is with respect to the *remarkably* consistent patterns of variability that occurred in the present study for the same test, albeit derived from completely different studies and cohorts (for example Digit Symbol as depicted in Figures 4.8b, p.115, and Figure 8.2, p.330; RAVLT Trial 1 and Recognition as depicted in Figure 7.7a and 7.7b, pp.213 and 216). This does indicate that there is a contributing influence to the variability patterns from the idiosyncratic nature of the tests in themselves. However, these idiosyncratic test effects were clearly overridden by the more powerful overall shuttle bulge effect of enhanced variability in cognitive function due to age (and a possible later decline in variability) which pervaded across most of the idiosyncratic effects of the many tests included in the analyses. The pervasiveness of the variability effect which has been identified in this study points to a much larger influence than a mere measurement artifact. Similarly, the commonly accepted effect of declining average performance across most cognitive tasks is attributed to the progression of older age, albeit similarly with minor variations due to error variance and particular task effects.

9.5.1.2 Sample artifacts

In the consideration of variability effects, artifactual sample influences such as unbalanced numbers of subjects in groups and uneven groups in terms of demographic characteristics, have also been identified as important potential confounding factors to be taken into account (Rabbitt, 1993c). Again this problem was discussed following the review on variability (see section 4.2.1.3), and it was pointed out that in explicating the variability patterns derived from the literature overview, where sample artifacts were seen to be contributing to effects this was taken into account. Similarly this was the case for the manner in which the empirical aspect of the present study was carried out.

The limitations of a data base derived retrospectively such as was the case for the present investigation were discussed in the method (see section 6.2.2.3). Such a data base clearly does not have the potency of a prospective study which is especially designed for the particular experimental purpose. Especially it was not possible to control completely in every data set for features which are important for a variability study such as carefully balanced age ranges, a full spectrum of age ranges, balanced and sufficient numbers of subjects between-groups, and matched levels of education, IQ and gender across age groups. However, whilst not being over stringent and thereby losing too many data sources from the meta-analysis, these factors were considered up to a point whilst compiling the inclusion criteria for the data sets (see Chapter 5, Sections 2.2.2.1, and 2.2.2.2), and any obvious residual aberrations along these dimensions were taken into account in the analyses and interpretation of the results. Thus for example, in Phase 1, whilst age groups within data sets were frequently not carefully matched for education or IQ, such demographic data as were available were carefully tabled and incorporated into the understanding of the emergent variability patterns; in Phase 2 the data for the Digit Symbol test were analyzed in two modes (five group comparisons, and three group comparisons) to examine the effect as far as possible of unbalanced age ranges across groups. It is apparent therefore, that whilst the type of data collection utilized for the purposes of this research study had a number of limitations, it was not such as to prohibit the gleaning of meaningful results. Moreover, the broad spectrum of data made available for analysis in this retrospective manner, provided the breadth of meta-analysis *requisite* for the development and evaluation of a new model. It is actually the extraordinary regularity in the patterns of variability that emerged *despite* such a mixed bag of data sets, that speaks more strongly than anything else for the potency of the overall effect.

Of particular note is the fact that most of the data sets in the analyses were based on samples comprising community dwelling, healthy, relatively well-educated individuals, a narrow elitist spectrum particularly with respect to the older populations. However, as pointed out by Rabbitt (1993c) and discussed in Chapter 2 (section 2.1.1) the use of such elitist elderly samples would serve to reduce rather than exaggerate variability effects. Hence the consistently significant and robust patterns that have emerged on the basis of relatively high functioning older adults, are likely to be *markedly* more pronounced for a data collection which was generally representative of the full older population. It is of *particular* relevance, therefore, that enhanced heterogeneity due to aging was so apparent for the present set of analyses on restricted samples. It truly serves to highlight the extraordinary degree of cognitive variability induced by neural aging in interaction with vastly differential biological, social and psychological threshold-altering influences between individuals.

9.5.1.3 Research Methodology

In accordance with recommendations for the ideal mode by which to conduct an investigation into variability in association with age (Wohlwill 1973), two distinct parameters were adopted for the present study with the purpose of (i) describing the *characteristics and course* of inter-individual differences via a between-groups meta-analysis incorporating multiple studies and (ii) addressing the *nature and origin* of individual differences via within-groups analyses of the characteristics of successful and unsuccessful scorers. A meta-analysis, as was adopted in the first phase of the present study, is a methodology recommended by Craik and Jennings (1992) as the most viable means of identifying overall trends in association with aging. The analysis of characteristics of subgroups of successful and unsuccessful agers, as was adopted in the second phase of the present study, has also been repeatedly emphasized as the route to comprehending variability with aging (Krauss, 1980; Owsley, 1991; Poon, 1985; Rabbitt et al., 1993d).

Prior to the present investigation there were a very limited number of studies which focused primarily on the issue of variability in cognition, and the most promising of these were the studies of Morse (1993) and Rabbitt (1993c) (see discussions in sections 1.3 and 4.2.1). With reference to the present investigation, however, both of these research endeavours had serious shortcomings. The study of Morse (1193) is the only broad-based meta-analysis to date which has set out to identify fundamental patterns of effects with respect to cognitive variability, and consequently has achieved almost seminal status in the literature. This serves to emphasize the paucity of adequate research in the area in that Morse's study, whilst being more informative than anything else available up to now, was severely limited. First it was only concerned with addressing the first of Wohlwill's recommended angles on variability, which was to describe its course, and there was no attempt to consider the origins of the variability effects which were identified. Furthermore the analysis was restricted only to the investigation of two composite functions of fluid versus crystallized abilities (which can obfuscate effects due to the inevitable multifaceted nature of many tests) and young versus old groups (thus not following through the entire progression of variability trends across the whole spectrum of age groups).

Hence the methodology of Phase 1 of the present study provided much more adequate scope than has been previously been achieved. The broad scope of the meta-analysis across the five key functional modalities of attention, memory, visual, language and motor skills, was complemented with an analysis of a wide spectrum of age groups which enabled the tracking of variability trends across the adult life span, and allowed for the identification of the shifting manner in which variability presents depending on differential test and demographic influences. This new clarity on the phenomenon represents a significant step forward in the area. There was however an important residual limitation in the present study. This relates to exact specification about what occurs to the variability pattern at the end of the life span. Generally the older age groups available for the meta-analysis only tapped age groups up to the 80's, and often only the 70's; there were hardly any data sets which tapped performance into the 90's. Hypothetically it appears probable that, as the natural course of physical demise approaches, there would be a dedifferentiation effect in all modalities. This is particulary so in view of the phenomenon of terminal decline in cognitive function (see discussion in section 2.2.1.3). However, what occurs in the very late years up to the point of death, for those few mainly verbal, problem solving and fine hand motor tasks which did not reveal the dedifferentiation effects at the point of the oldest age groups in either the review or empirical analyses, remains empirically uncertain.

From a methodological point of view, the only other study of major interest in the variability literature was that of Rabbitt (1993c), in that it provided a prototype for the examination of those individuals that are lost within any form of group statistics. Again, however, when compared with the present study, Rabbitt's (1993c) research is limited in that it was restricted to a relatively narrow spectrum with regard to age from the 50's through to the 70's. Furthermore his examination of outliers was narrowly focused on memory function with the main purpose being to contest a hypothesis from Salthouse's (1985a; 1985b) single factor model of aging (which has been the purpose of much of Rabbitt's research), rather than the identification of broad patterns of heterogeneity across the adult life span. In contrast, with the express purpose of describing the overall patterns in adult cognitive variability, the Phase

2 within-groups analyses of the present investigation, which included within-groups analyses of demographic data, as well as the analysis of raw score distributions in conjunction with the isolation of characteristics of the outliers, served a vital complementary function together with the present between-groups analyses. On the basis of the within-groups analyses it was possible to validate the shuttle patterns identified on the between-groups analyses, as well as to make detailed observations about the characteristics of outliers which would have been lost to group statistics (including correlational analyses) due to the neutralizing effect of observed later old age opposite tendencies. Furthermore, on the basis of the complementary Phase 2 analyses of raw score distributions and outliers, it was possible to exclude hypothetical indications of artifactual ceiling and floor effects, and thereby to verify the presence of genuine *aging* shuttle bulge and shuttle shift effects, which in turn provided potent corroboration for theoretical suppositions from BRC threshold theory.

Finally, from a methodological point of view, the present study was restricted exclusively to an examination of cross-sectional normative data sets, and thus the findings are only strictly generalizable in terms of age *differences* at particular points in time, and not age *changes* developmentally with the progression of time. Accordingly inferences need to be very cautiously drawn in view of the powerful confounding potential of generational effects. For example, college graduates in their 80's may not be representative of college graduates in their 20's due to vastly different educational experiences. However there are a number of pointers that the identified variability pattern across the entire adult age range is one which would not be in evidence only for cross-sectional analyses, but would also be apparent on the basis of longitudinal analyses. First, it is difficult to explain the variability curve itself purely in terms of an artifact of cross-sectional analysis since it occurs regardless of the set of generations spanned for data sets derived at markedly different periods of time (see for example the WAIS and WAIS-R data reported in Salthouse, Figure 4.4, p. 100). Second, the actual variability bulge itself, from the point of initial increase to the point of a subsequent decrease, is expressed over a relatively limited section of the adult age range of at most a few decades (see for example the narrow space between the up and subsequent down arrows in the graphical material in the present chapter). Specifically Schaie (1988) stresses that when a small section of the age range is studied there may be no reason to suspect the confounding artifact of cohort differences. Thirdly, data derived from the seminal Seattle aging study as reported by Schaie (1988b; 1983) (see discussion in section 4.2.1.1), revealed *broadly similar variability effects for both cross-sectional and longitudinal* *modes of research*. Taken together, these observations suggest that the effect of neural aging on the pattern of inter-individual variability across the adult life span, as identified on the basis of the present investigation, may be sufficiently potent to emerge *regardless* of the unavoidable limitations inherent in the adoption of either of these two basic research approaches to cognitive aging. (The reader is referred to section 2.2.1.2 for a detailed discussion of potential confounding factors in longitudinal research into aging and variability).

9.5.2 Overall evaluation

Thus, as was the case in the evaluation following the synthesis of patterns identified from the literature review on variability, it is concluded that measurement and sample artifacts, as well as the fact that the analyses were confined only to cross-sectional analyses, do not invalidate the particular patterns of variability that were revealed from the results of the Phase 1 and Phase 2 between and within-groups empirical investigations. Similarly as for the evaluation of the review on variability trends, it is considered that if this had been the case the direction of trends should make sense conceptually according to test measurement, sample or study characteristics, whereas the patterns that emerged revealed compelling conceptual validity in terms of functional modalities and their expected manifestation in terms of neural aging processes. Moreover there was firm support for the statistical hypotheses that were derived from the shuttle model of variability, which in turn were conceptually coherent in terms of the tenets and explanations available from BRC theory.

The most positive feature of the present research investigation, taken overall from both the theoretical and empirical angles, was the development of a theoretical prism (in the form of the specific shape and mechanism posed by the shuttle model of variability) through which to untangle the conglomeration of previous findings in relation to heterogeneity in association with aging. The effect of this was to create a higher level of conceptualization to explain prior contradictory renderings of the overall picture, to produce a hypothetical basis for a systematic empirical investigation, and to provide a mechanism for integrating and explaining a set of apparently highly regular events. Above all, having established a working model from this very broadly based meta-analysis of pre-existing studies - the appropriate choice of research methodology at the level of model building - the possibility has been opened up for further more specific hypothesis testing using tightly controlled empirical and even experimental studies.

9.6 IMPLICATIONS OF THE RESULTS

9.6.1 General implications

First and foremost the results of this study have indubitably supported the presence of enhanced cognitive variability in association with aging. Any attempts to see this effect as fallacious, or to dismiss it as an artifact (for example as Salthouse, 1991, has done) is fallacious in itself. The resoundingly clear outcome of the present study on this general issue concords with the tone of the most up to date aging literature in which any doubts about the matter have finally been dispelled. Finkel et al. (1996), for example, citing Birren and Schroots in the most recent Handbook on Aging of Birren and Schaie, state that "Increases in variance of performance with age have become an accepted truth" (p.85). The more particular contribution of this thesis has been to confirm the presence of enhanced variability specifically across all the fundamental cognitive modalities of attention, verbal and visual memory, visual, language and motor skills, and shown that in no case can this be explained away on the basis of measurement or sample artifacts. Further in more differentiated manner than has previously been attempted, the present study has identified and explained the particular *form* that enhanced cognitive variability takes (the *shuttle bulge* of inter-individual variability), and the manner in which it differs in relation to particular demographic and task-related influences (the *shuttle shift* effect).

The meaning of the variability *bulge*, which has been so clearly established on the basis of the present investigation, is that *when neural aging starts to take effect, it does so in a manner which is distinctly uneven across individuals.* The findings of the present research, therefore, are highly commensurate with the viewpoint repeatedly expressed by both Rabbitt and Schaie, which is that whilst the prevalence of a general decline in mental abilities has been reliably established with advancing age on the basis of group data, *this is not true for all individuals* (Holland & Rabbitt, 1991b; Rabbitt, 1986; Rabbitt, 1993b; Schaie 1990; 1988b) (see discussion in section 4.2.1.1). Furthermore the results stand in strong contradiction to the contention of Salthouse (1991), which is that there is no particular pattern in variability effects with aging, and that variability does not increase significantly with age (see also discussions in section 4.2.1.1). Thus from the present research, which has systematically examined a broad spectrum of normative data available in the literature, it is clear that, in association with aging, cognitive function does *not* "all go together when it goes" (to use the catchy phrase of Rabbitt's keynote address, 1993b). Across the comprehensive set of functional modalities tapped in the current study (attention, memory, visual, language, problem solving and hand motor skills), due to the almost invariable appearance of the initial differentiation phase of the shuttle bulge, there was, indeed, steady evidence for Rabbitt's observation that in association with the aging process, some individuals do fall off sooner than, and disproportionately more than others.

The meaning of the subsequent dedifferentiation effect in cognitive variability, which has also been clearly established on the basis of the present research, is that for *some* cognitive functions in later old age the consequences of aging are more *even* between individuals, and that for these modalities there would a notable proportion of individuals who are no longer spared the aging effect as occurs in the initial differentiation phases of the aging process. This appears to be particularly true for more age-sensitive tasks. This later old age dedifferentiation phenomenon is an effect which has been identified in the literature (Cunningham & Tomer, 1990; Schaie, 1983; Willis, 1985), but has received very limited attention. In their texts on inter-individual variability, both Rabbitt and Schaie have focused almost entirely on the initial differentiation effect. Salthouse, it appears, misperceived the subsequent dedifferentiation effect as the *absence* of any overall increase in variability with aging. In contrast, the formulation of the shuttle model in the present thesis has served very clearly to articulate this full variability effect, including both the initial differentiation and subsequent dedifferentiation phases in association with aging.

