Patterns of Cognitive Impairment in Dementia

Robert Taylor

Ph.D Thesis

University of Edinburgh

1986

Acknowledgements.

All the people who were of assistance in the completion of this study are too numerous to be named individually, but I would particularly like to thank:

Dr. Chris J Gilleard and Mr. Ralph J McGuire, my supervisors, for invaluable advice and guidance in the planning, conduction, and analysis of the study.
Drs. W D Boyd, J B Loudon, and A Phanjoo; Sisters L Taylor and H Willis; Charge Nurses R Harper and D Norrie; and other nursing staff at the Jardine Clinic, Royal Edinburgh Hospital.
Sisters F Cook and M Findon and other staff at the Balfour Pavilion, Astley Ainslie Hospital, Edinburgh.
Drs. B Ashworth, E Jellinek, and the late C Mawdsley; and other medical and nursing staff at the Northern General Hospital, Edinburgh.
Dr. R Cull and other medical and nursing staff at the Royal Infirmary of Edinburgh.
Mr. A Grimes, Miss E Jones, and Mr. W Randall at the Lothian Region Social Work Headquarters; Mrs E Brown and all other staff at a Local Authority old people's home near Edinburgh, and several local General Practitioners concerned with its residents.
Drs. P Fawcett and L Waddell; Charge Nurse B Taylor and other nursing staff at Herdmanflat Hospital, Haddington.
Drs. A Jacques, D Morrison, and A Stewart; Sister Clancy, and other medical and nursing staff at the Royal Victoria Hospital, Edinburgh.
Dr. A R Milne, Chairman of the North Lothian Ethics Committee.
Miss S M Mackay, Secretary to the Faculty of Medicine.
Mrs. M Dodd and Mrs. M Granger of the Edinburgh University Department of Psychiatry.
Miss E S Duncan, Medical Records Officer at the Royal Edinburgh Hospital.
Dr. J Naughton for previous advice and encouragement.
Sally Anderson, Ellen Taylor, Donald Taylor, and Keith Taylor for moral and material support.
And, last but by no means least, all the people who participated in this study as subjects.

With the exception of this acknowledged assistance, this study was conducted and the thesis composed entirely by myself.

Robert Taylor

The study was supported by an Edinburgh University Faculty of Medicine Postgraduate Research Scholarship.

Contents

	Page
Abstract	1
Chapter 1: Introduction	2
Dementia of the Alzheimer Type	6
Clinical Features and Cognitive Impairments	7
Structural Neuropathology	10
Chemical Neuropathology	12
Aetiology	13
Theoretical Perspectives	14
One Disease or More?	19
Clinical and Psychological Studies	20
Neuropathological Studies	29
Genetics	33
Accelerated Ageing?	34
Multi-Infarct Dementia	39
Prevalence of the Dementias	45
Clinical Diagnosis	47
Staging of Dementia	48
Other Conditions Involving Cognitive Impairment	50
The Absence of Normal Control Subjects	52
Other Methodological Problems in Age-Related Research	55
Why Assess the Cognitively Impaired Elderly	59
Chapter 2: Patterns of Cognitive Impairment	63
Tests	64
Subjects	83
Results: Patterns of Neuropsychological Performance	90
Focality of Performance	122
Inter-individual Variability	127
The Influence of Premorbid IQ	129
The Influence of Drugs	132
The Influence of Sex	134
The Influence of Duration of Illness	136
Other Possible Influences	137
The Factor Structure in the Three Groups Separately	138
Chapter 3: Decline Over Time	145
Subjects	146
The Representativeness of Retested Subjects	151
Results and Discussion	155
The Prediction of Decline	175
Changes in Focality	185
Test-Retest Reliability	187

Page

Chapter 4: Relationships Between Test Performance and Real Life 192

Measures 194
Subjects 197
Results 197
Changes over Time 206
Behaviour During Testing 210

Chapter 5: Variability Over Time in DAT and MID 216

Subjects 217
Method 218
Results 219
Test-Retest Reliability 229

Chapter 6: Further Analysis of Memory Impairment 230

Response Competition in Recognition Memory in DAT, MID and Korsakov's Syndrome 240
Subjects 241
Method 242
Results and Discussion 244
Encoding Preferences in DAT, MID, and Depression 253
Subjects 255
Method 256
Results and Discussion 258
A Signal Detection Analysis of Recognition Memory in DAT, MID, and Depression 264
Subjects 266
Method 266
Results and Discussion 267
Memory Performance in the Main Study:
Paragraph Recall 272
Face-Name Learning 277
Yes-No Picture Memory 281
Orientation 284

Chapter 7: Further Analysis of Language Impairment 289

Object Recognition and Nominal Ability 292
Sentence Production 304
Comprehension Deficit 310

Page

Chapter 8: Further Analysis of Other Abilities 318

Psychomotor Performance 319
Motor Apraxia 328
Spatial Span 332
Drawing Ability 337

Chapter 9: Concluding comments 341

References 352

Appendices

(1) Details of the test procedures 370
(2) Details of the test materials 395
(3) Inter-test correlations 422
(4) Details of factor analyses 432
(5) Individual subject graphs 438
(6) A published paper 458

Abstract.

A series of studies is presented investigating characteristics of the cognitive impairments found in Dementia of the Alzheimer Type (DAT), Multi-Infarct Dementia (MID), and other neurological and psychiatric conditions causing cognitive impairment.

Introductory material includes suitable definitions of the diagnostic conditions, descriptions of the nature of DAT and MID, and discussion of issues concerning whether important differences exist within the DAT category depending on the age at which the condition first appears.

The main initial study is a cross-sectional study of patterns of cognitive impairment, as assessed by a battery of neuropsychological tests, in groups of DAT and MID subjects of different ages and in a group of subjects with other conditions resulting in cognitive impairment. 58 DAT, 58 MID, and 58 other subjects are fully assessed; additional subjects receive only a short form of the test battery. Differences in patterns of cognitive impairment between diagnostic groups and between different age groups within diagnostic groups are described; the possible significance of these findings is discussed with reference to any possible artefacts arising from methodology or subject selection procedures.

A follow-up study of samples of these groups involves a second neuropsychological assessment 10 months after the first. Patterns of decline in different groups are described, and efforts are made to identify predictors of the extent of decline over 10 months based on subjects' test performances or personal characteristics at initial assessment. The extent of decline shows considerable variation even within diagnostic groups, and is only rather weakly predictable using the information gathered at initial assessment.

The remainder of the thesis comprises a number of relevant smaller studies. The first is an investigation of relationships between neuropsychological test performance and everyday functioning as assessed by a behaviour rating scale in ninety of the subjects initially tested as described above. The significance of the relationships found is discussed with reference to clinical psychological practice. The second is a study of short-term day to day variability in cognitive functioning in small groups of DAT and MID subjects involving three administrations of a brief test battery within a two-week period. Little variability is found in either group, and the notion that MID subjects characteristically fluctuate more from day to day than DAT subjects do is not supported. The rest of the studies involve detailed analyses of certain particular aspects of cognitive functioning in DAT, MID, and other cognitively impaired subjects. Some draw on data collected in the main initial study described above (previous presentation of such detailed analyses having been inappropriate when considering overall patterns of performance); others are separate experiments, with smaller numbers of subjects. The data presented principally concern aspects of memory, language function, and psychomotor performance. A variety of findings from these detailed analyses are noted and their importance discussed.

Finally the findings of all the studies are reviewed and some tentative conclusions drawn.

CHAPTER 1

Introduction.

The general purpose of the studies described in this thesis is to investigate characteristics and patterns of the cognitive impairments occurring in the two major categories of dementia in the later years of life: dementia of the Alzheimer type (DAT) and multi-infarct dementia (MID).

The layout of the thesis is as follows. This first chapter includes a selective review of relevant studies and general introductory material including suitable definitions of the diagnostic conditions, descriptions of the nature of DAT and MID, and discussion of issues concerning whether important differences exist within DAT depending on the age at which the condition first appears.

The main initial study, described in Chapter 2, is a cross-sectional study of patterns of cognitive impairment, as assessed by a battery of neuropsychological tests, in groups of DAT and MID subjects of different ages and in a mixed group of subjects with other conditions resulting in cognitive impairment. This primarily concerns differences in patterns of neuropsychological impairment between diagnostic groups and between different age groups within diagnostic groups. Such differences are considered in relation to variables such as overall level of impairment, duration of condition, estimated premorbid IQ, use of medication, gender, and possible artefacts arising from methodology or subject selection procedures.

A follow-up study of samples of these groups is described in Chapter 3. This involves a second neuropsychological assessment, 10 months after the first, of as many subjects as possible within the time limits of the study. This concerns comparison of the patterns

of decline over time in various diagnostic groups and sub-groups, with attempts to identify predictors of the extent of decline over 10 months based on subjects' test performances or personal characteristics at initial assessment.

The remainder of the thesis comprises a number of relevant smaller studies. An investigation of relationships between neuropsychological test performance and everyday functioning as assessed by a behaviour rating scale in a large sample of the subjects initially tested is presented in Chapter 4.

Chapter 5 contains an account of a study of short term day-to-day variability in cognitive functioning in small groups of DAT and MID subjects involving three administrations of a brief test battery within a two-week period. This tests the assertion that mental state commonly fluctuates from day to day in MID but not in DAT, with the implication that results of single cognitive assessments in subjects with MID are likely to be unreliable. It was not primarily designed as a study of test reliability, though it can be seen as such.

The rest of the studies involve detailed analyses of particular aspects of cognitive functioning in DAT, MID, and other cognitively impaired subjects. Some consist of detailed analyses of performance on certain tests which were 'built in' to the main neuropsychological test battery with such analyses in mind, but where previous presentation of such analyses would have been inappropriate when considering overall patterns of performance. Others are separate experiments, with smaller numbers of subjects. Chapter 6, on memory, describes a study concerning inhibition of competing responses in verbal and non-verbal recognition memory in

DAT, MID, and Korsakov's syndrome; a study of encoding preferences in verbal memory in DAT, MID, and depression; a signal detection analysis of non-verbal recognition memory in DAT, MID, and depression; and, from the main study, analyses of performance on a paragraph recall test, a face-name learning test, a picture-recognition test, and an orientation questionnaire in DAT, MID, and other conditions. In Chapter 7, on language, appear studies of recognition and nominal ability, sentence production, and comprehension deficit in DAT, MID, and other conditions. Chapter 8 includes studies of psychomotor performance in DAT, MID, and other conditions; and motor apraxia, spatial block span, and drawing ability in DAT and MID. General concluding comments appear in Chapter 9.

This chapter will contain material of particular relevance to the two main studies described in Chapters 2 and 3. Material specifically relevant to the remaining studies will appear later. Because of the size of the literature on ageing and dementia, reference to previous work is necessarily highly selective.

Dementia is a devastating condition which is becoming increasingly common as a result of increasing survival into old age as well as birth rate trends several decades ago. It can be considered one of the major health and social problems of industrialised societies.

Of the numerous definitions of dementia available, the one given in DSM-III (The Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition) is typical. Dementia is

"a loss of intellectual abilities of sufficient severity to interefere with social or occupational functioning. The deficit is multifaceted and involves memory, judgement, abstract thought, and

a variety of other higher cortical functions. Changes in personality and behaviour also occur. The diagnosis is not made if these features are due to clouding of consciousness, as in delirium."