The fact that the dedifferentiation effect is strongly in evidence for only some and not all modalities, probably explains why the more pervasive initial differentiation effect has been the focus of attention. Furthermore, the variable presence of the subsequent dedifferentiation effect explains why there has been considerable confusion in the literature about the presence or absence of heterogeneity in cognitive aging. It remains unclear whether there are functional modalities for which some individuals would sustain a disproportionately high level of ability right up to the point of demise - that is until extremely late old age. In particular from the present research it appears that this might possibly occur in the modalities of language and problem skills, and hand motor ability. However this uncertainty will need resolution via further research.

Following on the undoubted presence of significantly enhanced variability due to age, the following statement from Dannefer (1988, p. 359) has particular pertinence:

> "If age categories differ systematically in the amount of variability on a given characteristic, and especially if successive cohorts show similar patterns of variability over the life course, then it is a systematic distortion of the experience of age in a population to focus only on normative and central-tendency measures, as is usually done."

The focus on central-tendency measures has led to an overly pessimistic view of aging in which negative *differences* between the abilities of older relative to younger age groups are focused on, a point which has been most ardently endorsed by Rabbitt (1993b) and Schaie (1988a). When the heterogeneity of cognitive aging is acknowledged this forces attention on to the more positive angle of *overlapping* tendencies which exist for a sizeable proportion of older individuals when compared with their younger counterparts. Such overlapping tendencies which have been emphasized by a number of researchers (for example Schaie, 1988b; Weintraub et al., 1994), were also strongly apparent from the Phase 2 analyses in the present study, and particularly so for the timed test of hand motor dexterity right up to the 80's age group. The significant increases in variability that occur as neural aging encroaches strongly endorses the fallacy of understanding the aging process purely on the basis of normative data (as expressed by Dannefer in the quote above). Such central tendencies fail to represent the two extremes which occur of both disproportionately poor and disproportionately good scorers.

Acknowledgement of the certain increases in variability due to aging (the shuttle bulge effect) forces attention onto the investigation of the characteristics of those disproportionately successful and unsuccessful agers. This is the area of study which is increasingly being identified as probably more crucial than any other for future cognitive aging research (Rediess & Caine, 1996). In the present study the cohort-specific effects which were noted on the basis of the between-groups analyses (and supported by the within-groups analyses) provided a graphic series of indications that high education, high IQ (mostly by inference), and female gender were related to successful aging (the shuttle shift effects).

Although the number of studies available for the investigation of these cohort effects was limited, and in particular for IQ, there were a number of factors which cumulatively added weight to the validity of the findings. First, there was no data which stood in contradiction to the demonstrated direction of the cohort-specific shift effects. Second, there was a convincing level of cross-validation for educational/IQ effects when considering empirical reports from the literature review in conjunction with the results of both phases of the present investigation, and some cross-validation for gender effects across both phases of the present research. Furthermore there was a strong measure of conceptual validity in the direction of the cohort-specific shift patterns. That the high levels of education/IQ should act as protective factors against symptom onset due to neural aging, is entirely commensurate with BRC postulates of Satz (1993), in which he poses high levels of education (and by inference high levels of IQ) as protective factors against symptom onset in neurological disorder. The findings for gender are commensurate with the well-known phenomenon of greater life expectancy for females relative to males (Dulbecco, 1991; Kristof, 1996), and the report from Schaie (1994) which indicates that male gender was one of the most heavily weighted variables to predict earlier than average cognitive decline.

Importantly, however, it was also evident from the present research that high levels of education and IQ, and female gender, did not *necessarily* protect against being disproportionately poor scorers, with the implication that there are other factors which come into play which override these generally protective factors and result in opposite effects. In the case of longitudinal research this phenomenon would manifest as cross-over effects for particular individuals in association with aging (see earlier delineation of these terms, p.388). As Schaie (1988) has emphasized, these overriding influences which are responsible for such cross-over effects provide the crux of future research. They could be due to any of the multiple biological, psychological and social factors which have been identified as contributing strongly to variability in cognitive aging as discussed in Chapter 4 (section 4.2.2). These include such possibilities as particular genetic vulnerability to aging (see for example Finkel & McGue, 1993; Saudino et al., 1994; Swan et al., 1992), the potent influence of sensory deficits and other health-related vulnerabilities on cognitive function as discussed by Holland and Rabbitt (1991b), the indirect effects of related areas of cognitive deficit (Salthouse and Coon, 1994; Hultsch et al., 1990), the initial signs of an incipient Alzheimer's Dementia (Rediess & Caine, 1996; Smith et al., 1996), an aspect of the inherent cognitive dysfunction which occurs as part of normal aging particularly in memory functioning (Rediess & Caine, 1996; Smith et al., 1996), or seemingly subtle personal and social influences which nevertheless appear to predispose individuals to unsuccessful aging such as lack of generativity, generally negative attitudes towards self and others and towards life, marriage to an unstimulating partner (Olbrich, 1990; Schaie, 1994). Moreover it is likely that such vulnerability factors in opposite positive form, alone or in aggregation, might be interacting with the protective factors of high education and/or female gender (as identified on the basis of this study) to predispose individuals to successful aging as reflected in disproportionately *good* scorers. In the final analysis, as emphasized in particular by Finkel et al. (1996) and Uchiyama et al. (1996) research indicates that there is no simple explanation for the variability that occurs in association with aging, and that variables "should never be interpreted independently of each other due to their significant moderating influences" (Uchiyama et al. 1996, p. 299).

9.6.2 Social and clinical implications

The outcome from the present study which so strongly confirms the wide spread of ability in association with aging across a very wide spectrum of cognitive abilities, equally as strongly endorses the need (see Sterns & Miklos, 1995) for older individuals to be *individually* assessed in terms of their occupational and social potential, including career choices, planning as well as employment and promotional opportunities. Opposite effects (and by implication cross-over effects), imply that individuals may be impaired disproportionately in some cognitive areas whilst others are relatively spared, and such discrepancies also need to be identified on an individual basis. This would be particularly important for remedial interventions and abilities training such that older people maximize their potential, a point which has been emphasized by a number of researchers (see Hayslip, 1989; Hoyer, 1974; West, 1989). In particular, *much more emphasis needs to be directed at the upper limits of potential for older individuals in order to redress the negative stereotypes that surround aging*. Due to the degree of heterogeneity noted, the notion of *functional* age in the older populations, rather than chronological age (see Heron, 1987) becomes a vital notion in need of energetic promotion.

The implications with respect to gender from the present research are particularly potent. Across both phases of the study and a number of functional modalities including both verbal and visuo-perceptual skills, female gender emerged as a predominantly advantageous factor in cognitive aging. This finding strongly endorses the indications in the cognitive aging literature in which it appears that female advantage becomes increasingly clear with age (Heron & Chown, 1967; Robbins et al., 1994), and in particular Schaie's (1994) listing of male gender membership amongst the most highly weighted variables that predict earlier-than-average cognitive decline. Cumulatively this evidence for the earlier onset of cognitive decline due to neural aging for men is hardly surprising given their known diminished physical longevity in relation to women. As populations live longer this gap in the longevity between men and women is widening in the direction of further advantage to women (see figures cited by Kristof, 1996, from 2 years longer life expectancy for women in relation to men in 1850, to 6 years longer in 1996); and no doubt as educational and career opportunities for women continue to escalate and equalize in relation to their male counterparts, the present female cognitive advantage in later old age could become even more striking (see related discussions in Schaie 1988b; Thomae et al., 1981). On the basis of this, and from the specific indication in this thesis of earlier symptom onset for males by one to two decades when compared with women (the left shuttle effect for males relative to females), a number of *highly speculative* implications arise, as follows.

In general terms it might appear that for females the age at which retirement is routinely considered in equivalent manner to males, needs to be extended by at least 10 years. This type of consideration, however, requires much more thorough research such that the factors which cause females also to be amongst the disproportionately poor scorers are identified. Were the characteristics of cognitively unsuccessful females revealed as a combination of such elements as educational disadvantage, or having never followed a career consequent on college or high school education, or clearly identifiable negative health factors, then the reconsideration of an extended view of the appropriate age for occupational retirement of females would be indicated. On the other hand if genetic factors, the greater prevalence for females relative to males of incipient Alzheimer's Dementia, or the complex interaction of other less easily identifiable influences were in operation, this would be less feasible. Another implication which might be considered (and which is likely to be *highly* controversial amongst males in terms of the current stereotype!) is that women should advisedly choose a male marital partner at least 10 to 20 years their junior, rather than the opposite way around which stereotypically occurs. The present gender findings tentatively imply that they may in this way be more likely to have a compatible and stimulating mental companion into late old age, which in turn (see Schaie, 1990) would predispose them to yet further cognitive sparing! Overall the implementation of these practical gender-related indications, however logical in general terms, would be problematic in that they would appear highly discriminatory against males. Certainly they would be discriminatory against any isolated exceptional males who retain brilliance into late old age, and who are likely to be revealed (albeit not in as large proportions as females) in larger samples than were available for analysis in the present study. Cases in point, for instance, being statesmen Winston Churchill, Benjamin Franklin, and our own Nelson Mandela.

Finally, from a directly clinical perspective there is a crucial and undebatable message from the present research. *Normative data across the adult age spectrum should never be considered in isolation without reference to the variability data.* First, when reading normative data tables for clinical purposes, the point at which standard deviations start to widen provides confirmatory evidence for the onset of a marked decline in ability for some individuals, although this may not immediately be reflected in a declining central tendency. Overall as was evident from all the Phase 1 and Phase 2 between-groups analyses, enhanced variability occurs approximately in parallel with steadily declining average ability. Whilst the matching of the exact points of change between mean scores and standard deviations was not pursued as a central aspect of the present statistical investigation, in some instances it was clearly observable that variability increased dramatically *prior* to the onset of markedly diminished average ability. This occurred, as was noted above, for tasks of particularly high challenge (PASAT 1.2 second presentation, and TMT B, Figures 9.5a and 9.5b, respectively), and is an effect about which clinicians need to be vigilant. Au et al. (1995), with reference to their normative data set for the Boston Word Naming Test, report that for their data there was only a *significant* decline in mean scores by the *70's* age group. However from the present analysis of the same Au et al. data, it was evident from progressive comparisons of the standard deviations, that there was a steady increase in variability across the adult age range (see Table 7.8, p.278), which reached significance by the much earlier *50's* age decade. (The reader is reminded that the present study did not make its own progressive comparisons of means). Putting the finding of an early significant increase in variability, together with the Au et al. report of late average decline, indicates that a proportion of individuals lose word naming function disproportionately sooner than is generally depicted by the progression of the mean - an observation which is commensurate with the common anecdotal complaint from middle-aged individuals that they are aware of word finding difficulties. Thus, by making reference to the progression of standard deviations in relation to norms, important clinical information becomes available; conversely by *not* doing so, vital indications about cognitive aging are lost.

Second, once variability increases significantly as older age progresses, the indications from mean scores are distinctly unreliable. Specifically on the basis of the present study this occurrence has been confirmed for the entire spectrum of basic functional modalities (as defined by Spreen & Strauss, 1991, and Lezak, 1995), with the exception only of the least age sensitive modalities of pure visuo-spatial function and visual recognition. Hence in clinical assessments for older individuals, especially where declining ability is subtle, the use of normative data where there is an exceptionally wide spread of scores has minimal validity. Such practice in the case of medical diagnosis, head injury claims, or questions of the instatement or release of a curator bonus, could lead to aberrations of clinical judgement, which in turn may result in serious clinical and social consequences. A route around this dilemma is to manage such evaluations of older individuals as closely as possible in relation to their own highly individualized premorbid cognitive and behavioural indications. The use of normative data should be more cautiously employed than with younger age groups, paying conscientious attention to the variability data. Clearly in terms of the *differential* perspective promoted in this thesis, *the mode of test interpretation adopted for the assessment of older populations needs to be urgently reviewed and treated as an area for special management in its own right*. Assessment of older adults should not be grouped in an integrated manner with the rest of adult assessment for which normative data routinely acts as a key yardstick in the interpretation process. Such a review of the situation is particularly pressing when considered in association with the exponential rise in the aging population (cited by Kristof, 1996, as 30% in the 1990's). Highly *individualized* assessment of these older individuals, who are increasing in disparity in a fast-changing world, will become a necessity.

9.6.3 Implications for future research

The outcome of the present investigation highlights several routes for future research. In the first instance additional studies should be carried out in similar manner to the present study, merely examining the vast quantities of data already derived from aging populations for norming purposes. Such retrospective research would serve to supplement the current findings, and to refine the model that has been presented. It should focus in particular on the identification of variability patterns across those data sets for which *very late old age* *groups* are available, the absence of which was a limitation in the present study and prohibited certain definition of the latter limits of the model. The examination of subgroups of outliers in terms of their characteristics should be conducted (as was done in the present study) as far as is possible with retrospective data banks. Further it would be valuable to examine data sets for correspondence between the actual *points* of significant onset of enhanced variability and declining average performance since this reflects on the reliability of mean scores for clinical purposes. (The present study was limited in that it only examined the mean scores for evidence of an *overall* significant decline). Finally in terms of research with existing data, the examination of longitudinal data sets in addition to more cross-sectional data would be essential to add substance to the present findings.

The most important avenue for future research would be to design studies with the specific purpose of investigating variability, such that potential confounding variables can be properly controlled. This would serve to preclude artifactual effects and reveal true effects more definitively. For instance, it is likely that the shuttle shift effect due to differential levels of challenge between Trail Making Test Part A and Part B was revealed in the Shuttleworth-Jordan and Cornfield (1996) data set for which educational levels between-groups *were* well-controlled, whereas this effect was not apparent for the other two data sets examined (of Stuss et al., 1987 and van Gorp et al. 1990) which were *not* strictly controlled for these variables between age groups. Matched numbers of subjects for gender would be an important feature to control so as to scrutinize the major implications that have arisen in relation to male disadvantage from the present thesis, and as implied from Schaie's (1994) report. In addition it is imperative that over and above gender, such future studies should focus carefully on the psychological, social and physical characteristics of their subjects. This would be in order to identify those characteristics which contribute to successful versus unsuccessful aging, and especially to unravel those features which contribute to opposite effects in cross-sectional research, and cross-over effects in longitudinal research. An appropriate method for such research would be to derive detailed unique life histories from the subjects (as suggested by Dixon, 1992). Alternatively probing biographical questionnaires which address all possible known influences could be administered pre-experimentally (for example Thomae, 1990). The most ideal procedure would be to combine a structured biographical questionnaire with a detailed life history gleaned post-experimentally (as suggested by Perlmutter, 1978), so that specific avenues revealed in the characteristics of the test scores could be followed up more exactly. In order to identify those individuals who were at risk for a subsequent dementia, or in a 'silent' stage of terminal decline, longitudinal follow-up would also be imperative. In the view of Rowe (as cited in Crowley, 1996) the crucial question for future aging research is how to convert the "usual aging" group into the "successful" aging group and prevent them from experiencing intellectual decline. This is a question that can only be answered by pointedly searching research of the post hoc biographical type, and longitudinal follow-up.