A major American working party recently included in a similar definition a clear statement that

"Dementia is a diagnosis based on behaviour and cannot be determined by computerized tomography, electroencephalography, or other laboratory instruments, although specific causes of dementia may be identified by these means" (McKhann et al, 1984).

By far the most common types of dementia are dementia of the Alzheimer type and multi-infarct dementia, and these are the types principally studied in this thesis. The first will be dealt with in greatest detail.

Dementia of the Alzheimer Type.

Alzheimer's disease was first described by the eponymous German physician at the beginning of this century in a woman in her fifties who had a four year history of dementia and whose brain was subsequently found to contain numerous senile plaques and neurofibrillary tangles at post-mortem examination. According to Grufferman (1978), Alzheimer had previously recognised the presence of such plaques in senile dementia and considered that the disease shown by this woman represented the onset of senile dementia at an early age. The significance of such findings was not widely recognised: a belief that senile dementia was generally caused by arteriosclerotic disease (either with or without the presence of actual infarcts) persisted for many years, for reasons reviewed by Schneck et al (1982).

In the last 20 years, however, pathological studies have shaped the current viewpoint that the majority of cases of dementia in the senium are not attributable to arteriosclerotic disease (since the

arteriosclerotic changes are no more severe than those seen in normal elderly people): they are associated with the Alzheimer-type plaques and tangles. Dementia which is associated with vascular disease is described later. The existence of mixed forms of dementia involving both Alzheimer and vascular changes is well recognised, but this casts no doubt on the basic distinction.

Apparent similarities between Alzheimer's presenile dementia and the non-vascular senile dementia led to a view that the conditions are essentially similar, differing only in age of onset, and the term Dementia of the Alzheimer Type (DAT) has been widely adopted to refer to the condition regardless of the age of the sufferer. The arbitrariness of a distinction based on age (with 65 years the usual point of division) was never universally accepted, and some authorities preferred to retain a nominal distinction using PDAT to refer to the presenile form and SDAT to refer to the senile form. Classification difficulties and uncertainties are also reflected in the adoption of the term 'primary degenerative dementia' by some authors, particularly in America. DSM-III (The Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition) includes this term and maintains a distinction between forms with senile and presenile onset. Recent evidence suggests that adoption of a view of DAT as a single condition may have been premature. This is reviewed later, along with evidence relating to the question of whether DAT can be seen as some form of accelerated ageing.

Clinical features and cognitive impairments.

The clinical picture of DAT by definition involves a progressive deterioration in cognitive functioning. The earliest cognitive change generally concerns memory for new material or recent events,

probably accompanied by decline in general intellectual level. As the condition progresses all manner of neuropsychological deficits may appear. Various changes in behaviour, personality, emotional state, and social functioning are widely recognised. Descriptions of typical clinical presentations abound in papers and textbooks, and need not be repeated here. Semple et al (1982) give a representative summary. Eventual impairment is severe and wide-ranging, unless the sufferer dies before such a stage is reached. Life span is certainly reduced, though estimates of the size of the reduction vary: cause of death is usually attributed to whatever intercurrent condition was the 'final straw', but a number of authorities suggest that DAT itself deserves to be recorded as one of the leading causes of death in developed countries (Katzman, 1976).

The literature suggests that virtually every aspect of cognitive functioning which has been studied shows impairment in DAT, at least at some stage of its progression, compared to normal control populations. This tendency is so widespread that it has been suggested that studies which fail to show differences between DAT and control groups are of more theoretical interest (assuming that the lack of difference is not attributable to artefacts such as ceiling effects in control groups), though few such studies have been published. The most intensively studied aspects of the impairments have concerned general intellectual functioning (as assessed by I.Q. tests) and memory. Memory will be dealt with in Chapter 6.

Findings on I.Q. are reviewed by Miller (1977, 1981) and Woods and Britton (1985), and will be referred to here only briefly since

the studies which follow do not involve I.Q. assessments other than a single test estimate of premorbid intelligence. Both verbal and non-verbal I.Q. tend to be lowered in DAT, whether the measures used are Wechsler's Verbal and Performance I.Q. scales or the Mill Hill Vocabulary and Raven's Matrices combination. The impairment in non-verbal I.Q. is generally greater than that in verbal I.Q. With the Wechsler scales this may be partly attributable to slowness in view of the timed nature of many of the Performance tests, but this does not apply with the Mill Hill and Raven's Matrices combination since neither test is time-limited. Even with the Wechsler scales, allowance of extra time does not necessarily improve Performance scores to levels compatible with Verbal I.Q. The explanation probably also involves the fact that non-verbal tests often assess abilities relating to new problems or materials while the verbal tests tend to require old or overlearned knowledge or skills. The latter may simply be more resistant to the effects of dementia, at least until a relatively late stage. This view can be rephrased in terms of fluid versus crystallised intelligence, or in terms of tendencies towards abstract versus concrete thinking, with debatable explanatory benefit. A further possibilty is that the findings reflect some artefact of test construction or scaling such that certain non-verbal measures are inherently more sensitive to a given amount of deterioration in ability than are the verbal measures (Miller, 1984a). This possibility is naturally very difficult to test.

Clearly, overall I.Q. scores are global measures which encompass a range of abilities. Patterns of performance on various subtests are potentially of more interest, though even I.Q. subtests, like

most clinical tests, can be accused of measuring more than one ability. Studies concerning subtest patterns have produced varied findings and have not shown consistent identifiable patterns, perhaps partly as a result of major methodological difficulties in the study of ageing and dementia (which will be discussed later). Approaches involving calculation of deterioration indices based on relative levels of performance on various subtests have generally not been fruitful. Such patterns as have been described in studies using I.Q. subtests and other batteries of neuropsychological tests will be considered below in a sections on whether DAT can be considered as one or more conditions or as accelerated ageing.

Other aspects of cognitive impairment, such as those involving language, spatial, and psychomotor functions, have been studied less intensively. These will be dealt with in later chapters as appropriate. Changes related to personality, emotional state, and social functioning, which do not always correlate strongly with measurable cognitive impairment, are reviewed by Gilleard (1984b, 1984c) and Pearce (1984).

A brief outline of the neuropathology of DAT in general will now be given. Evidence concerning possible differences in neuropathology between subtypes of DAT will be reviewed later.

Structural neuropathology.

The three most commonly described structural neuropathological changes described in DAT are:

(1) Cerebral atrophy, which is now usually assumed to reflect in part a loss of neurons, but which varies greatly in degree between DAT subjects at post mortem examination. Atrophy tends to be most marked in temporal, certain parietal, and certain frontal areas:

primary sensory and motor areas and the occipital lobe in general are less affected (Roth, 1986). Controversy exists over the extent of the contribution of neuronal loss in DAT compared with normal ageing, and over which areas and layers of the cortex show greatest neuronal loss. Single indices of the degree of atrophy in life (such as measures of ventricular size) show at best weak correlations with the severity of dementia apparent clinically; newer measures such as CT regional density may be rather better.

(2) Senile plaques, consisting of discrete microscopic areas of amyloid surrounded by abnormal neuronal processes, found principally in the cerebral cortex and especially in temporal areas. The severity of observable dementia correlates with counts of the number of plaques per unit area (e.g. Tomlinson et al, 1970).

(3) Neurofibrillary tangles, consisting of microscopic 'paired helical filaments' of abnormal protein-based material, again found principally in the cerebral cortex and especially in temporal areas. Again, some correlation between the severity of observable dementia and tangle density has been reported (Farmer et al, 1976).

The presence of granulovacuolar bodies (particularly in the hippocampus) has also been repeatedly described. Other structural abnormalities (often observable only using electron microscopy) have been noted, though their significance is not yet known. Many of the structural neuropathological changes are seen (though not in combination with all the others) in similar or identical form in other pathological conditions.

Chemical neuropathology.

The longest-recognised and apparently most severe neurotransmitter defect in DAT affects the cholinergic system, as shown by a variety of studies concerning the activity of choline acetyltransferase (ChAT), a marker of cholinergic neurons. The degree of cholinergic deficit correlates with the severity of dementia shown in life (Roth, 1986). Other biochemical abnormalities have been found, though these are generally not as marked as the cholinergic one. They include abnormalities in 5-hydroxytryptamine (5-HT), gamma-aminobutyric acid (GABA), noradrenaline, dopamine, and the neuropeptide somatostatin, and are reviewed by Rossor & Iversen (1986). Of these changes, the reduction in cortical somatostatin appears to be most clearly related to severity of dementia. Where present, the chemical changes appear to predominate in frontal, temporal, and cingulate regions of the cortex (Roth, 1986). The extent to which the neurochemical abnormalities are specific to DAT as opposed to other brain diseases is not fully known. Studies comparing DAT with MID and Parkinson's disease show differences rather than similarities between conditions.

Correlations in the expected directions exist between different neuropathological measures (chemical deficits, neuronal loss, plaque and tangle counts and so on). Causal relationships between the various abnormalities have not so far been established. Correlations between neuropathological abnormalities and severity of dementia in life, though often statistically significant, tend to be small in absolute terms. Stronger actual relationships may exist but be obscured because of methodological factors reviewed by

Miller (1977). These include the inherent unreliability of measures; varying premorbid levels on any of the measured variables; non-linear relationships between variables (as will be found in the presence of threshold phenomena); restricted score ranges on measures (particularly psychological ones); small numbers of subjects; and the length of time between the taking of psychological and physical measurements in any study using post-mortem data.

Many of the neuropathological abnormalities seen in DAT are known to occur to some extent with normal ageing: their extent often correlates with age in samples of normal elderly people. The difference between DAT and normal ageing can therefore be seen as one of degree, and the hypothesis advanced that DAT represents an abnormal acceleration of the normal ageing process. A threshold effect is presumed to operate, whereby dementia only appears once a given level of neuropathological change is reached, and this does not depend on the chronological age of the person. This will be discussed below.

Aetiology.

Some genetic contribution to predisposition to develop DAT is undoubtedly present (Heston et al, 1982), but the size of its effect is debatable. Isolated or sporadic cases are by no means rare. Kay (1986) provides a recent review of the considerable and sometimes contradictory evidence. Different models of the mode of inheritance of any genetic factor have been proposed. Evidence concerning possible genetic differences between subtypes of DAT will be reviewed later. An association between Down's syndrome and DAT, both in terms of a tendency for Down's to be more common in

families where DAT is present and a tendency for most if not all Down's sufferers to develop Alzheimer changes after the age of 40, have led to speculation that any genetic abnormality might be located on the 21st chromosome.

Other possible aetiological factors receiving attention include slow viral agents of the sort found in Jakob-Creutzfeldt disease, Kuru, and animal scrapie; acquired auto-immune disorder involving the presence of antibodies reactive to brain substances; toxic agents such as aluminium; previous thyroid disease; previous head injury; and dietary factors. So far none of the aetiological hypotheses involving these factors is proven.

Pharmacological methods of treatment directed at bolstering the cholinergic system by administering precursors or trying to improve cerebral circulation using cerebral vasodilators have had limited success. Psychological and behavioural approaches to management and amelioration of deficits show promise, but await further development and evaluation: a review of such approaches is beyond the scope of the thesis.

Theoretical perspectives on DAT.

Miller (1977) points out that two broad types of theory about dementia can exist. In the first a given factor is proposed as contributing to dementia in a causal manner, while in the second consistent relationships between given factors and dementia are delineated in the manner of a law in physics.

Miller reviews attempts at causal psychological explanations of dementia. These have been unsuccessful: the widely accepted view is that dementia results primarily from the brain changes rather than from aspects of the person's psychological make-up. This of course

is not to deny the undoubted importance of psychological factors in the expression of and reaction to the condition. There is at present no adequate non-psychological causal theory of DAT in that the aetiology, as mentioned above, is unknown.

Descriptive psychological models of the nature of change in dementia are reviewed by Miller (1977) and Woods & Britton (1985, Ch 4). These include models relating to general arousal or cerebral excitation levels (including Kendrick's concept of a dual arousal system related to functioning of reticular and limbic brain systems), mental speed or speed of information processing (including versions couched in terms of impaired subjects' systems operating at low signal to noise ratios), concepts of 'core' memory disturbance, verbal-nonverbal and crystallised-fluid dichotomies in intellectual performance, changes in cognitive style, social and sensory deprivation, operant analyses in terms of adaptive responses to declining abilities, and developmental approaches including developmental reversal (where the progressive loss of abilities in dementia is supposed to mirror the progressive acquisition of those abilities in normal childhood development, abilities acquired earliest in childhood being last to be lost in dementia). The concept of accelerated ageing is described below. The various approaches are not mutually exclusive. They cannot all be dealt with here.

Miller (1981a) argues that, in the context of our present levels of knowledge about the nature of dementia, psychological research on dementia should either be descriptive or be based on explanatory models which are confined to specific aspects of impairment. Any attempts to consider dementia in terms of a single unifying, all-

encompassing, fundamental deficit are unrealistic at present, and may always remain so.

A number of broad perspectives exist on the nature of DAT and its relationship to normal ageing. The main possibilities are as follows (where PDAT and SDAT refer to DAT with presenile and senile onset respectively and NA refers to normal ageing defined in some ideal way):

(1) NA, PDAT, and SDAT are all aspects of a continuously distributed spectrum of age-related decline. Differences between them are differences of degree not type. PDAT is no more than the extreme end of a distribution (whether normal or skewed), where the biological clock is running particularly fast. SDAT reflects only moderately accelerated ageing.

(2) PDAT and SDAT are essentially the same as each other, differing only arbitrarily as regards age of onset, but are categorically different from NA. PDAT and SDAT constitute one disease.

(3) NA, PDAT, and SDAT are all categorically different from one another, with the last two being classifiable as separate diseases.

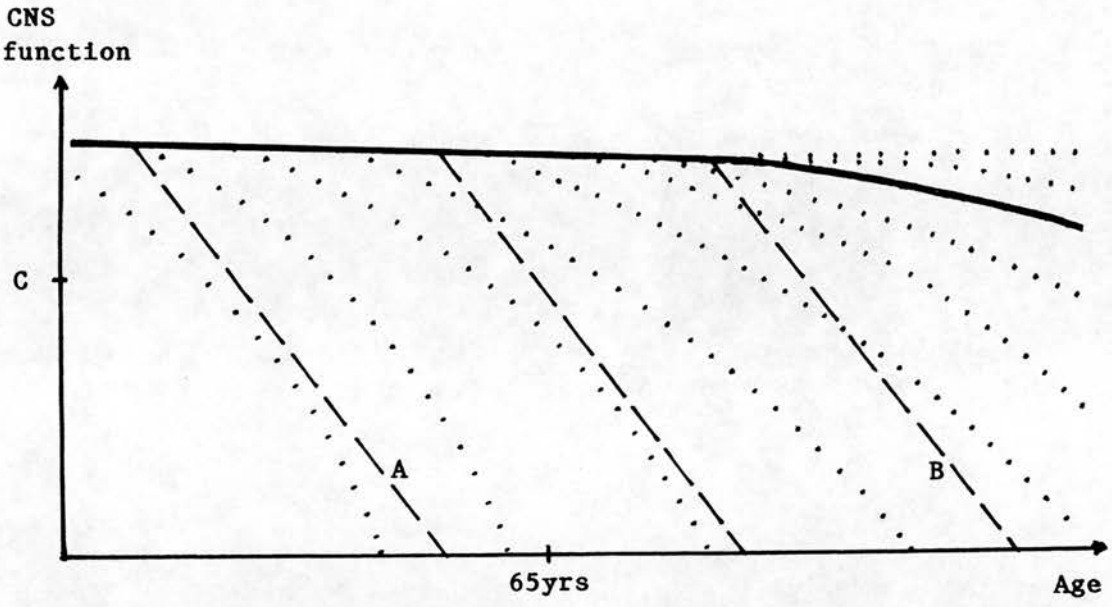
(4) DAT is an umbrella term for a collection of conditions which are all categorically different from NA, so constituting diseases, and which are all categorically different from each other even though methods which could accomplish this differentiation are not currently available. Age of onset may or may not be an important dimension along which the various conditions vary: SDAT and PDAT remain arbitrary labels, each potentially covering more than one condition.

Such crude outlines do of course ignore problems inherent in defining concepts of normality and disease. Threshold effects are

likely to be important in any distinction between DAT and NA. An analogy may be drawn with hypertension where the distinction between normality and disease can only be determined with reference to associated risk.

Some of these conceptions about the nature of DAT are represented schematically in Figure 1, with age along the abscissa and some measure of cognitive or CNS functioning on the ordinate. The solid line represents average performance in normal ageing. The dotted lines represent examples of continuously-distributed deviations from this average. (For simplicity the curves shown are smooth and regular, which might well not actually be the case.) The dotted curves above the solid average line represent 'supernormal' subjects, as described by Savage et al (1973), who show very little evidence of deterioration in cognitive functioning (or in other aspects of physical, social, and psychological well-being). Dotted curves below the solid average line represent subjects who are not so well preserved: those falling below a criterion level C would be regarded as having dementia (though this criterion level C might vary between individuals depending on their premorbid abilities). The onset of dementia cannot be identified except by arbitrary criteria. The dashed lines represent disease concepts of dementia - lines or curves with different gradients or shapes from the 'normal' dotted curves and with relatively identifiable points of departure from the normal curves (representing the onset of dementia). If PDAT and SDAT are in fact different illnesses then lines A and B must be different in some way other than just in their point of departure from normal. This might be in the gradient of the line (steeper in A than in B if PDAT is

Figure 1. Schematic representation of different conceptions of the nature of DAT. Age is represented on the abscissa and some measure of cognitive or CNS functioning on the ordinate. (See text for explanation.)



——— normal or average performance in an ideally representative group of normal people.

····· possible deviations from the idealised norm.

— — disease process.

C a criterion level below which dementia is recognised.

A, B see text.

characterised by more rapid decline); in the shape of the line or curve (straight lines being shown only for simplicity, with the assumption that real graphs might contain all manner of plateaux and other irregularities); or in some qualitative aspect of the measure which cannot be graphically illustrated.

The shapes of the decline curves might be different for different measured aspects of CNS function, and the relative shapes of such different decline curves within cases would have to be taken into account in any detailed syndrome analysis. Such pattern analysis is a daunting task and would require a large and lengthy cross-sequential study using a variety of measures of CNS function. The measures themselves would have to have exquisite sensitivity and ideal psychometric and scaling properties to identify differences between the types of decline curve shown in the figure (Gilleard, 1978).

Nevertheless, studies of manageable proportions can still shed some light on the issues. Some evidence relating to whether DAT can be seen as comprising more than one condition and whether it can be seen as some form of accelerated ageing will now be considered.

DAT: one disease or more?

Controversy over whether Alzheimer's disease and non-vascular senile dementia are one or two diseases is not new, beginning not long after Alzheimer's first description of the disease (Grufferman, 1978). According to Grufferman, Alzheimer himself considered that AD represented the onset of senile dementia at an early age. Various authors (e.g. Katzman, 1976; Terry, 1978; Butler, 1978) have argued that PDAT and SDAT are essentially similar in clinical character and in all aspects of physical

pathology, and that a distinction between PDAT and SDAT is so arbitrary and unsubstantiated that it should be dropped.

Recently, however, evidence has been accumulating that the two are perhaps not as identical as had come to be assumed. This may have important consequences as regards searching for causes and treatments. Some relevant clinical studies will be described first, followed by evidence from neuropathological studies.

Clinical and psychological studies.

Early attempts to identify specific patterns of cognitive impairment in dementia were unpromising. Brody (1942), for example, found the cognitive impairments in dementia to be so severe and non-specific that there was no opportunity to define specific meaningful patterns. Such conclusions might have been influenced by the use of subjects who were too severely impaired to achieve scores other than the minimum possible on some of the tests chosen, with subsequent problems involving floor effects and other artefacts, or by the use of subject groups which would be considered heterogeneous or otherwise inappropriate using present day criteria and definitions. Brody's group of subjects with dementia, for example, seems to have included many chronic psychiatric patients.

In a much-quoted paper entitled 'clinical heterogeneity in senile dementia' McDonald (1969) reported assessing female subjects with senile dementia using a cognitive test battery comprising five brief scales (now widely known as the Kew battery) called memory, aphasia, parietal, abstraction, and Weigl (the last consisting simply of the Weigl sorting test). He established a dichotomous classification based on 'parietal' performance in 51 subjects and

then studied a further 57 subjects using this classification. At 6-month follow-up 26% of the group with poor parietal performance were dead compared with only 4% of the other group, despite the fact that the former group were significantly younger. He concluded that two identifiable groups were present which had different clinical features and outcome. Efforts had been made to avoid using subjects where floor effects would hamper interpretation, and to ensure that subjects in an early stage of dementia were not being compared with those in a later stage. An assumption that the latter objective had been achieved was based on the relative ages of the two subject groups, but this is highly dubious in view of the likely variance in age of onset. Consideration of the limited test data presented concerning the original standardisation sample suggests that the group with poor parietal performance may have been worse on the other scales as well. McDonald chose the parietal scale as a basis for classification on the grounds that it provided a convenient division of subject numbers. It seems that the subsequent study sample were in fact only tested with the parietal scale alone. (Another incidental point is that criteria for excluding MID were less strict than would be applied today, so it is possibile that some of the subjects showing major deficit on the parietal scale had suffered infarcts and perhaps subsequently died from further strokes.) Nevertheless, the possibility remained that there might be more than one subgroup within DAT, even if any distinction was related to age of onset or severity of dementia rather than focal parietal deficit. This has been addressed in various other studies.

Kaszniak et al (1978) investigated relationships between

mortality and initial clinical characteristics and test performances in 35 patients followed up after 12 months. The dead and alive groups did not differ on age, duration of dementia, or degree of cerebral atrophy on CT scanning. They did differ on degree of EEG abnormality and on 8 out of 14 cognitive test measures (always in the expected direction). The most significant single difference was on a test of expressive language, sentence production. It is debatable whether the tests showing significant differences are more 'parietal' than those not showing differences (all of which showed non-significant trends in the same direction). The results can be interpreted in terms of an overall severity effect. The fact that discriminant function analysis using the cognitive test measures yielded highly successful discrimination between the dead and alive groups does not refute this possibility. Nevertheless, their conclusion that expressive language deficit may indicate a particularly poor prognosis for survival may have been justified. They also concluded, on the basis of their own data plus a reanalysis of those of Fox et al (1975), that the degree of functional brain impairment (i.e. cognitive impairment on testing) may be a better predictor of outcome than any structural variable such as the degree of atrophy on CT scanning.