Finally critical experiments could be designed to test the shuttle mechanisms of the variability model. Carefully matched samples could be compared with respect to the earlier and later onset of increased variability as a function of age in interaction with a range of influences including those already investigated here (that is education, IQ and gender). Other factors which might lower the threshold for symptom onset such as previous head injury, stroke, alcoholism, chronic life stress, major depression, or induced physical fatigue, alone and in aggregated form, could be investigated for the shuttle effect, all in association with the added cumulative stressor of various gradations of advancing neural attrition due to age. Hot off the press with regard to the aggregation of such threshold-lowering effects, is the highly topical report in Time magazine (Nash, March 1997), in which the work of epidemiologist David Snowdon from the University of Kentucky (USA), is described. Reportedly, out of a group of 61 deceased nuns, whose brains showed Alzheimer's disease at autopsy, 93% with a history of stroke developed clinical signs of dementia. In contrast, only 57% of stroke-free nuns showed functional signs of dementia. In Snowdon's research the influence of gender is automatically controlled. However, an important question that arises from the outcome of the present research, is whether the educational and IQ levels of the nuns, their chronological age, and/or other subtle yet significant moderating influences may be contributing to these effects. (Time constraints and limited library facilities have precluded access to Snowdon's actual research report at this point, thus it is not possible to comment on the extent to which account has been taken of such potentially crucial intervening effects in this Alzheimer's autopsy research).

An example of a threshold-related investigation, which has been specifically designed to test the shuttle model, is currently in progress at Rhodes University (South Africa) under the coordination of the present author. Routinely, physical stress tests are administered to top-level sportsmen at the South African Sports Science Institute in Cape Town, with the purpose of conducting physical fitness assessments, following which individuals are extremely fatigued and struggle to concentrate. Taking advantage of this, the Rhodes University research team are conducting a series of cognitive tests pre- and post a physical stress exercise with national level players of rugby (a contact sport which is associated with a high incidence of concussive and subconcussive head injuries). The data for the rugby players will be compared with cognitive test data collected under the same pre- and post physical stress conditions from national level players of cricket (a non-contact sport which is not associated with regular exposure to head trauma). Preliminary outcome from this rugby head injury research indicates that these national level rugby and cricket players, generally, are of superior intellectual potential. This can be seen a protective factor which, cumulatively with the additional protective factor of relatively young age, is likely to raise the threshold of symptom onset. However, in terms of BRC theory, and indications from this thesis on cognitive variability, it is anticipated that whether or not the post-stress testing precipitates significantly greater lowering for the rugby players compared with the cricketers as indicated by mean scores, it will reveal enhanced variability for the former group. This would be expected due to the potentiation of individual differences in the more vulnerable rugby playing (mild head injury) group, in aggregation with the superimposed brain stressor of induced physical fatigue. A further study is proposed, in which comparisons of cognitive tests scores will be made between ex-national rugby players and ex-national cricket players across the adult life span (either via cross-sectional analysis or longitudinal follow-up). It is anticipated that these data will reveal the shuttle shift effects of an *earlier* onset of enhanced variability for the rugby players compared with the cricket players, due to the aggregation of the threshold-lowering influences of multiple exposure to mild concussive and/or subconcussive head injuries, and aging. Given sufficiently high challenge of task, it is expected that increased variability may *signal* the onset of cognitive difficulties at an earlier age stage amongst rugby players compared with cricketers, before such disadvantage is reflected in comparisons of the central trends.

9.7 THE SHUTTLE MODEL: WIDER IMPLICATIONS

A model has been developed in the present thesis which depicts the apparently highly lawful recurring pattern of cognitive variability which occurs in association with aging. Progress in science advances by the development, refinement and possibly even the refutation of models (Kuhn, 1970), and from that point of view this thesis can be said to have served, and may continue to serve, an important heuristic purpose in the field of cognitive aging. It would seem very probable, however, that the shuttle model (including the bulge in variability and shifting mechanism of effects) as formulated here, is not exclusive to the process of cognitive aging. Rather it is likely to represent part of a more all-encompassing natural law that would apply with the onset and progression of any insidious disease process, or slowly progressive disruptive influence on a previously stable entity, in interaction with relevant influences which differentially raise or lower the thresholds of entropy until the point of restabilization. The wide array of possibilities for further research that flow from such a supposition, including the formulation of predictive mathematical equations for various phenomena and related disruptive forces, go beyond the scope of the present thesis.

REFERENCES

Albert, M.S. (1981). Geriatric neuropsychology. *Journal of Consulting and Clinical Psychology, 49*(6), 835-850.

Albert, M.S., Duffy, F. H., & Naeser, M. (1987). Nonlinear changes in cognition with age and their neuropsychologic correlates. Special Issue: Aging and cognition. *Canadian Journal of Psychology, 41*(2), 141-157.

Anastasi, A. (1958). *Differential psychology* (3rd ed.). New York: The Macmillan Company.

Anastasi, A. (1991). The gap between experimental and psychometric orientations. *Journal of the Washington Academy of Sciences, 81*(2), 61-73.

Arenberg, D. & Robertson-Tehabo, E.A. (1977). In In J.E. Birren, & K.W. Schaie (Eds.), *Handbook of the psychology of aging,* (pp. 421-449). New York: Van Nostrand Reinhold Company.

Armitage, P. (1971). *Statistical methods in medical research.* Oxford: Blackwell Scientific Publications.

Arsanow, R.F., Steffy, R.A., MacCrimmon, D.J., & Cleghorn, J.M. (1977). An attentional assessment of foster children at risk for Schizophrenia. *Journal of Abnormal Psychology. 86*(3), 267-275.

Au, R., Joung, P., Nicholas, M., Obler, L.K., Kass, R., and Albert. M.L. (1995). Naming ability across the adult life span. *Aging and Cognition. 2*(4), 300-311.

Axelrod, B.N., & Henry R.R. (1992). Age-related performance on the Wisconsin Card Sorting, Similarities, and Controlled Oral Word Association tests. *The Clinical Neuropsychologist, 6*(1), 16-26.

Baddeley, A. (1987). *Working Memory*. Bristol: J.W. Arrowsmith Ltd.

Beatty, W.W. (1993). Age differences on the California Card Sorting Test: Implications for the assessment of problem solving by the elderly. *Bulletin of the Psychonomic Society, 31*(6), 511-514.

Berg, C.A. (1992). Perspectives for viewing intellectual development throughout the life course. In R.J. Sternberg & C.A. Berg (Eds.), *Intellectual development* (pp. 1-15). New York: Cambridge university Press.

Berlot, H.H. & Volk, B. (1980) Age-related microheterogeneity of myelin basic protein isolated from human brain. In L. Amaducci, A.N. Davison, & P. Antuono (Eds.), *Aging of the brain and dementia,* (81-90). New York: Raven Press.

Bidell, T.R., & Fischer, K.W. (1992). Perspectives for viewing intellectual development throughout the life course. In R.J. Sternberg & C.A. Berg (Eds.), *Intellectual development* (pp. 100-140). New York: Cambridge university Press.

Bieliauskas, L.A. (1993). Depressed or not depressed? That is the question. Presidential address of the 21st annual International Neuropsychological Society meeting (1993, Galveston, Texas). *Journal of Clinical and Experimental Neuropsychology, 15* (1), 119-134.

Bieliauskas, L.A., & Lamberty, G.J. (1995). Simple reaction time and depression in the elderly. *Aging and Cognition, 2*(2), 128-131.

Birren, J.E., & Bengtson, V.L. (Eds.). (1988). *Emergent theories of aging*. New York: springer Publishing Co.

Birren, J.E., & Birren, B.A. (1990). The concepts, models, and history of the psychology of aging. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 3-20). New York: Academic Press, Inc.

Birren, J.E., & Cunningham, W.R. (1985). Research on the psychology of aging: Principles, concepts and theory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 3-34). New York: Van Nostrand Reinhold Company.

Birren, J.E. & Renner, V.J. (1977). Research on the psychology of aging: Principles and experimentation. In J.E. Birren, & K.W. Schaie (Eds.), *Handbook of the psychology of aging,* (pp. 3-38). New York: Van Nostrand Reinhold Company.

Birren, J.E., & Schaie, K.W. (Eds.). (1977). *Handbook of the psychology of aging*. New York: Van Nostrand Reinhold Company.

Birren, J.E., & Schaie, K.W. (Eds.). (1985). *Handbook of the psychology of aging* (2nd ed.). New York: Van Nostrand Reinhold Company.

Birren, J.E., & Schaie, K.W. (Eds.). (1990). *Handbook of the psychology of aging* (3rd ed.). New York: Academic Press, Inc.

Black, J.E., Greenough, W.T., Anderson, B.J., & Isaacs, K.R. (1987). Environment and the aging brain. Special Issue: Aging and cogntion. *Canadian Journal of Psychology, 41*(2), 111-130.

Blanchard-Fields, F., & Camp, C.J. (1990). Affect, individual differences, and real world problem solving across the adult life span. In T.M. Hess, (Ed.), *Knowledge organization and utilization. Advances in psycyology, 71* (pp. 461-497). Amsterdam, Netherlands: Elsevier Science Publishers B.V. (North-Holland).

Bode, S. G. & Shuttleworth-Jordan, A.B. (1993). The age effect on the SAWAIS Digit Symbol substitution and Digit Symbol Incidental Recall Tests. In R. Plunkett & S. Anderson (Eds.). *Proceedings of the 5th National Neuropsychology Conference*. South African Clinical Neuropsychology Association (SACNA).

Bolla, K.I., Lindgren, K.N., Bonaccorsy, C., & Bleecker, M.L. (1990). Predictors of verbal fluency (FAS) in the healthy elderly. *Journal of clinical Psychology, 46*(5), 623-628.

Bolla-Wilson, K., & Bleecker, M.L. (1986). Influence of verbal intelligence, sex, age, and education on the Rey Auditory Verbal Learning Test. *Developmental Neuropsychology, 2*(3), 203-211.

Boller, F. & Grafman, J. (Eds.). (1990). *Handbook of neuropsychology. Vol. 4.* The commisuratized brain; Aging and dementia. Amsterdam: Elsevier Science Publications.

Boller, F. & Grafman, J. (Eds.). (1991). *Handbook of neuropsychology. Vol. 5.* Aging and dementia; methodological and practical approaches. Amsterdam: Elsevier Science Publications.

Bondareff, W. (1985). The neural basis of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 95-112). New York: Van Nostrand Reinhold Company.

Bornstein, R., & Smircina, M.T. (1982). The status of the empirical support for the hypothesis of increased variability in aging populations. *Gerontologist, 22*(3), 258-260.

Botwinick, J. (1977). Intellectual abilities. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging,* (pp. 580-605). New York: Van Nostrand Reinhold Company.

Brayne, C. (1994). How common are cognitive impairment and dementia? An epidemiological viewpoint. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.), *Dementia and normal aging* (pp. 167 - 207). Cambridge, England: Cambridge University Press.

Brayne, C., &, Calloway, P. (1988). Normal ageing, impaired cognitive function, and senile dementia of the Alzheimer's type: a continuum? *Lancet ii*, 1265-6.

Brown, M.B., & Forsythe, A.B. (1974). Robust tests for the equality of variances. *Journal of American Statistical Association, 69*, 364-367.

Charnass, N., & Bosman, E.A. (1990). Expertise and aging: life in the lab. In In T.M. Hess (Ed)., *Aging and cognition: Knowledge organization and utilization,* (pp. 343-385). North-Holland: Elsevier Science Publishers B.V.

Charness, N., & Bosman, E.A. (1992). Human factors and age. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 495-551). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Christensen, H., Mackinnon, A., Jorm, A.F., Henderson, A.S., Scott, L.R., & Korten, A.E. (1994). Age differences and interindividual variation in cognition in community-dwelling elderly. *Psychology and Aging, 9*(3), 381-390.

Clarkson-Smith, L., & Hartley, A.A. (1989). Relationships between physical exercise and cognitive abilities in older adults. *Psychology and Aging, 4*(2), 183-189.

Cohen, G., Conway, M.A., & Maylor, E.A. (1994). Flashbulb memories in older adults. *Psychology and Aging, 9*, 454-463.

Cornfield, S.P., & Shuttleworth-Jordan, A.B. (1996). *Trail Making Test I: The age effect on a non-clinical population of high functioning older adults*. Manuscript in preparation.

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 Reinhold Company.

Craik, F.I.M., & Jennings, J.M. (1992). Human Memory. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51-109). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Craik, F.T.M., & Salthouse, T.A. (Eds.). (1992). *The handbook of aging and cognition*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Crook, T., Bartus, R.T., Ferris, S.H., Whitehouse, P., & Cohen, G.D. (1986). Age-associated memory impairment: Proposed diagnostic criteria and measures of clinical change - report of a National Institute of Mental Health work group. *Developmental Neuropsychology, 2* (4), 261-276.

Cronin-Golomb, A., (1990). Abstract thought in aging and age-related neurological disease. In F. Boller & J. Grafman (Eds), *Handbook of Neuropsychology, Vol. 4. The commisuratized brain; Aging and dementia*, (pp. 279-309). Amsterdam: Elsevier Science Publications.

Crowley, S.L. (1996). Aging brain's staying power. *AARP Bulletin, 37*(4), pp. 1 and 15. (A publication of the American Association of Retired Persons).

Cullum, C.M., Butters, N., Troster, A.I., & Salmon, D.P. (1990). Normal aging and forgetting rates on the Wechsler Memory Scale-Revised. *Archives of Clinical Neuropsychology. 5*, 23-30.

Cunningham, W.R., & Birren, J.E. (1980). Age changes in the factor structure of intellectual abilities in adulthood and old age. *Educational and Psychological Measurement, 40* (2), 271-290.

Cunningham, W.R., & Tomer, A. (1990). Intellectual abilities and age: Concepts, theories and analyses. In E.A. Lovelace (Ed.), *Aging and cognition: Mental processes, self-awareness, and interventions. Advances in psycholgy, 72* (pp.379-406). North-Holland, Amsterdam, Netherlands: Elsevier Science Publishers.

Dannefer, D. (1988). What's in a name? An account of the neglect of variability in the study of aging. In J.E. Birren & V. L. Bengtson (Eds.), *Emergent theories of aging* (pp. 356-384). New York: Springer Publishing Co., Inc.

De Kosky, S.T., & Bass, H. (1980). Effects of aging and senile dementia on the microchemical pathology of human cerebral cortex. In L. Amaducci, A.N. Davison, & P. Antuono (Eds.), *Aging of the brain and dementia*, (33-37). New York: Raven Press.

D'Elia, L.G., Boone, K.B. & Mitrushina, A.M., (in press). *Handbook of normative data for neuropsychological assessment*. New York: Oxford University Press.

Denckla, M.B. (1973). Development of speed in repetitive and successivefinger-movements in normal children. *Developmental Medical Child Neurology, 15*, 635-645.

Diamond, M.C., & Connor, J.R. (1982). Plasticity of the aging cerebral cortex. In S. Hoyer (Ed.), *The aging brain. Physiological and pathophysiological aspects,* (pp. 36-44).

Dixon, R.A. (1992) Contextual approaches to adult intellectual development. In R.J. Sternberg & C.A. Berg (Eds.), *Intellectual development* (pp. 350-380). New York: Cambridge University Press.

Dolen, L.S., & Bearison, D.J. (1982). Social interaction and social cognition in aging: a contextual analysis. *Human Development, 25*(6), 430-442.

Dulbecco, R. (Ed.). (1991). *Encyclopaedia of human biology. Volume 4.* New York: Academic Press, Inc.

Eckstrand, G.A. (1962). Individuality in the learning process: some issues and implications. *Psychological Record, 12* (4), 405-416.

Esiri, M.M. (1994). Dementia and normal aging: neuropsathology. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.). *Dementia and normal aging* (pp. 385-436). Cambridge, England: Cambridge University Press.

Ferris, S. H., & Kluger, A. (1996). Commentary on Age-Associated Memory Impairment, Age-Related Cognitive Decline and Mild Cognitive Impairment. *Aging, Neuropsychology, and Cognition, 3*(2), 148-153.

Filoteo, J.V., Delis, D.C., Massman, P.J., & Butters, N. (1994). Visuospatial dysfunction in dementia and normal aging. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.). *Dementia and normal aging* (pp. 366-381). Cambridge, England: Cambridge University Press.

Finkel, D., & McGue, M. (1993). The origins of individual differences in memory among the elderly: A behavior genetic analysis. *Psychology and Aging, 8*(4), 527-537.

Finkel., D., Pedersen, N.L., McClearn, G.E., Plomin, R., & Berg, S. (1996). Cross-sequential analysis of genetic influences on cognitive ability in the Swedish Adoption/twin study of aging. *Aging, Neuropsychology, and Cognition, 3*(1), 84-99.

Folstein, M.F., 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*(3), 189-198.

Fowler, F.G., & Fowler, H.W. (1959). *The pocket Oxford dictionary of current English.* Oxford: Clarendon Press.

Fozard, J.L., Vercruyssen, M., Reynolds, S.L., & Hancock, P.A. (1994). Age differences and changes in reaction time: The Baltimore Longitudinal Study of Aging. *Journals of Gerontology, 49*(4), 179-189.

Fromm, D., & Holland, A.L. (1989). Functional communication in Alzheimer's disease. *Journal of Speech and Hearing disorders, 54*(4), 535-540.

Gold, D.P., & Arbuckle, T.Y. (1990). Interactions between personality and cognition and their implications for theories of aging. In E.A. Lovelace (Ed.), *Aging and cognition: Mental processes, self-awareness, and interventions. Advances in psychology* (pp. 351-377). Amsterdam, The Netherlands: Elsevier Science Publishers.

Goul, W.R. & Brown, M. (1970). Effects of age and intelligence on Trail Making Test performance and validity. *Perceptual and Motor Skills, 30*(1), 319-326.

Hartley, J.T. (1988). Aging and individual differences in memory for written discourse. In L.L. Light & D. M. Burke (Eds.), *Language, memory, and aging* (pp. 36-57). New York: Cambridge University Press.

Hartley, A.A. (1992). Attention. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 3-49). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Hayslip, B. (1989). Fluid ability training with aged people: A past with a future? *Educational Gerontology, 15*(6), 573-595.

Hayslip, B., & Sterns, H.L., (1979). Age differences in relationships between crystallized and fluid intelligences and problem solving. *Journal of Gerontology, 34*(3), 404-414.

Heaton, R.K., Grant, I., & Matthews, C.G. (1986). Differences in neuropsychological test performance associated with age, education, and sex. In I. Grant & K.M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 100-119).

Heaton, R.K., Grant, I., & Matthews, C.G. (1991). *Comprehensive norms for an expanded Halstead-Reitan battery.* U.S.A.: Psychological Assessment Resources, Inc.

Henderson, A.S., & Sartorius, N. (1994). International criteria and differential diagnosis. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.). *Dementia and normal aging* (pp. 79-90). Cambridge, England: Cambridge University Press.

Heron, A. (1987). The concept of functional age. In L. Lennart (Ed.), *Society, stress, and disease* (pp. 31-37). Oxford: Oxford University Press.

Heron, A., & Chown, S. (1967). *Age and Function.* Boston: Little, Brown and Company.

Hindmarch, I. (1993). Age-associated memory impairment: real entity or statistical artefact? In R. Levy, R. Howard & A. Burns (Eds.), *Treatment and care in old age psychiatry* (pp. 77-93). Petersfield, UK and Bristol, PA, USA: Wrightson Biomedical Publishing Ltd.

Hinkin, C., Cummings, J.L., Van Gorp, W.G., Satz, P., Mitrushina, M., & Freeman, D. (1990). Frontal/Subcortical features of normal aging. *Canadian Journal on Aging, 9* (2), 104-119.

Hodges, J.R. (1994). Neurological aspects of dementia and normal aging. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.), *Dementia and normal aging* (pp. 118 - 129). Cambridge, England: Cambridge University Press.

Holland, A.L. (1990). Research methodology: I. Implications for speech-language pathology. Research Symposium: Communication sciences and disorders and aging (1988, Washington, DC). *ASHA Reports Series (American Speech-Language-Hearing Association),* 19, 35-39.

Holland, C.A., & Rabbitt, P.M. (1991a). Ageing memory: use versus impairment. *British Journal of Psychology, 82*(1), 29-38.

Holland, C.A., & Rabbitt, P.M. (1991b). The course and causes of cognitive change with advancing age. Social and psychological gerontology. *Reviews in Clinical Gerontology,* 1, 81-96.

Horn, J.L., & Cattell, R.B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*(2), 107-129.

Horn, J.L., & Hofer, S.M. (1992). Major abilities and development in the adult period. In R.J. Sternberg & C.A. Berg (Eds.), *Intellectual development* (pp. 44-99). New York: Cambridge University Press.

Howard, R. (1993). How benign is benign senescent forgetfulness? In R. Levy, R. Howard & A. Burns (Eds.), *Treatment and care in old age psychiatry* (pp. 135-138). Petersfield, UK and Bristol, PA, USA: Wrightson Biomedical Publishing Ltd.

Hoyer, W.J. (1974). Aging as intraindividual change. *Developmental Psychology, 10*(6), 821-826.

Huff, F.J. (1990). Language in normal aging and age-related neurological diseases. In F. Boller & J. Grafman (Eds), *Handbook of Neuropsychology, Vol. 4. The commisuratized brain; Aging and dementia,* (pp. 251-264). Amsterdam: Elsevier Science Publications.

Hultsch, D., Hertzog, C., & Dixon, R.A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychology and Aging, 5*(3), 356-368.

Huppert, F.A. (1991). Age-related changes in memory: learning and remembering new information. In F. Boller & J. Grafman (Eds), *Handbook of Neuropsychology, Vol. 5 Aging and dementia; methodological and practical approaches* (pp. 123 -147). Amsterdam: Elsevier Science Publications.

Huppert, F.A., & Brayne, C. (1994). What is the relationship between dementia and normal aging. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.), *Dementia and normal aging* (pp. 3-11). Cambridge, England: Cambridge University Press.

Huppert, F.A., Brayne, C., & O'Connor, D.W. (Eds.). (1994). *Dementia and normal aging*. Cambridge, England: Cambridge University Press.

Huyck, M.H. (1990). Gender differences in aging. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 124-132). New York: Academic Press, Inc.

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, R.C., Petersen, E.K., & Kurland, L.T. (1992). Mayo's Older Americans Normative Studies (MOANS). WAIS-R, WMS-R, and AVLT norms for ages 56 through 97. *The Clinical Neuropsychologist. 6* (Supplement).

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, R.C., & Petersen, E.K. (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE Token, WRAT-R Reading, AMNART, Stroop, TMT, and JLO. *The Clinical Neuropsychologist, 10*(3), 262-278.

Ivnik, R.J., Smith, G.E., Malec, J.F., Kokmen, E., & Tangalos. (1994) Mayo cognitive factor scales: distinguishing normal and clinical samples by profile variability. *Neuropsychology 8*(2), 203-209.

Ivnik, R.J., Smith, G.E., Petersen, E.K., Kokmen, E., & Tangalos, E.G. (1995). Long-term stability and intercorrelations of cognitive abilities in older persons. *Psychological Assessment, 7*, 155-16.

Ivy, O.G., Macleod, C.M., Petit, T.L., & Markus, E.J. (1992). A physiological framework for perceptual and cognitive changes in aging. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 273-313). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Johansson, B. (1991). Neuropsychological assessment in the oldest-old. International psychogeriatric association workshop. *International Psychogeriatrics, 3*(Suppl), 51-60.

Kausler, D.H. (1990). Motivation, human aging, and cognitive performance. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 171-182). New York: Academic Press, Inc.

Kemper, S. (1992). Language and aging. The organization of memory in old age. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 213-270). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Kempler, D., & Zelinski, E.M. (1994). Language in dementia and normal aging. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.), *Dementia and normal aging* (pp. 331-365). Cambridge, England: Cambridge University Press.

Kirkwood, T.B.L. (1994). How do risk factors for dementia relate to current theories on mechanisms of aging? In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.), *Dementia and normal aging* (pp. 230-243). Cambridge, England: Cambridge University Press.

Kliegl, R., Smith, J., & Baltes, P.B. (1989). Testing-the-limits and the study of adult age differences in cognitive plasticity of mnemonic skill. *Developmental Psychology. 25*(2), 247-256.

Kral, V.A. (1966). Memory loss in the aged. *Diseases of the nervous system, 27* (Suppl.), 51-54.

Kral, V.A. (1970). Clinical contribution towards an understanding of memory function. *Diseases of the Nervous System, 31* (1), 23-29.

Krauss, I.K. (1980). Between- and within-group comparisons in aging research. In L.W. Poon (Ed.). *Aging in the 1980's: Psychological Issues.* Washington, D.C.: American Psychological Association.

Kristof, N.D. (1996). *Aging world, new wrinkles.* The New York Times, Sunday, September 22, pp. 1 and 5.

Kuhn, T.S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: Chicago University Press.

Labouvie-Vief (1985). Intelligence and cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp.500-530). New York: Van Nostrand Reinhold Company.

Lamberty, G.J., & Bieliauskas, L.A. (1993). Distinguishing between depression and dementia in the elderly: A review of neuropsychological findings. *Archives of Clinical Neuropsychology, 8* (2), 149-170.

Larrabee, G.J., & Crook, T.H. III. (1994) Estimated prevalence of age-associated memory impairment derived from standardized tests of memory function. *International Psychogeriatrics, 6* (1), 95-104.

Lehr, U. (1980). Aging status and processes of aging-biographical determinants. *Zeitschrift fur Gerontologie, 13*(5), 442. Abstract.

Lenzi, G.L., & Jones, T. (1980). Cerebral metabolism-to-blood flow relationships with respect to aging and dementia. In L. Amaducci, A.N. Davison, & P. Antuono (Eds.), *Aging of the brain and dementia,* (99-102). New York: Raven Press.

Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lezak, M.D., and Gray, D.K. (1984). Sampling problems and nonparametric solutions in clinical neuropsychological research. *Journal of Clinical Neuropsychology, 6*(1), 101-109.

Light, L.L. (1990). Interactions between memory and language in old age. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 275-290). New York: Academic Press, Inc.

Light, L.L. (1992). The organization of memory in old age. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 111-165). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Light, L.L., & Burke, D.M. (1988). Patterns of language and memory in old age. In L.L. Light & D. M. Burke (Eds.), *Language, memory, and aging* (pp. 244-271). New York: Cambridge University Press.

Lindenberger, U., & Baltes, P.B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9*(3), 339-355.

Lishman, W.A. (1994). The history of research into dementia and its relationship to current concepts. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.), *Dementia and normal aging* (pp. 41-56). Cambridge, England: Cambridge University Press.

Lovelace, E.A. (1990). Cognitive aging: A summary overview. In E. A. Lovelace (Ed.), *Aging and cognition: Mental processes, self-awareness, and interventions. Advances in psychology* (pp. 407-434). Amsterdam, The Netherlands: Elsevier Science Publishers.

Maddox, G.L., & Douglass, E.B. (1974). Aging and individual differences: a longitudinal analysis of social, psychological, and physiological indicators. *Journal of Gerontology, 29*(5), 555-563.

Maylor, E.A. (1991). Recognizing and naming tunes: Memory impairment in the elderly. *Journal of Gerontology: Psychological Sciences, 46*, 207-217.

Maylor, E.A., & Rabbitt, P.M. (1994). Applying Brinley plots to individuals: effects of aging on performance distributions in two speeded tasks. *Psychology and Aging, 9*(2), 224-30.

McDowd, J.M., & Birren, J.E. (1990). Aging and attentional processes. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 222-233). New York: Academic Press, Inc.

Miller, R. (1981). *Simultaneous inference* (2nd Ed.). New York: Springer-Verlag.

Mitrushina, M., Patel, P.G., Satz, P., D'Elia, L., & McConnell, M.J. (1989). Changes in semantic memory processing in normal and at-risk aging adults. *Developmental Neuropsychology, 5*(4), 321-333.

Mitrushina, M., & Satz, P. (1989). Differential decline of specific memory components in normal aging. *Brain Dysfunction. 2*, 330-335.

Mitrushina, M., & Satz, P. (1991a). Stability of cognitive functions in young-old versus old-old individuals. *Brain Dysfunction, 4*(4), 174-181.

Mitrushina, M., & Satz, P. (1991b). Changes in cognitive functioning associated with normal aging. *Archives of Clinical Neuropsychology, 6*(1-2), 49-60.

Mitrushina, M., & Satz, P. (1995). Base rates of the WAIS-R intersubtest scatter and VIQ-PIQ discrepancy in normal elderly. *Journal of Clinical Psychology, 51*(1), 71-78.

Mitrushina, M., Satz, P., Chervinsky, A., & D'Elia, L. (1991). Performance of four age groups of normal elderly on the Rey Auditory-Verbal Learning Test. *Journal of Clinical Psychology. 47*(3), 351-357.

Mitrushina, M., Satz, P., Drebing, C.E., Van Gorp, W., Mathews, A., Harker, J., Chervinsky A. (1994). The differential pattern of memory deficit in normal aging and dementias of different etiology. *Journal of Clinical Psychology, 50*(2), 246-252.

Mitrushina, M., Satz, P., & Van Gorp. (1989). Some putative cognitive precursors in subjects hypothesized to be at-risk for dementia. *Archives of Clinical Neuropsychology, 4*(4), 323-333.