Naguib & Levy (1982) found evidence of an association between early death in DAT and a 'parieto-temporal score': however, the differences between survivors and non-survivors on a mental test score (assessing memory and orientation) was just as significant. Colgan (1985), in an interim report of a study of survival in DAT, found that survival was related to performance on the mental test score but not to performance on the parieto-temporal score used by

Naguib & Levy. Death was associated with the presence of lower parietal and occipital densities on previous CT scanning, in accord with Naguib & Levy's finding of an association with low right parietal density, though Colgan acknowledges the possibility of artefact.

Constantinidis (1978) sees presenile and senile forms of Alzheimer dementia as essentially similar but suggests that an aphasia-apraxia-agnosia syndrome seems to appear sooner, be more serious, and be less regular in chronological evolution in presenile than in senile cases. He suggests that his PDAT group generally had more serious clinical disabilities and a more rapid deterioration despite their greater longevity (which is partly attributable to a straightforward age effect). Interpretation is complicated by the fact that Constantinidis makes a distinction between Alzheimer senile dementia and 'simple' senile dementia and suggests that simple SD may evolve into Alzheimer SD.

Whitehead (1977) re-assessed a mixed group of elderly psychiatric patients after one year using a number of behavioural measures as well as a small battery of cognitive tests including tests of memory, vocabulary, fluency, and parietal function. In general changes were small but in those patients with uncomplicated chronic brain syndromes (presumably DAT) significant drops occurred on most of the test measures and some of the behavioural ones. At four year follow-up (Whitehead, 1982) survival was found to be related to age, with survivors being younger: this provides no support for the idea of younger subjects having a more malignant course, but this was in a mixed diagnostic group. Survival was also related in expected directions to initial level of impairment on

two rating measures and to the degree of decline over the first year on vocabulary, a total learning score, and the parietal score.

Seltzer & Sherwin (1983) compared 34 male DAT subjects having an age of onset before 65 years with 31 having a later age of onset. They found a greater prevalence of language disturbance and gait disorder, a disproportionate number of left-handers, and a shorter length of post-onset survival (relative to expected survival) in the early-onset group. They interpret the results as demonstrating a degree of heterogeneity within DAT and suggesting that an age-based distinction between forms of DAT is not entirely arbitrary. Their conclusions were based on Chi squared analyses of the numbers of subjects in each group with 'normal' and 'abnormal' scores on various tests. Not all subjects received every test. Unfortunately no data are presented to contradict the possibility that the early onset group were simply more demented (i.e. had reached a more severe stage of an identical condition to that suffered by the older group). Weak evidence that this was not the case consisted of observations that the groups did not differ significantly on estimated duration of condition and did not differ significantly on a few other clinical measures.

Hagberg & Ingvar (1976) assessed presenile subjects using a collection of tests of memory, constructional abilities, and language function. Post mortem examination confirmed DAT in most but not all subjects. They distinguished five groups of patients, which differed principally along the dimension of severity of dementia: differences between groups in terms of the presence of particular features was thought to reflect the natural history of the disorder, since there was an orderly appearance of such

features (cross-sectionally speaking) from groups 1 to 5: when impairment on a task was present in one group, all the more impaired groups also showed impairment on that task.

Rosen & Mohs (1982) similarly found a dimension of severity to be important in patterns of impairment found in senile DAT: they distinguished just two groups, classified simply as mild and moderate-to-severe respectively. They also found, however, that examination of patterns of performance (by inspection of profiles of z-transformed test scores) revealed considerable individual variation: for example, though two subjects were equally impaired on one task, their impairments on a second task might differ substantially. They suggest that the variations in performance profiles might reflect the existence of neuropathological subtypes within DAT, variations in premorbid abilities, or some combination of the two.

Brull et al (1979) and Wertheimer (1984) describe a study of performance on a wide range of neuropsychological tests in DAT, with subsequent follow-up and re-assessment over considerable if variable periods. They distinguished three groups. In the first group, subjects did well at initial assessment and showed little decline over time. In the other groups subjects did less well initially, but one group showed stability over time while the other showed decline. Subjects in the first group were younger on average than those in the other groups (providing no support for the idea of younger subjects having a more malignant form) but may have been in an earlier stage of the condition. The two poorer groups, one showing decline and the other not, did not differ from each other in initial characteristics. The authors comment that the patterns

(as opposed to the levels) of impairment at initial testing were on average very similar in the three groups. However, there seems to have been considerable individual variation in both the pattern of impairment at initial testing and the pattern of subsequent decline.

Recently, Neary et al (1986) studied 24 patients with presenile dementia due to cerebral atrophy, 74% of whom were proven on cerebral biopsy to have DAT, using a wide-ranging and well-chosen test battery. They claim to have identified distinctive patterns of neuropsychological impairment, generally based on the presence or relative severity of amnesia, aphasia, perceptuo-spatial disorder, and apraxia. They describe seven patterns in all, including four applying to the proven DAT group. These four patterns refer to subject numbers of 11, 3, 3, and 2 respectively. Subjects in the different categories were not matched for severity of dementia, and some of the group differences had disappeared at later re-assessment. The authors seem to have confirmed the existence of individual variability or heterogeneity in DAT rather than having discovered the 'neuropsychological syndromes' referred to in the title of their paper.

Drachman et al (1982) describe Multidimensional Assessment for Dementia (MAD, unfortunately) scales which cover a wide range of aspects of subjects' psychological and physical characteristics. The authors claim that different types of dementia can be distinguished using these procedures, but so far only preliminary data seem to have been published.

In an unpublished study referred to by Roth (1986), 35 early-onset cases were compared with 86 late-onset cases. Echolalia was

more common in the early-onset cases, and gait disturbance in the late-onset ones. There was no difference as regards 'parietal lobe' features. No information is provided concerning levels of overall severity in the two groups, and there is generally too little detail to judge the importance of the study: preparation for publication is stated to be under way. In the same article Roth comments that these findings with regard to parietal lobe features are in accord with findings of Lauter (published in German), though he later states that Lauter found parieto-temporal focal psychological deficits to be less common and less conspicuous as age of onset increases. The present author has not obtained translations of Lauter's studies. Roth distinguishes three broad stages of dementia. In the first, where memory disorder predominates, and in the last, where dementia is severe, he sees little difference between early and late onset cases; but he feels that some clinical differences probably do exist in the intermediate stage. He sees early and late onset as mere correlates of a distinction between Type I and Type II DAT, which will be mentioned below.

There is then some evidence from clinical and psychological studies to suggest that subtypes of DAT exist. There are suggestions that onset at an early age and the existence of parietal or linguistic (or at least non-memory) disturbances are features signalling a malignant form of DAT, and that the early onset and focal features are themselves associated with each other (especially in particularly familial forms of DAT; Breitner & Folstein, 1984).

One hypothesis relevant to such suggestions (and popular with

the present author at the beginning of this project) is that focal features are more often found in younger (or familial) cases simply because they are more often looked for: it is widely accepted that younger cognitively impaired subjects are more likely to undergo intensive investigation (perhaps including neuropsychological testing) than are older ones. The assumption is that older subjects would show similar impairments if assessed in the same detail as younger ones. In the light of more recent evidence, however, this no longer seems a plausible explanation of many of the findings.

Another important point concerns the significance of parietal or other focal features. As Gilleard (1984b) puts it, a number of the studies suggest that the appearance of certain cognitive impairments 'ahead of time' in dementia (i.e. focal impairments of crystallised cognitive skills in the absence of major overall impairment) indicates a more severe form of the condition with poorer prognosis. Again the point is made that these focal features are most fruitfully considered in the context of abilities in other areas. Any kind of parietal scale, if used in isolation, may be no more than an index of the severity of dementia: the presence of parietal signs might simply indicate a more advanced stage of the condition, one step on from a stage in which impairments are largely confined to the area of memory (assuming that the natural history of DAT might involve such a progression). In these circumstances an association between parietal impairment and early death is trivial, since one would expect more severely demented subjects to die sooner. Any finding of more rapid deterioration in subjects with parietal signs could also be accounted for if the natural history involved steeper decline in later stages of the

condition than in earlier ones. Parietal impairment must be seen in the context of other neuropsychological impairments or in the context of the overall severity or stage of the condition. Some assessment of relative levels of performance in different areas is required rather than the simple identification of impairments in any one area. The studies to be presented adopt such an approach.

The expression 'parietal signs' is often used loosely (despite passing reference to an aphasia-apraxia-agnosia triad), sometimes apparently to refer to any cognitive deficits other than memory deficit. This may help explain why the presence of 'parietal' signs is rarely mirrored by neuropathological reports of marked parietal atrophy (the frontal and temporal lobes often being more affected; Perry, 1986). Gilleard (1984b) suggests that the malignant feature may be disruption of overlearned or established skills involving speech, writing, dressing, handling tools, and so on rather than parietal lobe features per se. Roth (1986) puts forward an explanation in terms of functional disconnection involving the major projections between parietal and fronto-temporal areas, whereby fronto-temporal dysfunction produces deficits traditionally associated with parietal damage. Part of the explanation may lie in the nature of the chemical rather than the structural changes in subjects' brains. Possible neuropathological distinctions between DAT subtypes will now be considered.

Neuropathological studies.

Studies generally rely on post-mortem neuropathological data. Naturally the studies are cross-sectional. The distinction between old and young DAT subjects is generally made on the basis of age at death. This of course correlates with age of onset, but far from

perfectly because of the variance in length of survival after onset (which may itself be slightly different in young and old groups). This is an inevitable limitation if ethically dubious brain biopsy studies are to be avoided. The hypothesis that younger subjects might simply live longer than older ones and so have more time in which to develop severe neuropathology before death and post-mortem examination does not seem capable of accounting for all the findings. Some studies can be criticised for not making adequate comparisons with age-matched control subjects, and for being unable to control for the severity of dementia shown in life by the DAT subjects. There are now enough adequate studies to suggest that all the results cannot be entirely explained in terms of artefact, but conclusions deserve to be cautious.

Hubbard & Anderson (1981) suggest that the pattern of atrophy associated with DAT may be age-related. They found that DAT subjects dying before the age of 80 years showed relatively global cerebral atrophy while those dying after that age usually showed a more selective atrophy affecting primarily the temporal cortex. Perry (1986) agrees that over the age of 80 atrophy (where it is in fact present to a greater degree than in age-matched normal controls) is often restricted to particular regions such as the temporal lobe. Mountjoy et al (1984) report being unable to find significant cortical neuronal loss in DAT subjects dying after the age of 80 years. Significant loss was present (particularly in frontal, temporal, and cingulate regions) in subjects dying younger. Bondareff et al (1982) found that the loss of neurons in the locus ceruleus, an important nucleus in the origins of adrenergic projections to cerebral cortex, was generally more

marked in younger patients, though changes could be present in older ones. The younger patients also had more severe dementia in life according to scores on a dementia scale administered at some unspecified time before death: as in many of these studies, the direction of relationships between age, severity of dementia, and degree of neuropathological change cannot be ascertained. Roth (1986) reviews evidence suggesting that neuronal loss or tangle density in other sub-cortical nuclei (such as the nucleus of Meynert) also tends to predominate in younger patients. Perry (1986) suggests that such findings so far should be considered tentative. Cortical plaques and tangles may be more numerous and more widespread in younger subjects, though evidence relating to this is equivocal (Roth, 1986).