Mitrushina, M., Uchiyama, C., & Satz, P. (1995). Heterogeneity of cognitive profiles in normal aging: implications for early manifestations of Alzheimer's Disease. *Journal of Clinical and Experimental Neuropsychology, 17*(3), 374-382.

Montgomery, R.J., & Borgatta, E.F. (1986). Plausible theories and the development of scientific theory: the case of aging research. *Research on Aging, 8*(4), 586-608.

Morris, R.G., & Kopelman, M.D. (1986). The memory deficits in Alzheimer-type dementia: A reveiw. *The Quarterly Journal of Experimental Psychology, 38*(4-A), 575-602.

Mortimer, J.A. (1994). What are the risk factors for dementia? In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.), *Dementia and normal aging* (pp. 209-229). Cambridge, England: Cambridge University Press.

Mortimer, J.A., van Duijn, C.M., Chandra, V., Fratiglioni, L., Graves, A.B., Heyman, A., Jorm, A..F., Kokmen, E., Kondo, K., Rocca, W.A., Shalat, S.L., Soininen, H., Hofman, A. (1991). Head trauma as a risk factor for Alzheimer's disease: a collaborative re-analysis of case-control studies. *International Journal of Epidemiology, 20*(2), Suppl. 2, S28-35.

Morse, C.K., (1993). Does variability increase with age? An archival study of cognitive measures. *Psychology and Aging, 8* (2), 156-164.

Moscovitch, M. & Winocur, G. (1992). The neuropsychology of memory and aging. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 315-372). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Nash, M. (1997, March). Gift of love. A group of nuns are denoting their brains to science to help fight Alzheimer's disease. *Time Magazine, 149*(12), 43-44.

Nesselroade, J.R. (1991a). Interindividual differences in intraindividual change. In L.M. Collins & J.L. Horn (Eds.), *Best methods for the analysis of change. Recent advances, unanswered questions, future directions* (pp. 92-105). Washington, D.C.: American Psychological Association.

Nesselroade, J.R. (1991b). The warp and the woof of the developmental fabric. In J.F. Wohlwill; R.M. Downs, L.S. Liben, D.S. Palermo (Eds.), *Visions of aesthetics, the environment and development: The legacy of Joachim F. Wohlwill* (213-240). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Nesselroade, J.R., & Featherman, D.L. (1991). Intraindividual variability in older adults' depression scores: some implications for developmental theory and longitudinal research. In D. Magnusson, L.R. Bergman, G. Rudinger, & B. Torestad (Eds.), *Problems and methods in longitudinal research: Stability and change* (pp. 47-66). Cambridge, England: Cambridge University Press.

Nesselroade, J.R., & Labouvie, E.W. (1985). Experimental design in research on aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 35-60). New York: Van Nostrand Reinhold Company.

Nettelbeck, T., & Rabbitt, P.M.A. (1992). Aging, cognitive performance, and mental speed. *Intelligence, 16* (2), 189-205.

Nolan, K.A., and Blass, J.P. (1992). Preventing cognitive decline. *Clinics in Geriatric Medicine, 8*(1), 19-34.

Obler, L.K., & Albert, M.L. (1985). Language skills across adulthood. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging*, (2nd ed., pp. 463-473). New York: Van Nostrand Reinhold Company.

O'Connor, D.W. (1994). Mild dementia: a clinical perspective. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.), *Dementia and normal aging* (pp. 91-117). Cambridge, England: Cambridge University Press.

Offenbacher, D.I., & Poster, C.H. (1985). Aging and the baseline code: An alternative to the "normless elderly". *Gerontologist, 25*(5), 526-521.

Ogden, J.A. (1990). Spatial abilities and deficits in aging and age-related disorders. In F. Boller & J. Grafman (Eds), *Handbook of Neuropsychology, Vol. 4. The commisuratized brain; Aging and dementia,* (pp. 265-278). Amsterdam: Elsevier Science Publications.

Olbrich, E. Potentials in old age: Personal and social process in their development. *Zeitschrift fur Gerontologie, 23*(5), 246-251.

Owsley, C., & Burton, K.B. (1991). Aging and spatial contrast sensitivity; Underlying mechanisms and implications for everyday life. In P. Bagnoli, & W. Hodos (Eds.), *The changing visual system: maturation and aging in the central nervous system. NATO ASI series. Series A, Life Sciences, 222,* 119-136.

Park, D.C. (1992). In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 449-493). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Peress, N.S., Kane, W.C., & Aronson, S.M. (1973). Central nervous system findings in a tenth decade autopsy population. In *Neurobiological aspects of maturation and aging. Progress in Brain Research. Volume 40* (pp. 473-483). Amsterdam: Elsevier Science Publications.

Perlmutter, M. (1980). An apparent paradox about memory aging. In L.W. Poon, J.L. Fozard, L.S. Cermak, D. Arenberg, & L.W. Thompson (Eds.,), *New directions in memory and aging,* (pp. 345-353). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Pirsig, R.M. (1984). *Zen and the art of motorcycle maintenance.* New York: Bantam books.

Plomin, R., and McClearn, G.E. (1990). Human behavioral genetics of aging. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 67-78). New York: Academic Press, Inc.

Poitrenaud, J., Barrere, H., Darcet, P., & Driss, F. (1983). Aging of cognitive functions. Results of a longitudinal study. *Presse Medicale, 12*(48), 3110-23. Abstract.

Poon, L.W. (1985). Differences in human memory with aging: nature, causes, and clinical implications. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 427-462). New York: Van Nostrand Reinhold Company.

Query, W.T., & Megran, J. (1983). Age-related norms for AVLT in a male patient population. *Journal of Clinical Psychology, 39*(1), 136-138.

Rabbitt, P. (1983). How can we tell whether human performance is related to chronological age? In D. Samuel, A. Algeria, S. Gershon, V.E. Grimm, & G. Toffano (Eds.), *Aging of the Brain* (pp. 9-18). New York: Raven Press.

Rabbitt, P. (1986). Memory impairment in the elderly. In P.E. Bebbington, & r. Jacoby (Eds.), *Psychiatric disorders in the elderly*, (pp. 101-119). London: Mental Health Foundation.

Rabbitt, P. (1990). Applied cognitive gerontology: some problems, methodologies and data. Special Issue: Applying cognitive psychology in the 1990's. *Applied Cognitive Psychology 4*(4), 225-246.

Rabbitt, P. (1993a). Crystal quest: a search for the basis of maintenance of practised skills into old age. In A.D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: a tribute to Donald Broadbent* (pp. 188-230). Oxford, England: Oxford University Press.

Rabbitt, P. (1993b). Does it all go together when it goes? The nineteenth Bartlett Memorial lecture. *The Quarterly Journal of Experimental Psychology, 46A*(3), 385-434.

Rabbitt, P. (1993c). Baseline changes in cognitive performance with age. In R. Levy, R. Howard & A. Burns (Eds.), *Treatment and care in old age psychiatry* (pp. 11-30). Petersfield, UK and Bristol, PA, USA: Wrightson Biomedical Publishing Ltd.

Rabbitt, P., Donlan, C., Bent, N., McInnes, L., & Abson, V. (1993). The University of Manchester Age and Cognitive performance Research Centre and North East Age Research Longitudinal Programmes, 1982 to 1997. *Zeitschrift fur Gerontologie, 26*, 176-183.

Rabbitt, P., & Goward, L. Age, information processing speed, and intelligence. *The Quarterly Journal of Experimental Psychology.* *47*(3), 741-60.

Rabbitt, P., & McInnis, L. (1988). Do clever old people have earlier and richer first memories? *Psychology and Aging, 3*(4), 338-41.

Rabbitt, P., Watson, P., Donlan, C., Bent, N., & McInnis. (1994). Subject attrition in a longitudinal study of cognitive performance in community-based elderly people. In B. Vellas, J.L. Albarede, P.J. Garry (Eds.), *Facts and research in gerontology: Epidemiology and aging,* (29-34). New York: Springer Publishing Co.

Rediess, S., & Caine, E.D. (1996). Aging, cognition, and DSM-IV. *Aging, Neuropsychology, and Cognition.* *3*(2), 105-117.

Reese, H.W., & Rodeheaver, D. (1985). Problem solving and complex decision making. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 474-499). New York: Van Nostrand Reinhold Company.

Reitan, R.M. (1979). *Manual for administration of neuropsychological test batteries for adults and children.* Tucson, Arizona: Author.

Reitan, R.M., & Wolfson, D. (1995). Influence of age and education on neuropsychological test results. *The Clinical Neuropsychologist, 9*(2), 151-158.

Reitan, R.M., & Wolfson, D. (1996). Relationships of age and education to Wechsler Adult Intelligence Scale IQ values in brain-damaged and non-brain-damaged groups. (1996). *The Clinical Neuropsychologist, 19*(3), 293-304.

Riegel, K.F. (1977). History of psychological gerontology. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging,* (pp. 70-102). New York: Van Nostrand Reinhold Company.

Riegel, K.F., Riegel, R.M., & Meyer, G. (1967). Socio-psychological factors of aging: a cohort-sequential analysis. *Human Development, 10*(1), 27-56.

Robbins, T.W., James, M., Owen, A.M., Sahakian, B.J., McInnes, L., & Rabbitt, P. (1994). Cambridge neuropsychlogical test automated battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dementia 5*(5), 266-281.

Rogosa, D. (1988). Myths about longitudinal research. In *Methodological issues in aging research,* (pp. 171-209). New York: Springer Publishing Company.

Roman, D.D., Edwall, G.E., Buchanan, R.J., & Patton, J.H. (1991). Extended norms for the Paced Auditory Serial Addition task. *The Clinical Neuropsychologist, 5*(1), 33-40.

Rowe, J.W., & Kahn, R.L. Human aging: Usual and successful. (1987). *Science, 237,* 143-149.

Roth, M. (1994). The relationship between dementia and normal aging of the brain. In F.A. Huppert, C. Brayne, & D.W. O'Connor, (Eds.), *Dementia and normal aging* (pp. 57-76). Cambridge, England: Cambridge University Press.

Rush, M.C., Panek, P.E., & Russell, J.E. (1990). Analysis of individual variability among older adults on the Stroop color word Interference Test. *International Journal of Aging and Human Development,* *30*(3), 225-36.

Russell, E.W. (1986). The psychometric foundation of clinical neuropsychology. In S.B. Filskov & T.J. Boll (Eds.), *Handbook of clinical neuropsychology. Volume 2,* (pp. 45-80). New York: John Wiley & Sons.

Ryan, J.J., Paolo, A.M., & Brungardt, T.M. (1992). WAIS-R test-retest stability in normal persons 75 years and older. *Clinical Neuropsychologist, 6*(1), 3-8.

Sagar, H.J. (1990). Aging and age-related neurological disease: remote memory. In F. Boller & J. Grafman (Eds), *Handbook of Neuropsychology, Vol. 4. The commisuratized brain; Aging and dementia,* (pp. 311-324). Amsterdam: Elsevier Science Publications.

Salthouse, T.A. (1985a). Speed of behavior and its implications for cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp.400-426). New York: Van Nostrand Reinhold Company.

Salthouse, T.A. (1985b). A theory of cognitive aging. In G.E. Stelmach, & P.A. Vroon (Eds.), *Series: Advances in Psychology 28.* New York: Elsevier Science Publishing Co. Inc.

Salthouse, T.A. (1990). Cognitive competence and expertise in aging. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 310-319). New York: Academic Press, Inc.

Salthouse, T.A. (1991). *Theoretical perspectives on cognitive aging.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Salthouse, T.A. (1992a). Reasoning and spatial abilities. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 167-211). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Salthouse, T.A. (1992b). The information-processing perspective on cognitive aging. In R.J. Sternberg & C.A. Berg (Eds.), *Intellectual development* (pp. 261-277). New York: Cambridge University Press.

Salthouse, T.A., and Coon, V.E. (1994). Interpretation of differential deficits: The case of aging and mental arithmetic. *Journal of Experimental Psychology Learning, Memory, and Cognition, 20*(5), 1172-1182.

Satz, P. (1993). Brain reserve capacity on symptom onset after brain injury: a formulation and review of evidence for threshold theory. *Neuropsychology. 7,* (3), 273-295.

Saudino, K.J., Plomin, R., Pedersen, N.L., & McClearn, G.E. (1994). The etiology of high and low cognitive ability during the second half of the life span. *Intelligence, 19*(3), 359-371.

Schaie, K.W. (1983). The Seattle longitudinal study: A 21-year exploration of psychometric intelligence in adulthood. In K.W. Schaie (Ed.), *Longitudinal studies of psychological development* (pp. 64-135). New York: Guildford Press.

Schaie, K.W. (1988a). Ageism in psychological research. *American Psychologist, 43*, 179-183.

Schaie, K.W. (1988b). Variability in cognitive function in the elderly: implications for societal participation. *Basic Life Sciences, 43*, 191-211.

Schaie, K.W. (1990). Intellectual development in adulthood. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 291-309). New York: Academic Press, Inc.

Schaie, K.W. (1994). The course of adult intellectual development. 101st Annual convention of the American Psychological Association Distinguished Scientific contribution Award Address (1993), Toronto, Canada. *American Psychologist, 49*(4), 304-313.

Schaie, K.W., & Hertzog, C. (1985). Measurement in the psychology of adulthood and aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 61-92). New York: Van Nostrand Reinhold Company.

Schaie, K.W., & Willis, S.L. (1986). Can decline in adult intellectual functioning be reversed? *Developmental Psychology, 22*(2), 223-232.

Schroots, J.T., & Birren, J.E. (1990). Concepts of time and aging in science. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 45-64). New York: Academic Press, Inc.

Schwartzman, A.E., Gold, D., Andres, D., Arbuckle, T.Y., & Chaikelson, J. (1987). Stability of intelligence: a 40-year follow-up. Special Issue: aging and cognition. *Canadian Journal of Psychology, 41*(2), 244-256.

Seitelberger, F. (1980). Dementia associated with aging processes: functional neuropathologic study. In L. Amaducci, A.N. Davison, & P. Antuono (Eds.), *Aging of the brain and dementia,* (239-244). New York: Raven Press.

Sheppard, H.L. (1978). *Research and development strategy on employment-related problems of older workers.* Thomas Jefferson Street, NW, Washington, DC 20007: American Institutes for Research.

Shock, N. W. (1977). Biological theories of aging. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging,* (pp. 103-115). New York: Van Nostrand Reinhold Company.

Shuttleworth, S.G. (1994). *Age shall not weary you.* Port Elizabeth, South Africa: Cedar.

Shuttleworth-Jordan, A.B. (1982, October). *Complementarity as a philosophy of psychological science - implications for the practice of psychotherapy.* Paper presented at the annual conference of the Psychological Association of South Africa (PASA), Bloemfontein, South Africa.

Shuttleworth-Jordan, A.B. (1984). *Complementarity and the practice of psychotherapy: An argument against eclecticism.* M.Sc (Clin Psych) Dissertation, University of Cape Town.

Shuttleworth-Jordan, A.B. (1992, June). *Age and second language effects on Digit Span backwards relative to Digit Span forwards, and Digit Supraspan.* Paper presented at the 14th International Neuropsychological Society European Meeting, Durham, England.

Shuttleworth-Jordan, A.B. (1995). Against mutton masquerading as lamb: Additional commentary on the interpretation and misinterpretation of the SAWAIS. *South African Journal of Psychology, 25*(4), 244-246.

Shuttleworth-Jordan, A.B. (1996). On not reinventing the wheel: A clinical perspective on cultural aspects of test usage in South Africa. *South African Journal of Psychology. 26(2), 96-102.*

Shuttleworth-Jordan, A.B. (in press). Age and education effects on brain-damaged subjects: 'Negative findings revisited'. *The Clinical Neuropsychologist.*

Shuttleworth-Jordan, A.B., & Bode, S.G. (1995a). Use of the SAWAIS Digit Symbol test of Incidental Recall. *South African Journal of Psychology, 25*(1), 57-60.

Shuttleworth-Jordan, A.B. & Bode, S.G. (1995b). Taking account of the age effect on Digit Symbol and Incidental Recall for diagnostic purposes. *Journal of Clinical and Experimental Neuropsychology, 17*(3), 439-448.

Shuttleworth-Jordan, A.B. & Bode, S.G. (1996). *Normative data for Finger Tapping across the 20 to 89 year old adult age range.* Unpublished data, Rhodes University, Grahamstown, South Africa.

Shuttleworth-Jordan, A.B. & Cornfield, S.P. (1996). *Trail Making Test II: Review with particular reference to performance in older adult populations.* Manuscript in preparation.

Shuttleworth-Jordan, A.B., & Saayman, G. (1989). Differential effects of alternative strategies on psychotherapeutic process in group dream work. *Psychotherapy, 26*(4), 514-519.

Smith, G.E., Petersen, R.C., Parisi, J.E., Ivnik, R. J., Kokmen, E., Tangalos, E.G., & Waring, S. (1996). Definition, course, and outcome of Mild Cognitive Impairment. *Aging, Neuropsychology, and Cognition, 3*(2), 148-153.

South African Wechsler Adult Intelligence Scale manual. (1969). Johannesburg: National Institute for Personnel Research, Human Sciences Research Council (HSRC).

Spirduso, W.W. & MacRae, P.G. (1990). Motor performance and aging. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 183-200). New York: Academic Press, Inc.

Spreen, O. & Strauss, E. (1991). *A Compendium of Neuropsychological Tests. Administration, norms and commentary.* New York: Oxford University Press.

Stanton, B.A., Jenkins, C.D., Savageau, J.A., & Zyzanski, S.J. (1984). Age and educational differences on the Trail Making Test and Weschler Memory Scales. *Perceptual and Motor Skills, 58,* 311-318.

Sterns, H.L., & Miklos, S.M. (1995). The aging worker in a changing environment: Organizational and individual issues. Special Issue: Careers from midlife. *Journal of Vocational Behavior 47*(3), 248-268.

Stones, M.J. (1978). Aging and semantic memory: structural age differences. *Experimental Aging Research, 4*(2), 125-132.

Stuss , D.T., Stethem, L.L., and Poirier, C.A. (1987). Comparison of three tests of attention and rapid information processing across six age groups. *The Clinical Neuropsychologist, 1*(2), 139-152.

Swan, G.E., La Rue, A., Carmelli, D., Reed, T. E., & Fabsitz, R.R. (1992). Decline in cognitive performance in aging twins: Heritability and biobehavioral predictors from the National Heart, Lung, and Blood Institute twin study. *Archives of Neurology. 494*(5), 476-481.

Sykes, J.B. (Ed.). (1982). *The concise Oxford dictionary.* (7th Ed.). Oxford: Clarendon Press.

Thomae, H. (1990). Stress, satisfaction, competence: Findings from the Bonn Longitudinal study on aging. In M. Bergener & I.F. Sanford (Eds.), *Clinical and scientific psychogeriatrics, Vol. 1: The holistic approaches* (pp. 117-134). New York: Springer Publishing Co, Inc.

Thomae, H., Lehr, U., & Schmitz-Scherzer, R. (1981). Research on aging in behavioral sciences. *Zeitschrit fur Gerontologie, 14*(3), 204-223.

Tomlinson, B.E., Blessed, G., & Roth, M. (1968). Observations on the brains of non-demented old people. *Journal of the neurological sciences, 7*(2), 331-356.

Troyer, A.K., Graves, R.E., & Cullum, C.M. (1994). Executive functioning as a mediator of the relationship between age and episodic memory in healthy aging. *Aging and Cognition, 1*(1), 045-053.

Tuokko, H., & Woodward, T.S. (1996). Development and validation of a demographic correction system for neuropsychological measures used in the Canadian Study of Health and Aging. *Journal of Clinical and Experimental Neuropsychology, 18*(4), 479-616.

Uchiyama, C.L., Mitrushina, M., Satz, P., & Schall, M. (1996). Direct and indirect effects of demographic, medical, and psychological variables on neuropsychological performance in normal geriatric persons: a structural equation mode. *Journal of the International Neuropsychological Society, 2*(4), 299-305.

Van Gorp, W.G., Satz, P., Kiersch, M.E., & Henry, R. (1986). Normative data on the Boston naming test for a group of normal older adults. *Journal of Clinical and Experimental Neuropsychology, 8*(6), 702-705.

Van Gorp, W.G., Satz, P., & Mitrushina, M. (1990). Neuropsychological processes associated with normal aging. *Developmental neuropsychology.* *6*(4), 279-290.

Vazzana, a.D., Bylsma, F.W., Kawas, C. (1995, February). *Performance of normal elderly persons on Tey-Osterrieth Complex Figure Test (CFT).* Paper presented at the 23rd annual meeting of the International neuropsychological Society, Seattle, Washington, U.S.A.

Verran, J.A., Mark, B.A., and Lamb, G. (1992). Focus on psychometrics: Psychometric examination of instruments using aggregated data. *Research in Nursing and Health, 15*, 237-240.

Walsh, K. (1991). *Understanding Brain Damage. A Primer of Neuropsychological Evaluation* (2nd ed.). Melbourne: Churchill Livingstone.

Weintraub, S., Powell, D.H., & Whitla, D.K. (1994). Successful cognitive aging: Individual differences among physicians on a computerized test of mental state. *Journal of Geriatric Psychiatry, 27*(1), 15-34.

Welford, A.T. (1977). Motor performance. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging,* (pp. 450-496). New York: Van Nostrand Reinhold Company.

West, R. L. (1989). Planning practical memory training for the aged. In L.W. Poon, D.C. Rubin, & B.A. Wilson (Eds), *Everyday cognition in adulthood and late life* (pp. 573-597). New York: Cambridge University Press.

Willis, S. L. (1985). Towards an educational psychology of the older adult learner: intellectual and cognitive bases. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 818-847). New York: Van Nostrand Reinhold Company.

Winocur, G., Moscovitch, M., & Freedman, J. (1987). An investigation of cognitive function in relation to psychosocial variables in institutionalized old people. special Issue: Aging and cognition. *Canadian Journal of Psychology, 41*(2), 257-269.

Wohlwill, J.F. (1973). Individual differences in development. In J.F. Wohlwill. *The study of behavioral development*, (pp. 333-375). New York: Academic Press.

Woodruff-Pak, D.S. (1990). Mammalian models of learning, memory, and aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 234-257). New York: Van Nostrand Reinhold Company.

Woods, A.M., & Rusin, M.J. (1988). Stress management and the elderly. In M.L. Russell (Ed.), *Stress management for chronic disease. Pergamon general psychology series, Vo. 152*, (pp. 49-62). Oxford, England: Pergamon Press.

Yesavage, J.A. (1989). Techniques for cognitive training of memory in Age-Associated Memory Impairment. Symposium: Memory and aging (1988), Lausanne, Switzerland). *Archives of Gerontology & Geriatrics, Suppl, 1*, 185-190.

Zarit, S.H., Eiler, J., & Hassinger, M. (1985). Clinical assessment. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 725-754). New York: Van Nostrand Reinhold Company.

Zung, W.W., & Zung, E.M. (1986). Use of the Zung Self-rating Depression Scale in the elderly. *Clinical Gerontologist, 5*(1-2), 137-148.

APPENDIX

TABLE A. PHASE 1 GROUP DATA

NUMBERS OF SUBJECTS (N), MEANS (X), AND STANDARD DEVIATIONS (SD)

General Note: This table contains all the data sets that were included in the Phase 1 analysis. Each data set is reported as it appears in the relevant main text. Where data have been combined from a number of sources to span the adult age range (in order to make up a more complete data set), *these are not the present author's compilations*. In all such instances they reflect pre-existing combinations of normative data in the seminal sources of Spreen & Strauss (1991), or Lezak (1983), which are used regularly as valid indicators of normative trends in current neuropsychological practice. Thus, whereas Table A reflects the author's overall combination of existing data sets used in the present analysis, the compilations of data sets in themselves are not of the author's making. The reader is referred to the method (Sections 6.2.1; 6.2.2.1; 6.2.2.3; 9.5.1.2, p.419, middle paragraph) for a clear definition of what constitutes a data set, how they were accumulated for the present study, and a critical discussion of the collection with respect to sample effects. (Further specific annotations are listed together at the end of the table, p.17).

ATTENTION

Paced Auditory Serial Addition Test (PASAT)

Stuss, Stethem, & Pelchat, 1988, in Spreen & Strauss, 1991[1]

PASAT, presentation rate 2.4 sec

Age	16-29	30-49	50-69
N	30	30	30
X	47.4	43.4	43.5
SD	10.1	10.2	13.6

PASAT, presentation rate 1.2 sec

Age	16-29	30-49	50-69
N	30	30	30
X	27.4	24.6	21.2
SD	9.9	10.6	14.4

Stuss, Stethem, & Poirier, 1987

PASAT, presentation rate 2.4 sec

Age	20-29	30-39	40-49	50-59	60-69
N	10	10	10	10	10
X	51.3	44.8	49.1	35.3	47.7
SD	6.2	8.1	7.3	15.6	8.8

PASAT, presentation rate 1.2 sec

Age	20-29	30-39	40-49	50-59	60-69
N	10	10	10	10	10
X	33.6	27.6	23.0	8.2	30.7
SD	9.2	8.1	15.0	11.5	8.7

Roman, Edwall, Buchanan & Patton, 1991

PASAT, presentation rate 2.4 sec

Age	18-27	33-50	60-75
N	62	40	41
X	45.0	44.0	37.0
SD	4.3	4.9	9.1

PASAT, presentation rate 1.2 sec

Age	18-27	33-50	60-75
N	62	40	41
X	28.0	28.0	20.0
SD	6.7	9.3	6.1

Digit Span

Shuttleworth-Jordan, 1992

Digit Span Forward

Age	18-25	40-49	60-69	70-79	80-89
N	33	15	15	33	25
X	7.48	7.53	7.40	7.18	6.92
SD	1.06	1.19	1.45	1.10	0.86

Digit Span Backward

Age	18-25	40-49	60-69	70-79	80-89
N	33	15	15	33	25
X	6.24	5.60	5.73	5.52	5.08
SD	1.17	1.30	1.83	1.00	1.07

VERBAL MEMORY

Wechsler Memory Scale (WMS) Associate Learning

Strauss & Spreen, 1989, in Spreen & Strauss, 1991

WMS Associate Learning

Age	20-29	30-39	40-49	50-59	60-69	70+
N	23	13	14	12	14	14
X	18.33	18.21	18.29	17.30	14.30	15.89
SD	1.45	2.28	2.82	3.34	2.52	3.10

Data collated from Wechsler, 1945 (20-29 years), Hulicka, 1966 (30-39, 60-69, 70-79), and Klonoff & Kennedy (80-92), 1966, in Lezak, 1983[2]

WMS Associate Learning

Age	20-29	30-39	40-49	60-69	70-79	80-92
N	50	53	46	70	46	115
X	15.72	15.48	13.91	11.94	10.98	10.15
SD	2.81	3.48	3.12	4.53	4.78	3.80

Wechsler Memory Scale (WMS) Logical Memory

Abikoff, Alvir, Hong, Sukoff, Orazio, Soloman, et al. 1987, in Spreen & Strauss, 1991

WMS Logical Memory Immediate

Age	18-29	30-39	40-49	50-59	60-69	70+
N	74	67	41	54	56	46
X	22.99	24.57	23.44	23.63	20.48	19.11
SD	6.66	6.97	5.01	6.14	6.42	6.74

WMS Logical Memory Delayed

Age	18-29	30-39	40-49	50-59	60-69	70+
N	74	67	41	54	56	46
X	19.84	22.16	21.07	20.13	17.34	15.33
SD	6.67	7.57	5.91	6.48	6.71	7.57

Van Gorp, Satz., & Mitrushina, 1990

WMS Logical Memory Immediate

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	9.75	8.47	9.27	8.05
SD	2.40	2.65	3.15	2.19

WMS Logical Memory Delayed

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	7.80	5.73	6.31	5.32
SD	2.16	2.24	3.16	2.48

Data collated from Wechsler, 1945 (20-29 years), Hulicka, 1966 (30-39, 60-69, 70-79), and Klonoff & Kennedy (80-92), 1966, in Lezak, 1983

WMS Logical Memory

Age	20-29	30-39	40-49	60-69	70-79	80-92
N	50	53	46	70	46	115
X	9.28	7.99	8.09	7.34	7.35	5.72
SD	3.10	2.95	2.52	2.90	3.83	2.91

Rey Auditory-Verbal Learning Test (RAVLT)

Geffen, Hoar, & O'Hanlon 1990, in Spreen & Strauss, 1991

RAVLT-1 Males

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	8.4	6.0	6.4	6.5	4.9	3.6
SD	1.2	1.8	1.8	2.0	1.1	0.8

RAVLT-5 Males

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	12.2	11.4	10.9	11.8	8.9	8.2
SD	2.2	2.6	2.0	2.6	2.0	2.5

RAVLT-Total Males

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	54.9	46.0	47.5	47.6	36.7	32.6
SD	7.0	10.9	8.3	8.5	8.4	8.3

RAVLT-Delayed Males

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	10.6	10.4	10.5	10.0	7.1	5.6
SD	2.4	2.3	2.7	2.6	3.8	2.6

RAVLT-Recognition Males

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	14.2	13.5	14.2	13.9	12.4	11.5
SD	0.8	1.5	1.0	0.9	2.8	2.6

RAVLT-Trial 1 Females

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	7.7	8.0	6.8	6.4	6.0	5.6
SD	1.0	2.0	1.5	1.5	2.2	1.4