Bowen et al (1979) found that the cholinergic deficit was less severe in older patients (those dying over the age of 80 years) than in younger ones, but interpreted this in terms of the frailer, older cases dying at an earlier stage in the pathogenesis of the disease. Rossor et al (1982, 1984) found that the cortical cholinergic deficit was more severe and widespread (affecting all cortical areas) in younger DAT subjects (those dying before the age of 79 years) than in older ones, where the milder deficit remained principally confined to temporal cortex and hippocampus and spared the frontal lobes. Significant reductions in cortical noradrenaline and GABA were found only in younger patients (those dying before the age of 79 years) by Rossor et al (1984). Again the reductions seemed to be greatest in frontal and temporal areas. Somatostatin showed widespread reduction in the younger patients while in the older ones any reduction was confined to temporal areas. Rossor &

Iversen (1986) suggest that 5-HT deficits in DAT are again largely confined to younger sufferers. Mountjoy et al (1984) felt that the difference between older and younger DAT subjects was elegantly demonstrated by the fact that, in a cross-sectional study, ChAT and GABA concentrations tended to decrease with age of subject in the normal control group but to increase with age of subject in the DAT group. Rossor & Iversen (1986) conclude from a review of the evidence to date that there is a trend for all neurotransmitter abnormalities to be more severe in younger patients.

Significant correlations between different pathological measures (chemical deficits, neuronal loss, plaque and tangle counts etc) tend to be present in more areas of the cortex in younger patient groups (Roth, 1986). The significance of this is not entirely clear, and some artefact may be present if older subjects actually have very low scores or restricted score ranges on some of the indices of abnormality.

Trapp et al (1978) reported increased levels of brain aluminium in PDAT but not in SDAT subjects when comparison was made with age-matched controls. As with other studies relating to aluminium, the significance and reliability of this finding have not been established.

Recently Roth (1986) has described a dichotomous classification of DAT into Type I and Type II, with the latter term referring to the form showing more severe and widespread pathological abnormalities. This Type II generally has an earlier age of onset, but this is not invariable: age of onset, he says, provides only a crude criterion of differentiation between the two syndromes.

Genetics.

A variety of suggestions have appeared to the effect that subtypes of DAT could be discriminated on genetic grounds. Some studies, as reviewed by Grufferman (1978), have suggested that relatives of patients with PDAT have increased risk of developing PDAT but not of developing SDAT, while relatives of patients with SDAT have increased risk of developing SDAT but not of developing PDAT. This would appear to support a distinction between conditions, but interpretation of the data is not straightforward (Katzman, 1976) and other studies have produced inconsistent findings. Heston et al (1982) concluded that the genetic contribution in cases with onset after the age of 70 was small. Support still exists for the possibility suggested by Larsson et al (1963) of a single genetic model involving a predisposing autosomal dominant gene with age-related penetrance, where the failure of some genetically prediposed individuals to develop the condition is due to the fact that they do not live long enough. The existence of particularly familial types with characteristic features including young onset, prominent parietal and dysphasic signs, and rapid deterioration (Breitner & Folstein, 1984) seems to suggest that a unitary genetic model is not tenable. Morris et al (1984) suggest that some cases of DAT can be seen as belonging to a spectrum of hereditary dysphasic dementia which also includes Pick's disease and where the mode of transmission is autosomal dominant. Genetic factors in DAT show variations between different countries, for example in the existence of more and stronger family pedigrees, and of more white matter change on neuropathological examination, in Scandinavia than in this country (Mountjoy C Q, unpublished

observation).

Kay (1986) concludes that two opposing sets of observations exist, one suggesting that the risk of a relative developing DAT varies with the severity and age of onset of DAT in the proband (a pattern suggestive of polygenetic transmission) and the other that DAT is an autosomal dominant disease, characterised by language disorder, with penetrance dependent on age. Other more complex genetic models have been proposed to account for the disparate findings, though to the uninitiated these seem irrefutable and pragmatically equivalent to a statement that some people get it and some do not.

The clinical, pathological, and genetic evidence does therefore suggest some differences within DAT. However the differences between possible subtypes remain smaller than the similarities between them, and Miller's (1977) assertion that any differences between subtypes are likely to remain much smaller than the differences between DAT and other conditions certainly still holds. The possibility remains that DAT comprises not just one or two forms, but a group of related disorders which might be differentiable on the basis of a number of dimensions. Age is certainly the dimension which has received most attention, but others (concerning pathological characteristics, genetics and so on) could be important. Further search for different patterns of cognitive impairment within DAT seems justified.

DAT: accelerated ageing?

Before turning to the question of whether DAT can be seen as accelerated ageing, some consideration must be given to changes associated with normal ageing.

Theoretical psychological perspectives on normal age-related cognitive changes are reviewed by Miller (1977, Ch 2) and Woods & Britton (1985, Ch 2). Many of the ideas are similar to those previously listed as relevant to psychological descriptions of dementia and will not be considered here. Again the various views are not mutually exclusive. Purely biological explanations of ageing in terms of inefficient functioning or division of cells are clearly important but do not necessarily shed much light on the psychological changes associated with ageing.

The same authors provide summary reviews of the topic of normal age-related decline. Woods & Britton (1985) suggest that, allowing for methodological problems in studying such phenomena and the scarcity of adequate studies, the overall trend as regards IQ scores is one of stability until the age of about 60, with some decline in Performance abilities after 60 and then some decline in Verbal as well as Performance abilities after about 70. They suggest that age-related decline in cognitive abilities is a feature of old age, not middle age as was once thought, and that its extent varies considerably in different areas of functioning. Cohort differences are probably of greater significance than are age changes within groups, at least for the majority of the life span. 'Normal' age-related decline generally seems not to be sufficiently severe to compromise everyday functioning.

Attention has been paid in the literature to specific areas of cognitive function in normal ageing, including aspects of memory, perception, visuospatial functioning, information processing, attention, problem solving abilities, motor skills, and language. Again it is not appropriate to review such research here.

Particular relevant points will be included in later chapters as appropriate. The phenomena of terminal decline and terminal drop will be referred to in a later section of this chapter. Similarities between the types of neuropathological change seen in normal ageing and in DAT were noted earlier.

As with the question of whether DAT has more than one form, the question of whether it represents an abnormal acceleration of the normal ageing process is relevant to the search for causes and treatments. Such a hypothesis does not imply that DAT with all its dreadful features should be considered normal, as some critics of the idea seem to assume. Certainly the extent of acceleration must be large to account for the severity of impairments seen.

Terry (1978) suggests that all differences between DAT (particularly SDAT) and normal ageing are quantitative not qualitative, differences of degree rather than of type. He considers SDAT to represent a threshold phenomenon, with a strong possibility of a continuous spectrum for normalcy to dementia. Authors such as Katzman (1976) and Butler (1978) argue against a distinction between PDAT and SDAT but feel that DAT is a disease entity distinct from normal ageing.

Psychological studies comparing Wechsler subtest performances of DAT subjects with those of normal elderly subjects have produced no conclusive answer to the question of whether the nature of the cognitive impairment in dementia is qualitatively rather than just quantitatively different from that seen in normal ageing, i.e. whether the patterns are essentially similar or different (Miller 1977; Woods and Britton, 1985). Whitehead (1973a), for example, suggests that the general pattern of impairment on WAIS subtests is

similar in normal ageing and dementia in that performances are highest on Vocabulary and other Verbal subtests and lowest on Performance subtests. However, clear deficits appear in the Verbal tests themselves in dementia while in normal ageing this may only occur in terminal decline phases. Botwinick and Birren's (1951) results suggested some differences between dementia and ageing in patterns of WAIS impairment. Problems of interpretation inevitably arise since many patients with dementia will be performing at or close to the floor level on certain WAIS subtests, particularly the Performance ones. Use of any of the various WAIS deterioration indices has not been fruitful.

Adequate testing of an accelerated ageing hypothesis is methodologically complex and would require use of normal control subjects who are older than the study group of DAT patients and matched on some criterion variable thought to be a reasonable index of general degree of CNS deterioration (Miller, 1977). Cohort effects are an inherent possibility with such a design, but might be overcome in a cross-sequential study. The problem of choosing appropriate matching instruments (given the impracticability of neuropathological matching) is not insurmountable. Nevertheless few studies seem to have adopted this suggestion. One example is a study by Kaszniak et al (1979) in which groups of relatively young and relatively old elderly people, with and without dementia, were tested on paired-associate learning and digit span. The first was impaired in both DAT and normal ageing while the second was impaired only in dementia, suggesting a difference between DAT and ageing.

Of course, if threshold phenomena occur in DAT, whereby certain

behavioural or cognitive impairments only become manifest after certain thresholds of organic deterioration have been passed, then a smooth and normal distribution of apparent age-related decline may not be found: in these circumstances, empirical differentiation of accelerated ageing and disease process becomes extremely difficult.

Some forms of DAT may be more compatible with an abnormal ageing hypothesis than are others. Rossor et al (1984) and Roth (1986) review abnormal ageing hypotheses in the light of recent neuropathological findings. They feel that such a hypothesis is not compatible with DAT as a whole. Certain features remain consistent with abnormal ageing in that they differ between DAT and normal ageing only in degree or number: this applies to cerebral atrophy, the presence of plaques and tangles in the cortex, and changes in the cholinergic neurotransmitter system. However, some features observed (primarily in patients with early-onset DAT) are so far thought to have no counterparts in normal ageing: this applies to neuronal loss in sub-cortical nuclei and some of the non-cholinergic neurotransmitter deficits. The most striking discrepancies between DAT and normal ageing, then, are seen in younger DAT patients. Roth concludes that while an abnormal ageing hypothesis at present seems untenable as regards Type II or early onset DAT, it remains a possibility as regards Type I or late onset DAT. He stresses that the issue is by no means settled: "Views about the nature of the relationships between these phenomena have had so often to be revised during the last half century that it would appear prudent to maintain an open mind on the subject."

Multi-Infarct Dementia.

The condition now generally referred to as multi-infarct dementia has been known by a variety of names in the past, including arteriosclerotic dementia, atherosclerotic dementia, atheromatous dementia, and cerebrovascular dementia. It has also at times been included under the umbrella term of senile dementia. The underlying assumption behind the older names was that the dementia was caused by a general hardening and narrowing of the cerebral blood vessels, due to a loss of elasticity and the accumulation of atheromatous porridge on their inner surfaces. This was presumed to cause an overall lack of adequate perfusion of blood, and therefore oxygen and glucose, to the tissues of the brain and so cause impaired function. The change in nomenclature arose from findings that, at post mortem examination, the brains of subjects who had suffered from the condition showed not just hardening and narrowing of blood vessels but also accumulations of strokes, large and small. It is now generally thought that the condition always involves such an accumulation of infarcts and that vessel disease in itself is not sufficient to cause it (though of course the vessel disease will contribute to the production of infarcts and may also still produce poor perfusion. Cerebral blood flow is known to be reduced but, as in DAT, it is difficult to know to what extent reduced supply results simply from the reduced demand of a damaged brain).

The infarcts themselves result from three main types of cerebrovascular accident, each of which can of course vary greatly in severity and consequence: (1) Haemorrhage, where a vessel bursts or leaks: adjacent areas are damaged by the flooding, and areas

served by the vessel and distal to the burst may be damaged through starvation of blood. (2) Thrombosis, where a vessel is plugged by the eventual accumulation of atheroma and clotting blood at some particular location, and areas served by the vessel and distal to the plug are starved. (3) Thromboembolism, where a vessel (perhaps already narrowed to some extent) is plugged by an embolus of clotted blood or atheroma carried by the circulation from elsewhere in the cardiovascular system such as the heart itself, and distal areas are again starved.

In all cases the severity of the consequences will depend to some extent on the availability of alternative blood supply routes which might reduce the extent of infarction. Thrombosis and thromboembolism are said to be more common than haemorrhage. The distinction is difficult to make clinically in life.

The clinical features of MID must to some extent depend on the particular areas of the brain which have suffered infarction. One would expect these to vary considerably from person to person, though Matthews & Miller (1979, p156) suggest that atheroma, for reasons unknown, has a predilection for particular parts of particular vessels. Perhaps because of the assumption of a straightforward relationship between impairments and sites of damage, and the fact that MID subjects will be considered a heterogeneous group because of the varying locations of their pathology, much less research has been carried out on the cognitive impairments seen in MID than on those seen in DAT or single major stroke. (In earlier times, of course, MID subjects were doubtless included in studies of subjects with dementia, but without being differentiated as a separate group.)

Differences between typical clinical presentations of DAT and MID are embodied in the Hachinski index (Hachinski et al, 1975), the most common instrument used clinically to differentiate the two conditions once other causes of the dementia or apparent dementia have been excluded. MID subjects are considered more likely to show: a relatively abrupt and dateable onset (as opposed to the insidious onset of DAT), resulting from the first significant infarct; stepwise rather than gradual, steady decline (as a result of serial infarcts perhaps with intervening periods of no change or slight recovery); short term fluctuations in course, probably as a result of variable general inefficiences in the cardiovascular system and the occurrence of transient ischaemic attacks; nocturnal confusion (though this is debated by some authors), perhaps again involving associated general cardiovascular abnormalities; relative preservation of personality, often including insight into the condition and awareness of the problems and disabilities; depression, perhaps as a result of direct effects of infarcts (especially subcortical ones) as well as a psychological reaction involving the maintained awareness of declining abilities; somatic complaints, again perhaps connected to the retained awareness of problems and symptoms in general (since many old people are likely to have physical problems of some sort) and possibly exaggerated by any depression (since somatic symptoms are known to be common in the presentation of depression in older people); emotional incontinence, thought to result from deep infarcts as in pseudobulbar palsy, where tearfulness and occasionally laughter begin and end abruptly with little apparent relationship to circumstances or the content of any conversation; and finally, as

one would expect considering the nature of the condition, a history of hypertension (because of its association with infarction), a history of strokes, evidence of significant atherosclerosis anywhere in the body, any focal neurological symptoms reported by the patient, and any focal neurological signs found on examination. Points are assigned to each of the thirteen features as follows:

Abrupt onset	2
Stepwise progression	1
Fluctuating course	2
Nocturnal confusion	1
Relative preservation of personality	1
Depression	1
Somatic complaints	1
Emotional incontinence	1
History of hypertension	1
History of strokes	2
Evidence of associated atherosclerosis	1
Focal neurological symptoms	2
Focal neurological signs	2

A score of 7 or more is considered to be compatible with a diagnosis of MID, 4 or less with one of DAT. Intermediate scores are ambiguous, and do not necessarily indicate a mixed form of dementia (i.e. DAT plus MID): a mixed form might be present with any high score on the scale.

Otherwise, the clinical features and cognitive impairments seen in MID are not known to be strikingly dissimilar to those seen in DAT. Perez et al (1975) analysed performance on the WAIS in patients with MID, DAT, and vertebrobasilar insufficiency; they found differences between groups and reported that 74% of the subjects could be correctly classified using the results of discriminant function analyses of WAIS subtest scores. The present study does not make use of WAIS subtests, and so details of Perez et al's results are not of direct relevance; it also remains to be seen whether such apparently impressive classification could be

achieved in any sort of prospective study using the discriminant equations with a new group of patients. Even if it could, such discrimination is unlikely to be reliably applicable to the individual case in clinical settings (Miller, 1981b). In addition, though the groups could be differentiated with considerable accuracy, 'typical' patterns of cognitive impairment for each diagnostic group were not observed.

Gainotti et al (1980) review a number of WAIS-based studies which have failed to differentiate characteristic patterns of impairment in various types of dementia. Gainotti et al themselves used a 'mental deterioration battery', consisting of three verbal and three visuospatial tasks yielding eight standardised scores, to assess performance in DAT and MID as well as in Huntington's chorea, Parkinson's disease, normal pressure hydrocephalus, and depression. They claim to have found some differences between the diagnostic groups, though in fact the differences appear to be largely related to severity rather than pattern of impairment. The numbers of subjects are small, and little statistical analysis is presented. All the reservations stated above with respect to the Perez study also apply here.

Gustafson & Nilsson (1982) suggest methods of differentiating DAT, Pick's disease, and MID using a combination of behavioural and basic neuropsychological characteristics, but here again separation on such clinical grounds is retrospective and still far from perfect.

Constantinidis (1978), after summarising some clinical features of DAT, states that all of the clinical symptoms may also be found in dementias of vascular (i.e. MID) or combined (i.e. DAT plus MID)

origin. However, he says, "in pure degenerative cases (i.e. DAT alone), simultaneous disintegration of operational, amnesic, instrumental, and motor functions is observed. This homogeneity of joint disintegration is characteristic of degenerative dementias. It implies that when a sign of defect appears in one function, another one can be found at the same moment in another function." He also states, after a brief description of four stages of dementia, that "the progressive onset of the aphasia-apraxia-agnosia syndrome is scheduled according to relatively precise rules when the disease is of degenerative aetiology". Results will be presented which suggest that these assertions are only very loosely true.

Significant aetiological factors in the production of MID are assumed to be those thought relevant to cardiovascular disease in general, including factors relating to genetic predisposition, diet, exercise, tobacco smoking, personality and behaviour, and hypertension itself. Suggestions that the incidence of single major stroke has fallen over a period during which the incidence of myocardial infarction has risen indicate that the relationship is not straightforward. As with DAT, life expectancy is undoubtedly reduced though the average size of the reduction is not known accurately.

Theoretical perspectives on MID would seem to be much more straightforward than those relating to DAT. It is widely recognised that some accumulation of arterial atheroma is ubiquitous beyond youth in people living in Western societies, but actual strokes seem relatively discrete and clearly abnormal events: it is tempting to assume that some people have strokes (and live long

enough to suffer an accumulation of several of them) and others do not. However there is evidence of cerebral infarction in approximately 50% of normal subjects over 65 who retain normal intellect, though the volume of infarcted tissue rarely exceeds 50 ml in these subjects (Wells, 1978; Tomlinson et al 1970). Presumably some sort of threshold effect is operating (as may be the case in DAT) whereby clearly apparent cognitive deficit only occurs once a certain level of brain damage is reached.

Methods of treatment are largely aimed at controlling hypertension, reducing blood viscosity, and altering lifestyle factors in the hope of preventing further infarcts. Drugs intended to improve mental functioning, such as cerebral vasodilators, have had limited success. Psychological and behavioural approaches to management will be similar to those appropriate for use with DAT subjects: nothing is known about the relative efficacy of psychological interventions in DAT as opposed to MID. One might assume that factors such as the higher incidence of depression in MID would be relevant, but this is only to say that clinically people should be considered individually according to their particular features.

Prevalence of the dementias.

Estimates of the prevalence of dementia vary considerably, partly because of sampling factors and differing criteria for defining the presence or severity of dementia as well as perhaps real differences in prevalence in different parts of the world. It is undoubtedly common, increasingly so with advancing age. Katzman (1976), in reviewing a number of epidemiological studies, states that the average estimate of the prevalence of dementia in people

over 65 in Europe is around 4% for severe dementia (defined in terms of incapacity for independent functioning) and around 11% for mild or moderate degrees of dementia. Within this age group, of course, estimated rates are higher in older groups than in younger ones. Henderson (1986), in a similar review of studies, gives figures of around 5% for those over 65 and around 15% for those over 80.

Epidemiological studies are often not able to distinguish between different categories of dementia. Jellinger (1976) reported the results of autopsy examinations of the brains of 1,009 patients who had suffered from dementia after the age of 55 years. 52.8% showed primarily DAT changes, 22.5% primarily vascular changes (i.e. MID), and 13.6% showed both types of abnormality. In the 12% found to have other types of pathology, the most common findings were cerebral tumour (4.7%), Pick's disease (2.3%), and hydrocephalus (1.3%). The figures are similar to those reported by Tomlinson et al (1970) in a series of 140 patients with dementia beginning after the age of 65.

It is often suggested that DAT is more common in women than in men, even when allowances are made for the greater longevity of women. Schneck et al (1982) suggest that there may still be some artefact, resulting from the use of hospital samples: cognitively impaired women are more likely to be under the care of hospital services than are cognitively impaired men for reasons concerning the likelihood of the availability of an able carer. MID, like other cardiovascular diseases, is more common in men.

Clinical diagnosis.

Renvoize et al (1985) thouroughly investigated 150 consecutive psychiatric hospital admissions of patients with dementia, and found that less than 5% had potentially reversible causes of the dementia. The remainder fell within the usual main categories. In a similar study Smith & Kiloh (1981) found that less than 4% of 200 consecutive admissions had a potentially treatable dementia. Gilleard (1984b) presents combined data from a number of studies similarly showing that, in typical series, dementias other than DAT and MID are rare. In a review article Schneck et al (1982) concluded that "the vast majority of dementia cases in the elderly are not due to treatable causes. Most of these cases are in the nosologic categories of Alzheimer's disease and multi-infarct dementia."

Renvoize et al (1985) mention other studies showing higher percentages of potentially treatable causes of dementia in patients seen in general medical services. A cautionary note is also sounded by Ron et al (1979) who found that the initial diagnosis had to be rejected in a third of 51 patients initially diagnosed, between 5 and 15 years earlier, as having presenile dementia. Nott & Fleminger (1975) similarly found that an initial diagnosis of presenile dementia often later proved to be questionable. Differences of detail in the various studies may partly reflect local patterns of referral.

Findings of unreliability are particularly associated with initial early diagnosis. It is worth pointing out here that all subjects in the DAT and MID groups in this study were considered to show dementia for at least six months before and at least six

months after the time of initial testing. Use of appropriate exclusion criteria in conjunction with scales such as the Hachinski index reduces the likelihood of mis- classification. Rosen et al (1980) have provided some pathological verification of the validity of Hachinski's (1975) ischaemic score in differentiating between types of dementia. Similarly Harrison et al (1979) found the scale to be successful in separating vascular dementia from other dementias. A major recent American working party (McKhann et al, 1984), while stressing that dementia is diagnosed on clinical and behavioural grounds, concluded that existing knowledge about DAT is insufficient to allow formulation of rigid and detailed operational criteria for diagnosing it. Subjects in this study are diagnosed according to commonly accepted clinical criteria, including the Hachinski index, described in the following chapter. No neuropathological data are available. Is is therefore possible that a proportion of the subjects in the studies to be presented were misdiagnosed, but this proportion is likely to be very small.

Staging of dementia.

Various attempts have been made to define identifiable stages or phases through which dementia progresses. This may have value as regards improving the accuracy of classification and labelling, particularly for research purposes. It is relevant to data to be presented later. The stage descriptions generally refer to DAT though it is not unreasonable to consider stages in MID in a similar way for practical purposes.