RAVLT-Trial 5 Females

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	12.9	12.7	12.8	11.6	11.9	10.1
SD	1.5	1.3	1.4	2.1	1.6	1.2

RAVLT-Total Females

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	55.3	55.9	52.1	47.6	49.0	41.6
SD	6.6	6.3	7.1	7.7	7.1	6.6

RAVLT-Delayed Females

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	11.0	12.2	11.1	10.2	10.3	8.3
SD	2.0	2.5	2.3	2.7	2.3	2.1

RAVLT-Recognition Females

Age	20-29	30-39	40-49	50-59	60-69	70+
N	10	10	11	11	10	10
X	14.4	14.2	14.4	13.7	13.8	13.6
SD	0.8	1.7	0.8	1.1	1.1	2.0

Mitrushina, Satz, Chervinsky & D'Elia, 1991

RAVLT-Trial 1

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	6.4	5.9	5.1	5.1
SD	1.5	1.6	1.8	1.6

RAVLT-Trial 5

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	12.1	11.5	10.3	9.7
SD	2.4	3.0	2.9	2.8

RAVLT-Recognition

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	13.2	13.0	12.7	12.6
SD	1.3	1.1	1.9	1.9

Buschke Selective Reminding Test (BSRT)

Larrabee, Trahan, Curtiss, & Levin, 1988, in Spreen & Strauss, 1991[3]

BSRT-Total

Age	20's	30's	40's	50's	60's	70's	80's
N	51	29	31	24	50	59	27
X	128	125	125	122	115	105	98
SD	9.2	13.4	12.0	10.5	15.8	16.7	17.5

BSRT-Delayed

Age	20's	30's	40's	50's	60's	70's	80's
N	51	29	31	24	50	59	27
X	11.5	10.7	11.0	10.8	9.6	9.1	8.4
SD	.8	2.0	1.43	1.4	2.5	2.6	2.5

BSRT-RLTR (Words recalled twice in a row)

Age	20's	30's	40's	50's	60's	70's	80's
N	51	29	31	24	50	59	27
X	8.1	10.1	11.2	10.8	14.7	20.7	22.2
SD	9.4	9.7	11.3	9.3	11.8	14.4	10.7

BSRT-CLTR (Continuous recall)

Age	20's	30's	40's	50's	60's	70's	80's
N	51	29	31	24	50	59	27
X	115	108	107	102	89	70	55
SD	19.7	27.6	26.6	22.4	35	36	29

VISUAL MEMORY

Wechsler Memory Scale (WMS) Visual Reproducation:

Data collated from Trahan, Quintana, Willingham, & Goethe, 1988 (18-69 years), and Haaland et al., 1983 (65+ years), in Spreen & Strauss, 1991

WMS Visual Reproduction Immediate

Age	18-29	30-49	50-64	65-69	70-74	75-79	80+
N	97	81	51	49	74	40	13
X	10.48	10.10	8.73	6.0	5.1	4.9	3.3
SD	1.93	2.55	2.59	2.1	2.0	2.0	2.3

WMS Visual Reproduction Delayed

Age	18-29	30-49	50-64	65-69	70-74	75-79	80+
N	97	81	51	49	74	40	13
X	9.84	9.26	7.35	5.4	4.3	4.2	2.8
SD	2.21	2.74	2.46	2.5	2.3	1.9	1.9

Van Gorp, Satz., & Mitrushina, 1990

WMS Visual Reproduction Immediate

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	9.60	8.67	7.79	4.09
SD	2.76	3.81	2.99	2.26

WMS Visual Reproduction Delayed

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	6.60	6.33	4.38	1.64
SD	2.46	3.83	3.93	1.69

Data collated from Wechsler, 1945 (20-29 years), Hulicka, 1966 (30-39, 60-69, 70-79), and Klonoff & Kennedy (80-92), 1966, in Lezak, 1983.

WMS Visual Reproduction

Age	20-29	30-39	40-49	60-69	70-79	80-92
N	50	53	46	70	46	115
X	11.0	10.09	8.35	6.03	4.95	3.76
SD	2.73	3.01	3.17	3.72	3.42	2.70

Rey-Osterrieth Complex Figure Test (RCF) Recall

Data collated from Kolb & Whishaw, 1985 (16-49 years), and Strauss & Spreen, unpublished study (50+ years), in Spreen & Strauss, 1991

RCF Recall Delayed

Age	16-30	31-44	50-59	60-69	70+
N	67	26	14	13	10
X	22.7	19.5	18.82	16.65	11.80
SD	7.00	6.70	7.37	8.70	6.20

Van Gorp, Satz., & Mitrushina, 1990

Rey Complex Figure Recall Delayed

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	14.45	14.13	11.13	8.41
SD	5.33	7.83	6.66	5.86

Troyer, Graves, & Cullum, 1994

Rey Complex Figure Recall Delayed

Age	60-69	70-79	80-91
N	19	17	15
X	18.9	15.7	13.9
SD	5.9	6.0	4.1

Benton Visual Retention Test - Revised (BVRT-R)

Robertson-Tehabo & Arenberg 1989, in Spreen & Strauss, 1991

BVRT-R: Male/No Degree

Age	20's	30's	40's	50's	60's	70's	80's
N	45	57	51	42	28	47	15
X	2.02	3.56	3.75	5.07	6.25	7.72	11.13
SD	1.53	3.04	2.46	2.79	2.41	3.92	4.29

BVRT-R: Male/College Degree

Age	20's	30's	40's	50's	60's	70's	80's
N	92	181	155	156	123	139	40
X	2.47	2.49	2.94	3.67	4.66	6.03	8.15
SD	1.87	1.87	1.89	2.57	2.81	3.20	4.25

BVRT-R: Female/No Degree

Age	20's	30's	40's	50's	60's	70's	80's
N	19	27	23	35	42	32	9
X	2.90	3.04	3.87	5.11	6.02	7.53	9.44
SD	2.08	2.21	1.82	1.95	2.58	3.37	3.84

BVRT-R: Female/College Degree

Age	20's	30's	40's	50's	60's	70's	80's
N	37	70	28	38	56	43	13
X	2.70	2.67	2.61	3.68	4.57	6.79	7.62
SD	1.82	2.11	2.17	1.99	2.17	2.97	4.43

Rey Visual Design Learning Test (RVDLT)

Strauss & Spreen, unpublished study, in Spreen & Strauss, 1991

RVDLT-Trial 1

Age	20's	30's	40's	50's	60's	70+
N	23	14	14	10	10	13
X	6.09	5.36	5.64	5.00	4.70	4.29
SD	2.11	2.10	1.91	1.56	1.16	1.64

RVDLT-Trial 5

Age	20's	30's	40's	50's	60's	70+
N	23	14	14	10	10	13
X	13.48	11.64	10.93	9.30	8.20	7.46
SD	1.70	2.21	2.92	2.16	2.62	1.66

RVDLT-Total

Age	20's	30's	40's	50's	60's	70+
N	23	14	14	10	10	13
X	53.65	46.64	44.36	37.00	35.20	31.38
SD	8.02	9.30	12.36	7.77	7.32	6.67

RVDLT-Recognition

Age	20's	30's	40's	50's	60's	70+
N	23	14	14	10	10	13
X	14.61	14.07	14.21	13.30	13.10	13.46
SD	0.78	1.27	0.89	1.70	0.99	2.02

LANGUAGE SKILLS

Controlled Word Association (FAS)

Read,1987 in Spreen & Strauss, 1991

FAS Oral/ Education < 12 years

Age	50-54	55-59	60-64	65-69	70-74	75+
N	42	67	70	56	55	55
X	41.52	37.42	37.57	39.25	36.47	35.20
SD	12.33	9.55	10.90	11.50	13.21	11.90

FAS Oral/ Education > 13 years

Age	50-54	55-59	60-64	65-69	70-74	75+
N	56	78	87	90	60	59
X	41.16	45.96	40.76	44.16	41.00	39.08
SD	11.42	11.22	9.58	10.77	10.44	14.17

Axelrod & Henry, 1992

FAS Oral

Age	50-59	60-69	70-79	80-89
N	20	20	20	20
X	41.1	39.6	36.0	37.8
SD	9.9	10.7	9.3	14.0

Animal Name Fluency

Read, 1987 in Spreen & Strauss, 1991

Animal Name Fluency - Education < 12 years

Age	50-54	55-59	60-64	65-69	70-74	75+
N	42	67	70	56	55	55
X	18.64	18.81	18.61	17.75	16.35	15.09
SD	4.06	3.83	4.49	4.02	4.36	4.25

Animal Name Fluency - Education > 13 years

Age	50-54	55-59	60-64	65-69	70-74	75+
N	56	78	87	90	60	59
X	19.98	20.88	19.97	19.13	18.28	17.07
SD	4.08	5.14	5.00	5.01	3.58	4.93

Boston Word Naming Test

Van Gorp, Satz, Kiersch, & Henry, 1986, in Spreen & Strauss, 1991

Boston Word Naming

Age	60-64	65-69	70-74	75-79	80-85
N	15	37	47	23	14
X	56.7	55.6	54.5	51.7	51.5
SD	3.0	2.4	5.1	6.2	7.0

Au, Joung, Nicholas, Obler, Kass, & Albert, 1995

Boston Word Naming

Age	30's	50's	60's	70's
N	33	32	41	39
X	92.7	89.6	89.2	83.1
SD	4.2	6.5	8.4	13.2

VISUAL SKILLS

Right Left Orientation

Spreen & Strauss, unpublished study, in Spreen & Strauss, 1991

R/L Orientation

Age	20-30	50-59	60-69	70-79	80+
N	20	19	27	23	15
X	17.82	17.16	16.78	15.17	15.00
SD	2.40	2.89	2.98	2.66	2.70

Embedded Figures Test

Spreen & Strauss, unpublished study, in Spreen & Strauss, 1991

Embedded Figures Test, Total Correct

Age	20's	30's	40's	50's	60's	70+
N	23	14	14	10	10	14
X	14.5	15.7	14.1	15.4	14.5	15.5
SD	1.0	1.0	1.0	0.9	0.8	0.5

Embedded Figures Test, Total Credit

Age	20's	30's	40's	50's	60's	70+
N	23	14	14	10	10	14
X	31.78	31.64	31.14	31.60	30.30	30.36
SD	0.85	0.74	1.46	0.80	2.50	2.41

Rey Osterrieth Complex Figure Test (RCF) Copy

Data collated from Kolb & Whishaw, 1985 (16-49 years), and Strauss & Spreen, unpublished study (50+ years), in Spreen & Strauss, 1991

RCF Copy

Age	16-30	31-44	50-59	60-69	70+
N	67	26	14	13	10
X	35.10	33.20	35.57	33.15	32.90
SD	1.5	6.1	.76	4.02	2.69

Van Gorp, Satz., & Mitrushina, 1990

RCF Copy

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	32.50	32.93	31.73	30.14
SD	4.67	3.38	3.40	5.75

Trail Making Test (TMT)

Van Gorp, Satz., & Mitrushina, 1990

TMT A

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	41.50	43.20	50.08	59.73
SD	7.38	14.98	12.88	15.95

TMT B

Age	57-65	66-70	71-75	76-85
N	28	45	57	26
X	84.40	105.20	97.79	153.09
SD	24.60	43.43	30.40	62.60

Stuss, Stethem, & Poirier, 1987

TMT A

Age	20-29	30-39	40-49	50-59	60-69
N	10	10	10	10	10
X	18.5	21.9	29.2	38.5	37.3
SD	5.1	6.3	9.0	18.2	14.7

TMT B

Age	20-29	30-39	40-49	50-59	60-69
N	10	10	10	10	10
X	41.6	46.3	64.1	83.1	73.3
SD	11.4	13.7	16.3	44.3	20.3

Symbol Digit Modalities

Centofanti, unreferenced study, in Lezak, 1983

Symbol Digit Modalities, Oral

Age	18-24	25-34	35-44	45-54	55-64	65-74
N	69	72	76	75	67	61
X	62.7	61.2	59.7	54.5	48.4	46.2
SD	9.1	7.8	9.7	9.1	9.1	12.8

Symbol Digit Modalities, Written

Age	18-24	25-34	35-44	45-54	55-64	65-74
N	69	72	76	75	67	61
X	55.2	53.6	51.1	46.8	41.5	37.4
SD	7.5	6.6	8.1	8.4	8.6	11.4

PROBLEM SOLVING

Wisconsin Card Sorting Test

Data collated from Heaton, 1981 (< 60 years), and Strauss & Spreen, unpublished study (60-94 years), in Spreen & Strauss, 1991

Wisconsin Card Sorting, Categories achieved

Age	< 40	40's	50's	60's	70's	80+
N	100	19	16	28	19	13
X	5.6	4.8	5.6	5.5	5.0	4.23
SD	1.0	1.8	1.1	1.1	1.3	1.5

Wisconsin Card Sorting, Perseverative errors

Age	<40	40's	50's	60's	70's	80's
N	100	19	16	28	19	13
X	10.4	16.0	11.3	12.25	15.9	25.77
SD	8.0	13.9	6.9	10.91	9.8	12.23

Beatty, 1993

Wisconsin Card Sorting, Categories achieved

Age	18-34	35-59	60+
N	21	21	21
X	5.8	5.8	4.6
SD	0.7	0.8	1.9

Wisconsin Card Sorting, Perseverative errors

Age	18-34	35-59	60+
N	21	21	21
X	9.9	9.9	20.5
SD	5.8	8.5	21.5

Axelrod & Henry, 1992

Wisconsin Card Sorting, Categories achieved

Age	50-59	60-69	70-79	80-89
N	20	20	20	20
X	5.4	3.8	4.2	3.6
SD	1.3	2.1	1.3	2.4

Wisconsin Card Sorting, Perseverative errors

Age	50-59	60-69	70-79	80-89
N	20	20	20	20
X	14.0	22.4	25.5	33.0
SD	10.0	14.7	9.5	22.2

Stroop Test

Data collated from Regard, 1981 (20-35 years), and Spreen & Strauss, unpublished study (50-94 years), in Spreen & Strauss, 1991

Stroop Test, Dots

Age	20-35	50-59	60-69	70-79	80+
N	40	19	28	24	15
X	10.10	13.74	12.71	15.00	18.87
SD	2.01	2.58	1.90	5.07	4.67

Stroop Test, Words

Age	20-35	50-59	60-69	70-79	80+
N	40	19	28	24	15
X	12.00	16.58	16.32	19.04	24.13
SD	2.49	3.34	3.33	5.10	5.13

Stroop Test, Colour

Age	20-35	50-59	60-69	70-79	80+
N	40	19	28	24	15
X	19.25	28.90	31.82	38.33	61.13
SD	5.18	7.62	9.86	13.29	30.94

Troyer, Graves, & Cullum, 1994

Stroop Test Interference

Age	60-69	70-79	80-91
N	19	17	15
X	2.2	2.6	3.2
SD	0.4	0.7	1.1

Similarities

Axelrod & Henry, 1992

Similarities

Age	50-59	60-69	70-79	80-89
N	20	20	20	20
X	20.3	19.5	14.9	14.2
SD	4.0	4.5	5.4	8.6

MOTOR ACTIVITY

Finger Tapping Test

Bornstein, 1985, in Spreen & Strauss, 1991

Finger Tapping Preferred Males < 12 years ed

Age	20-39	40-59	60-69
N	21	13	16
X	49.7	42.3	39.1
SD	6.0	5.2	5.7

Finger Tapping Preferred Males > 12 years ed

Age	20-39	40-59	60-69
N	86	17	23
X	48.5	43.4	43
SD	6.5	7.9	4.7

Finger Tapping Nonpreferred Males < 12 years ed

Age	20-39	40-59	60-69
N	21	13	16
X	47.0	39.8	35.2
SD	5.5	3.6	5.2

Finger Tapping Nonpreferred Males > 12 years ed

Age	20-39	40-59	60-69
N	86	17	23
X	44.8	39.5	39.3
SD	6.4	5.8	6.2

Finger Tapping Preferred Females < 12 years ed

Age	20-39	40-59	60-69
N	13	22	22
X	45.2	36.3	29.7
SD	6.0	7.8	6.2

Finger Tapping Preferred Females > 12 years ed

Age	20-39	40-59	60-69
N	49	43	34
X	44.3	40.5	32.2
SD	5.8	7.1	6.0

Finger Tapping Nonpreferred Females < 12 years ed

Age	20-39	40-59	60-69
N	13	22	22
X	40.7	35.2	29.8
SD	5.0	5.8	5.6

Finger Tapping Nonpreferred Females > 12 years ed

Age	20-39	40-59	60-69
N	49	43	34
X	40.6	37.8	32.0
SD	5.6	6.0	4.9

---

[1] In all instances data from the studies reported in Spreen & Strauss (1991) are specified in the table as they appear in that text, and not from the original sources.