Schneck et al (1982) describe in a review article a simple three-stage classification of the kind commonly found in general accounts of dementia. The first stage is a 'forgetfulness' phase, in which

the cognitive deficit is largely subjective and may initially be indistinguishable from Kral's (1962) benign senescent forgetfulness; the second is a 'confusional' phase where the cognitive impairment becomes evident to others and interferes significantly with life; and the third is a 'dementia' phase where cognitive impairment is gross and other physical, behavioural, and social problems are often marked. Roth (1986) uses a similar three-stage model. Such classification, though simple and fairly uncontroversial, is rather imprecise and loosely defined. Others have attempted more specific stage definitions with at least some operational criteria for discriminating between stages.

Reisberg (1982) describe a 7-point scale for use in primary degenerative dementia, with stages ranging from 'no cognitive decline' (i.e. normal) to 'very severe cognitive decline'. The clinical features of the various stages are described along with associated characteristics of performance on psychometric tests. Discrimination between the early stages on testing requires relatively detailed assessment instruments (e.g. WAIS, Guild Memory Battery), but for later stages only a simple mental status questionnaire is said to be necessary. The authors report validation of their scale against independent psychometric, behavioural, and neuroradiological measures.

Hughes et al (1982) propose a 5-point scale, including 'normal' and 'questionable dementia' categories as well as three stages of 'definite dementia', defined according to cognitive performance and behaviour rating criteria. Minimal use is made of actual scores on cognitive tests, but good inter-rater reliability is reported. Hagberg & Ingvar (1976) describe a 5-stage model which is broadly

similar to other models described but which also takes into account the possible occurrence of separate patterns of focal impairments within the context of progressive overall reduction in cognitive or intellectual performance. In other words they allow for greater variation or heterogeneity within stages, making their approach perhaps more realistic though less neat. Models such as that of Reisberg et al (1982), perhaps for the sake of simplicity and clarity, appear to involve assumptions of greater uniformity or homogeneity within stages. Other staging strategies are reviewed by Levin & Peters (1982).

As with many studies concerning characteristics of the impairments seen in dementia, these attempts at staging are largely based on cross-sectional rather than longitudinal data, and are therefore necessarily tentative in view of the inherent problems in cross-sectional approaches to the study of ageing (Gilleard 1984b). The later, clearer definitions appeared in the literature after the studies to be described here were under way, too late for their operational criteria to be adopted (and relevant measures taken) on a prospective basis. Attempts are made to analyse patterns of cognitive performance according to stage of dementia defined according to criteria, based on level of performance on an orientation questionnaire, which are loosely similar to those of Reisberg et al (1982).

Other conditions involving cognitive impairment.

The original intention in the main study to be presented was to assess adequate-sized samples in a number of conditions (in addition to DAT and MID) where cognitive impairment of some sort is likely, and so to compare patterns of performance and decline in

the various groups. Conditions of particular interest were to include Huntington's chorea, Parkinson's disease, and multiple sclerosis (in each condition restricting study to patients retaining adequate motor abilities for valid testing). Study of depressed patients was also envisaged. These aims were not achieved, for practical reasons concerning time and sometimes also the availability of adequate numbers of suitable subjects. Once it became apparent that this would be the case, however, testing of subjects in this 'other' category continued since data from them could still fulfil useful functions.

Data from these subjects have been retained for three principal purposes, which are (1) to add to the DAT and MID subjects in the main study to provide good numbers of subjects upon which to carry out factor analysis to determine which tests (not which subjects) are sufficiently strongly associated to be lumped together to allow comprehensible pattern analyses. The number of tests used is large, so maximisation of subject numbers is desirable; (2) to assess the degree of variability in patterns of test performance between subjects in a mixture of diagnostic categories, against which the inter-individual variability within DAT and within MID can be compared; and (3) to provide some general control data from a mixed group of cognitively impaired subjects when looking at patterns of performance. This is of limited inherent interest in many respects, but can be of use in potentially identifying artefactual results or patterns - results which have arisen simply through the testing and statistical analysis of performance in any group of cognitively impaired subjects, rather than being peculiar to DAT or MID or to subgroups within those categories.

Diagnoses and exact subject numbers are listed in the next chapter. None of the Other subject groups is large enough to merit independent statistical analysis and no more than a few brief descriptive comments will be made about the separate groups, in due course. Detailed descriptions of the nature and characteristics of each of the conditions seem unnecessary in view of the uses to which the data will be put. Accounts of the aetiology, pathology, clinical features, and cognitive impairments found in the conditions, as far as these details are known, appear in texts and journals.

Many conditions other than those represented in the subject groups used in this thesis are known to be capable of causing cognitive impairment in elderly subjects. These include various primary CNS diseases, other physical illnesses and infections, nutritional and endocrine disorders, drug misuse, and psychiatric and psychological disorders. No further mention will be made of these, other than to note that efforts were made to exclude from study all subjects with possibly relevant conditions other than their principal diagnosed condition.

The absence of normal control subjects.

The studies presented in this thesis compare performance in different groups and subgroups of cognitively impaired subjects. No data concerning test performances in normal elderly groups are presented. There are a number of inter-related reasons for this. The first concerns tests and possible test artefacts. Pilot testing suggested that tests pitched at an appropriate level for the subjects of principal interest (i.e. those with confirmed dementia) were often too easy for 'normal' elderly subjects. This would not

be disastrous as regards straightforward description of the existence of deficits in dementia, but would cause serious problems in any pattern analyses of the kind to be attempted here. The inevitable ceiling effects on some tests in normals would produce distorted and misleading pictures of patterns of ability. This problem could not have been adequately overcome simply by extending upwards the range of each test's difficulty: with some tests this is impossible because of the nature of the test, while with others any such extension may result in the test measuring different things, i.e. different kinds of ability, at different ends of the range. (WAIS Arithmetic provides a crude example of this kind of phenomenon: early items assess straightforward calculation while late items assess more complex and abstract problem-solving ability.) Similar problems appear in much research comparing dementia sufferers with normals.

The second problem concerns practical aspects of the collection of an adequate sample of normal elderly people. Such people will not be found gathered together in hospitals or other care institutions or even in any of the various activity and social clubs organised for the essentially normal elderly. They will largely be found living in their own homes with little or no contact with organised services. The collection of a genuinely representative sample therefore requires a large (and enormously time-consuming) community survey, though even then the sample may be biased through aspects of self-selection such as readiness to participate. Such a project was completely beyond the resources of the present study.

The third and most fundamental problem in gathering normal

elderly control data in a study such as this concerns definitions of normality. In one strict if rather trivial sense of the term, of course, a sample of (for example) normal 90-year-old men cannot be assessed since such subjects have been dead for many years. The more practical approach of assessing living elderly people is however still valid, even though living 90-year-olds are a special survivor group who may always have had characteristics different from those of their deceased contemporaries, since the fact of survival can be assumed to be controlled for where cognitively impaired and normal subjects are age-matched. But the problem of defining normal for the purposes of the type of studies carried out here is insurmountable. Given that a proportion of old people, regardless of their location, is likely to be classifiable as showing dementia (as shown in the Newcastle studies), one has to select for a normal sample only those subjects who do not show signs of dementia. If this is not done, any comparison of normality and dementia becomes meaningless. Such selection can only be done on the basis of some external criterion involving assessment of subjects' cognitive or behavioural competence. An obvious candidate for such a purpose would be a cognitive test such as a brief mental status questionnaire: other types of assessment (such as behavioural ratings) could be used, but a priori one would expect any such measures to correlate with a cognitive test measure, i.e. to be contaminated by level of cognitive functioning. In the present context such considerations could only lead to tautological and inane conclusions that old people who show cognitive impairment do badly on tests and that mentally well-preserved people do well. Psychological test criteria can reasonably be used to help define

normal control populations in some kinds of study, but not in this one. As stated above, any apparently interesting differences between so-defined normal and dementing subject groups concerning patterns of performance on various different tests might also be attributable to artefacts arising from problems inherent in the nature of test construction.

For these various reasons, then, the studies which follow make no direct comparisons between characteristics of performance in cognitively impaired and normal subjects. Comparisons are generally confined to those between subjects with different types of dementia and between different sugroups of subjects (broken down according to age, severity of dementia, and so on) within these types.

Other methodological problems in age-related research.

Methodological problems arising in the study of ageing are reviewed by Bromley (1974) and Botwinick (1978). A number of important difficulties exist in addition to those concerning selection of normal control subjects and the accurate classification of abnormal groups, as discussed earlier. One concerns separation of actual results of ageing or maturation itself from cohort effects reulting from life experiences peculiar to different age groups. The limitations of cross-sectional studies in this respect are now well recognised. As mentioned above, Woods & Britton (1985, Ch 2) suggest that cohort differences are probably of greater significance than are age changes within groups, at least for the majority of the life span, in normal ageing.

Cohort influences can only be controlled for, and then perhaps only partially, in very long longitudinal studies requiring enormous resources (well beyond the scope of the present

investigation). A single cohort may still be subjected to particular experiences, events, or other environmental influences which lead to atypical patterns of decline and hence unrepresentative conclusions about the nature of normal or usual changes in the type of group being studied. Non-random drop-out of subjects can also be a major problem: more able or better preserved subjects tend to self-select themselves for continued repeated assessment, as shown by Siegler & Botwinick (1979). Non-random death is also relevant of course: especially in older groups, subjects available for study are survivors who may have (and who may always have had) special characteristics. Longitudinal studies may also suffer from complicating practice effects with retesting.

Baltes (1968) and Schaie (1967) suggest trying to overcome the limitations and difficulties of cross-sectional and longitudinal studies using cross-sequential designs, i.e. a group of cross-sectional studies with each age cohort re-assessed over time. In full form this too requires large resources, and is not free from limitations. Analysis of the results of such studies is complex: cohort effects and ageing effects can be inextricably confounded, and non-random drop-out remains a problem.

Liston (1979) comments on the problems of interpretation arising in studies where patients with varying durations of dementia and varying levels of impairment are investigated at only one point in the temporal progression of their illness. The studies presented in Chapters 2 and 3 involve some attempts to overcome such problems in that duration of illness is estimated wherever possible, level of impairment is staged using performance on a mental status questionnaire as an index, and a sample of subjects is reassessed

after an interval of ten months. Nevertheless, the studies fall well short of the ideal cross-sequential type of design described by Schaie.

A further methodological problem arising particularly in cross-sectional studies concerns the phenomena of terminal drop and terminal decline, and some comments on these are in order. The relationship between cognitive functioning and subsequent survival has been investigated in retrospective analyses of longitudinal studies, using time span between assessments and death as the main variable rather than time span between assessments and birth (i.e. chronological age). Some authors draw a distinction between terminal decline, a linear relationship between functioning and time before death, and terminal drop, a curvilinear relationship between the two involving accelerating decline with closeness to death. 'Critical loss', a specified rate of annual decline on certain tests, thus shows relationships with time until death. Siegler (1980), in a review of relevant evidence, concludes that most longitudinal studies do show a relationship between intellectual performance and subsequent survival, though this in itself does not prove the existence of terminal deterioration. Mixed results have appeared concerning the duration of the terminal deterioration phenomenon, with estimates ranging from a few months to several years, and even concerning whether it actually exists (Woods & Britton, 1985, p41). Clearly it is difficult to differentiate cognitive decline (as some isolated or self-contained phenomenon) from the influences of physical illnesses and disabilities common in old age, particularly in the period soon before death. Data from Savage et al (1973, reanalysed as described

by Woods & Britton, 1985, p41) suggest that terminal deterioration may begin around three years before death: Performance abilities deteriorate first, most noticeably in the two years before death, while striking loss of Verbal abilities occurs only very shortly before death. The available evidence principally concerns terminal deterioration in normal ageing. It is reasonable to suppose that some similar phenomenon might occur in people with dementia, though its measurement may be difficult in many cases because of the severity of impairment leading to floor effects on almost any conceivable kind of cognitive or behavioural measure.