[2] In all instances data from the studies reported in Lezak (1983) are specified in the table as they appear in that text, and not from the original sources.

[3] For the purposes of tabulation with high numbers and a large number of age groups, Larrabee et al.'s (1988) BSRT data were rounded to the nearest integer for the means, and to the first decimal place for the standard deviations. The statistical analyses were conducted on the data in this form.

TABLE B. PHASE 1 DEMOGRAPHIC DATA
AGE RANGES, EDUCATIONAL/IQ LEVELS

(Annotations are listed together at the end of the table on p.21)

TEST	AGE RANGES	EDUC(YRS)/IQ[1]	STUDY
ATTENTION			
PASAT	16-29, 30-49, 50-69	Nil reported[2]	Stuss et al., 1988, in Spreen & Strauss, 1991
PASAT	20's, 30's, 40's 50's, 60's	Ed 16, 17, 16 12, 14	Stuss et al., 1987
PASAT	18-27, 33-50, 60-75	IQ 110,110,107 Ed 12, 15, 15	Roman et al., 1991
Digit Span Forward	18-25, 40's, 60's 70's, 80's	Ed 14, 14, 14, 14, 14	Shuttleworth-Jordan, 1992
Digit Span Backward	18-25, 40's, 60's 70's, 80's	Ed 14, 14, 14, 14, 14	Shuttleworth-Jordan, 1992
VERBAL MEMORY			
WMS Assoc Learn	20's, 30's, 40's, 50's, 60's, 70-84	Nil reported	Strauss & Spreen, in Spreen & Strauss, 1991
WMS Assoc Learn	20's, 30's, 40's, 60's, 70's, 80-92	Generally representative[3]	Collated data in Lezak, 1983, (3 studies compiled)
WMS Log Mem	18-29, 30's, 40's, 50's, 60's, 70-80+	Ed estimate relatively high[4]	Abikoff et al, 1987, in Spreen & Strauss, 1991
WMS Log Mem	57-65, 66-70, 71-75, 76-85	IQ 117[5] Ed 14	Van Gorp et al., 1990
WMS Log Mem	20's, 30's, 40's, 60's, 70's, 80-92	Generally representative[3]	Collated data in Lezak, 1983, (3 studies compiled)
RAVLT	20's, 30's, 40's, 50's, 60's, 70+	IQ Above Ave	Geffen et al., 1990, in Spreen & Strauss, 1991
RAVLT	57-65, 66-70, 71-75, 76-85	IQ 117[5] Ed 14	Mitrushina et al., 1991
BSRT	18-29, 30's, 40's, 50's, 60's, 70-91	Nil reported	Larrabee et al, 1988, in Spreen & Strauss, 1991

VISUAL MEMORY

WMS Vis Reprod	18-29, 30-49, 50-64; 65-69, 75-79, 80⁺	Nil reported; Nil reported	Collated data in Spreen & Strauss, 1991 (2 studies compiled)
WMS Vis Reprod	57-65, 66-70, 71-75, 76-85	IQ 117[5] Ed 14	Van Gorp et al., 1990
WMS Vis Reprod	20's, 30's, 40's, 60's, 70's, 80-92	Generally representative[3]	Collated data in Lezak, 1983, (3 studies compiled)
RCF Recall	16-30, 31-44; 50's, 60's, 70⁺	Nil reported; Ed 13 years	Collated data in Spreen & Strauss, 1991 (2 studies compiled)
RCF Recall	57-65, 66-70, 71-75, 76-85	IQ 117[5] Ed 14	Van Gorp et al., 1990
RCF Recall	60's, 70's, 80's	Ed 14, 13, 12	Troyer et al., 1994
BVRT-R	20's to 80's in decades	No Degree vs College Degree; Male vs Female	Robertson-Tehabo & Arenberg, 1989, in Spreen & Strauss, 1991
RVDLT	20's, 30's, 40's, 50's, 60's, 70⁺	Nil reported	Strauss & Spreen, unpublished study, in Spreen & Strauss, 1991

LANGUAGE SKILLS

FAS Word Assoc Oral + Written	50 to 75⁺ in 5 year intervals	Ed <12 vs >13	Read, 1987, in Spreen & Strauss, 1991
FAS Word Assoc Oral	50's, 60's, 70's, 80's	Ed 15, 14, 15, 15	Axelrod & Henry, 1992
Animal Name Fluency	50 to 75⁺ in 5 year intervals	Ed <12 vs >13	Read, 1987 in Spreen & Strauss, 1991
Boston Word Naming	60 to 85 in 5 year intervals	Nil reported	Van Gorp et al., 1986, in Spreen & Stauss, 1991
Boston Word Naming	30's, 50's, 60's 70's	Ed estimate relatively high[6]	Au et al. 1995

VISUAL SKILLS

R/L Orientation	20's, 50's, 60's 70's, 80⁺	Nil reported	Spreen & Strauss, unpublished study, in Spreen & Strauss, 1991
Embedded Figures	20's, 30's, 40's, 50's, 60's, 70-84	Nil reported	Spreen & Strauss, unpublished study, in Spreen & Strauss, 1991
RCF Copy	16-30, 31-44; 50's, 60's, 70⁺	Nil reported; 13 years	Collated data in Spreen & Strauss, 1991 (2 studies compiled)
RCF Copy	57-65, 66-70, 71-75, 76-85	IQ 117[5] Ed 13	Van Gorp et al., 1990
Trail Making A + B	57-65, 66-70, 71-75, 76-85	IQ 117[5] Ed 13	Van Gorp et al., 1990
Trail Making A + B	20's, 30's, 40's 50's, 60's	Ed 16, 17, 16 12, 14	Stuss et al., 1987
Symbol Digit Modalities	18-24, 25-74 in 10 year intervals	Nil Reported	Centofanti, unreferenced study, in Lezak, 1995

PROBLEM SOLVING

Wisconsin Card Sorting	<40, 40-49,50-59; 60's, 70's, 80⁺	IQ Above Ave; Well-educated	Collated data in Spreen & Strauss, 1991 (2 studies compiled)
Wisconsin Card Sorting	50's, 60's, 70's, 80's	Ed 15, 14, 15, 15	Axelrod & Henry, 1992
Wisconsin Card Sorting	18-34, 35-59, 60⁺	Ed 15, 16, 16	Beatty, 1993
Stroop Test	20-35; 50's, 60's, 70's 80⁺	Average; Ed 13	Collated data in Spreen & Strauss, 1991 (2 studies compiled)
Stroop Test Interference	60's, 70's, 80's	Ed 14, 13, 12	Troyer et al., 1994
Similarities	50's, 60's, 70's, 80's	Ed 15, 14, 15 15	Axelrod & Henry, 1992

MOTOR ACTIVITY

Finger Tapping	20-39, 40-59, 60-69	Ed < 12 vs > 12	Bornstein, 1985, in Spreen & Strauss, 1991

---

[1] For the purposes of this tabulation all figures reported for Education and IQ have been rounded to the nearest integer. Where only *one* figure (or statement) is reported in the table for Education or IQ, it applies to the *total* sample, and indicates that *no separate specifications per age group were provided*. Where separate specifications are available for different age groups within one test (or for distinct clusters of data collated from different studies for one test), the educational/IQ data are ordered to match the order of the age groups (or clusters of age groups) to which they relate.

[2] In all instances the demographic details from the studies reported in Spreen & Strauss (1991) were indicated in the table as delineated in that text, and not from the original sources.

[3] For the Wechsler Memory Scale Associate Learning, Logical Memory and Visual Reproduction subtests, Lezak (1983) provides no information about the educational/IQ levels of the samples on which the normative data she presents are based. Lezak's compiled data for these three tests were compared with data from a nonclinical Caucasian sample restricted to at least 12 years of education, and a high mean educational level of 14 years (Shuttleworth-Jordan, 1996). Consistently across all the tests, for comparable age groups, Lezak's norms were markedly lower than the Shuttleworth-Jordan norms. This suggests that the data collated in Lezak (1983) for these three Wechsler Memory Scale subtests were based on samples of a lower educational mean than the Shuttleworth-Jordan sample. Thus it appears likely that Lezak's compiled norms relate to samples which are representative of the general population as distinct from a specifically well-educated population of above average IQ. In support of this supposition, Spreen & Strauss (1991) report on the same studies which make up Lezak's collation for these three subtests, and in each case state that the data are derived from samples of average IQ.

[4] Figures reported in Spreen & Strauss (1991) reveal that in Abikoff et al.'s (1987) study 45% of the sample had some college education; 30% were graduates of postgraduates; 24% were High School graduates; 6% were non-High School graduates. Thus, since 75% of the sample had at least some College education the educational level of this sample is estimated to be relatively high.

[5] Van Gorp et al. (1990) only report the FSIQ for the total sample. However their normative data table in which VIQ and PIQ figures are presented reveals a significant peak for VIQ and PIQ in the 71-75 year-old group (123 and 115, respectively), and relative to this a substantial dip in the 76-85 year-old group (111 and 101, respectively). The younger two age groups are equivalent with VIQ and PIQ falling around 115 and 110, respectively for both groups. This note applies to all the van Gorp (1990) data included in the present study, and probably also to that of Mitrushina et al., 1991, which report data from the same cohort.

[6] Au et al. (1995) do not report on the educational level of the larger cross-sectional sample which was used for analysis in the present study. They report educational levels (which across age groups range from 13.0 to 15.6 years), for a reduced sample which was used for the longitudinal follow-up. Thus, it can be extrapolated that the larger sample also constituted a well-educated group.

TABLE C. PHASE 2 GROUP DATA

NUMBERS OF SUBJECTS (N), MEANS (X), AND STANDARD DEVIATIONS (SD)

VISUAL SKILLS

Trail Making Test (TMT)

Cornfield & Shuttleworth-Jordan, 1996

TMT A

Age	60-69	70-79	80-89
N	33	38	34
X	28.77	35.64	47.85
SD	7.78	7.63	13.15

TMT B

Age	60-69	70-79	80-89
N	33	38	34
X	69.32	85.00	131.05
SD	20.90	38.71	43.73

SAWAIS Digit Symbol Subtest

Shuttleworth-Jordan & Bode, 1995b

Digit Symbol

Age	20-39	40-59	60-69	70-79	80-89
N	23	21	28	31	28
X	58.24	51.95	43.55	35.92	31.38
SD	8.82	8.02	10.96	9.81	5.98

VISUAL MEMORY

Digit Symbol Incidental Recall

Shuttleworth-Jordan & Bode, 1995b

Digit Symbol Incidental Recall

Age	20-39	40-59	60-69	70-79	80-89
N	23	21	28	31	28
X	7.30	5.48	6.16	3.50	3.11
SD	1.54	1.85	2.15	1.48	2.14

HAND MOTOR ABILITY

Finger Tapping Test

Shuttleworth-Jordan & Bode, 1996

Finger Tapping Preferred Hand

Age	20-39	40-59	60-69	70-79	80-89
N	22	21	28	26	26
X	5.28	5.49	5.93	6.49	6.71
SD	0.87	O.97	1.29	1.40	1.32

Finger Tapping Nonpreferred Hand

Age	20-39	40-59	60-69	70-79	80-89
N	22	21	28	26	26
X	5.26	5.36	5.95	6.62	6.95
SD	0.77	0.86	1.28	1.35	1.44

TABLE D. PHASE 2 DEMOGRAPHIC DATA

AGE RANGES, EDUCATIONAL/IQ LEVELS

TEST	AGE RANGES	EDUC(YRS)/IQ[1]	STUDY
VISUAL SKILLS			
Trail Making A+B	60's, 70's, 80's	Ed 15, 15, 15 IQ 128, 129, 128	Cornfield & Shuttleworth-Jordan, 1996
Digit Symbol	20-39, 40-59, 60's, 70's, 80's	Ed 16, 15, 14, 15, 15	Shuttleworth-Jordan & Bode, 1995b
VISUAL MEMORY			
Digit Symbol Inc Recall	20-39, 40-59, 60's, 70's, 80's	Ed 16, 15, 14, 15, 15	Shuttleworth-Jordan & Bode, 1995b
MOTOR ACTIVITY			
Finger Tapping	20-39, 40-59, 60's, 70's, 80's	Ed 16, 15, 14, 15, 15	Shuttleworth-Jordan & Bode, 1996

[1] The data for education and IQ are ordered to match the order of the age groups to which they relate. For the purposes of this summary all figures reported for Education and IQ have been rounded to the nearest integer. (Statistical comparisons across groups for more detailed Education and IQ levels are tabulated in the results section).

TABLE E. LEVELS OF SIGNIFICANCE FOR THE GROUP PAIRWISE COMPARISONS USING BONFERRONI'S ADJUSTMENT FOR SIMULTANEOUS TESTING.

Number of Groups	Number of Comparisons	10% *	5% **	1% ***
3	2	0.05	0.025	0.005
4	3	0.033	0.0166	0.0033
5	4	0.025	0.0125	0.0025
6	5	0.02	0.01	0.002
7	6	0.0166	0.0083	0.0016