The existence of any terminal deterioration phenomenon introduces bias into cross-sectional and even longitudinal group studies of age changes. The proportion of subjects in a terminal deterioration phase (i.e. close to death) is likely to be higher in older age groups, leading to an exaggerated estimate of the average decline in functioning with advancing age in normal groups and distortion of the patterns found in different age groups in dementia.

A final methodological note concerns sampling. The use of hospital- based samples introduces bias into data. Factors other than illness itself influence whether some people go to a particular hospital at a particular time. Social support factors, personality factors, age, sex, and the nature of the condition are all likely to play a role. The well known Newcastle studies showed that a proportion of elderly people living at home (and not in contact with specialised services) suffer from dementia which is objectively just as severe as that seen in hospitalised patients. The studies to be described used subjects from a range of sources

in an attempt to minimise the influence of selection biases, but the samples still cannot be considered ideally representative. Problems associated with collecting such a truly representative sample are great, and have much in common with the practical problems of collecting representative samples of normal elderly subjects, as described earlier.

Why assess the cognitively impaired elderly.

Despite a trend in British clinical psychology to make less use of formal tests, which has occurred for resons other than just pragmatic ones, various justifications exist for assessing elderly cognitively impaired subjects. Reviews are provided by Miller (1981b), Gilleard (1984), Woods & Britton (1985), and Albert (1981).

General themes of clinical relevance concern: the explanation of observed behavioural deficit; clear and reliable identification of a subject's cognitive strengths and weaknesses; contributions to management as regards appropriate placement, treatment, rehabilitation strategies, retraining, activities, environmental modification, medication, and so on; assessment of the effects of interventions, and ideally the prediction of response to interventions in individual cases; prediction of ability to cope in other situations or in the future; giving advice and explanation to the person and his or her family or carers; contributing to diagnosis (e.g. in attempting to detect early dementia or differentiate dementia from depression); and avoiding pejorative mis-labelling. Examples of other specific purposes are given by Kendrick (1982; and ensuing replies in same journal).

The ultimate aim is a clarity of definition of specific deficits

which will assist in management of the disorder (Woods & Britton, 1985); and Miller (1977, p144) notes that "There is a pressing need to understand more fully those processes that could underlie practical attempts to improve the functioning of demented patients." A better understanding of patterns of impairment may lead to improvements in assessment procedures for dementia, which can then be put to use in the various aspects of management.

Clearly many research endeavours warrant psychological testing of some sort. Many will relate, directly or indirectly, to the clinical aims listed above, though there is considerable scope for further basic research concerning the nature of the cognitive impairments seen in various categories of elderly people, with and without dementia. Currently, major research efforts are justifiably directed at physical aspects of dementia (particularly DAT) in studies of structural and chemical pathology and searches for possible causes. Dementia nevertheless remains by definition a state which is diagnosed on clinical, psychological grounds. Particular brain changes may be necessary conditions for dementia to exist, but they are not sufficient (and not all of them can be identified or measured in life in any case). Studies concentrating on physical aspects must rely on psychological assessment of some sort in identifying appropriate cases in the absence of reliable physical markers in vivo. Physical discoveries leading to treatments capable of arresting (rather than reversing) a dementing process will if anything make psychological approaches more important in dealing with the arrested sufferers: physical approaches capable of reversing, preventing, or eradicating any of the major dementias seem far in the future.

Of particular relevance to the present studies are papers by Moscovitch (1982) and Rosen & Mohs (1982), who have made attempts to relate patterns of neuropsychological impairment to particular areas of brain pathology in dementia. Moscovitch (1982) argues that at present the psychological consequences of pathological changes may be best understood by focussing on the larger systems (e.g. hippocampal functioning) rather than on the microstructure of those systems (e.g. the precise nature of the plaques or tangles within the hippocampus): the particular deficits shown by an individual will be determined by the brain structures that are most severely affected. Intuitively, more severe or more global dementia can be seen as resulting from more widespread pathological change, either structural (involving cortical atrophy or the density of plaques and tangles) or chemical (involving the extent to which different neurotransmitter systems are affected). Rosen & Mohs (1982) note that only relatively recently has systematic research using neuropsychological methods to evaluate decline in dementia been carried out. The rationale of the approach, they say, is that evaluation of performance on tasks which have been found to be sensitive to the effects of focal lesions reflects the functional status of cortical and subcortical regions of the brain: the functions of specific regions might or might not be compromised by the pathological changes characteristic of the dementia. Controversy continues over whether the neuropsychological impairments seen in dementia are or are not similar to those seen after focal lesions. The literature on language and memory disturbance suggests that they are not identical, and it is probably not reasonable to view the wide-ranging impairments of

dementia as simply a sum of the focal deficits identifiable in different patients with focal lesions. In dementia, areas of greatest brain pathology probably cannot be precisely defined on the basis of particular neuropsychological test deficits; but the correlational approach of looking for relationships between patterns of neuropsychological impairment and structural or chemical abnormalities (as estimated from pathological data obtained from the same or other similar groups of subjects) remains promising as regards describing the range of abnormalities encountered in dementia or defining subgoups. Rosen & Mohs (1982) specifically recommend neuropsychological comparisons of presenile and senile DAT.

No apology for conducting test-based studies therefore seems necessary. Wells (1982) argues that a greater research effort should be directed towards all clinical aspects of dementia, including affective, behavioural, and social components as well as the cognitive components which have typically received greatest attention. Symonds (1981) and Gilleard (1984b) view dementia in humanistic terms involving loss of the 'person' or the continuity of self, while acknowledging that the accumulation of cognitive impairments are bound to contribute to this loss. The studies presented in this thesis are largely confined to the area of cognitive impairment, for practical reasons concerning the time and resource limits of the study rather than any belief that these other components or perspectives are of minor importance.

CHAPTER 2

Patterns of Cognitive Impairment.

Issues addressed by the studies presented in this chapter and the next include: whether different patterns of neuropsychological impairment are present in DAT and MID; whether different patterns are present in subgroups within DAT and within MID; whether patterns of impairment are more 'focal' in nature in MID than in DAT (or in subgroups within these conditions); how great the variation is between subjects within defined groups and subgroups; which aspects of function show decline in different groups and subgroups over a 10 month period; to what extent the degree of decline over 10 months in different groups can be predicted from subjects' test performances or personal characteristics at initial assessment; and whether performances become less 'focal' in nature as the conditions progress, i.e. whether cognitive abilities exhibit de-differentiation with progression of dementia. The studies are essentially descriptive.

Tests.

Authorities on clinical assessment such as Lezak (1983), Albert (1981), and Lishman (1978) agree that an adequate picture of neuropsychological functioning should include some assessment of memory (verbal and non-verbal); language function (including comprehension, expression, repetition, naming, reading, and writing); visuospatial abilities (including perceptual and constructional abilities); so-called frontal lobe functions (involving planning, abstraction, and judgement); and general features such as concentration and mental speed (whether assessed objectively by specific tests or subjectively in terms of general performance on testing). Opinions differ as regards the importance of calculating general intellectual levels. Naturally they also

differ as regards the most appropriate tests to use in assessing the above abilities and as regards routine assessment of other specific abilities.

The choice of tests which could be used in a study such as this is enormous. Often, a choice has to be made between the merits of a test which has a long history of use (and hence which is more likely to have adequate normative data) and a test which seems better for the purpose in mind, but which has been used little with elderly populations. The latter of these two options has often been taken because this study is primarily intended to compare patterns of impairment in different impaired populations rather than comparing the performances of impaired groups against normative data.

A collection of tests was assembled on the basis of accounts in the literature and the author's previous experience of assessment in dementia. Piloting reduced this to the group of tests which will now be described. It is hoped that these satisfy a number of essential requirements for the intended study: that they (1) Give some idea of patients' levels of ability or deficit in the areas listed above; (2) are pitched at an appropriate level and cover an appropriate range of levels for the patient groups of interest, and are sensitive to impairment and to changes in level of functioning; (3) do not include requirements or general instructions likely to be incomprehensible to subjects; and (4) can be completed in a reasonable time without tiring or stressing subjects excessively.

General descriptions of the tests, and reasons or justifications for their inclusion where appropriate, will now be given. Detailed accounts of the procedures, verbatim instructions, and scoring

methods used are given in Appendix 1, and precise descriptions and illustrations of the materials appear in Appendix 2. The tests will be grouped under broad category headings for convenience, though this was not the order in which they were actually given; certain tests could just as easily be included under a different heading. For each test, appropriate introductory comments (not reproduced here) are made about the test and its requirements. Many tests include 'cut-off' criteria so that subjects do not needlessly receive challenges which are beyond them. In all tests the subject is encouraged to guess or 'have a go' where appropriate.

Tests of Memory and Learning.

1. ORIENTATION The subject is asked 13 questions covering orientation for person (name, age, date of birth); orientation for place (type of place currently in, its name, name of the city); orientation for time (weekday, month, year); and some other information (name of the present Prime Minister, the previous one, the U.S President; colours of the Union Jack), scoring 1 point for each correct answer. This is the Information/Orientation section of the Clifton Assessment Procedures for the Elderly (CAPE, Pattie & Gilleard, 1979) plus one question (on the name of the previous Prime Minister) in the hope of increasing sensitivity to impairment or change in subjects performing at a high level. It contains more appropriate items than some alternative mental status questionnaires.

2. PARAGRAPH RECALL A prose paragraph is read out to the subject, who must then say as much as he or she can remember about it, scoring 1 point for each idea correctly recalled. He is then asked a series of 6 questions about the paragraph, scoring (separately

from the first score) 1 point for each question answered correctly but 2 points for those questions already 'answered' in the spontaneous attempt. After a number of other tests (and on average about half an hour later) the subject is asked about the paragraph again, with the same 6 questions following the spontaneous recall attempt. With this in mind, during the initial questioning the subject is actually given the correct answers to the 6 questions if he does not know them. The paragraph used comes from Talland (1965), the 6 questions being made up around the main points of the story. This paragraph is used in preference to one of those in the Wechsler Memory Scale 'Logical Memory' subtest (which this test obviously resembles) simply because the author had used it satisfactorily in previous research with dementing patients, and also just in case any subjects might recently have been tested using the WMS (though this turned out not to be the case). The use of questions is to try to probe how much information the subject actually retains rather than how much he is able or willing to come out with spontaneously; these amounts are not necessarily similar. (This will be returned to in Chapter 6.) The scoring system based on the questions is a compromise giving some extra credit for spontaneous recall, and it will be used in the main analyses.

3. MEMORY FOR DESIGNS The subject is shown 4 designs and required to draw each one from memory immediately after a 10 sec. presentation, scoring points according to accuracy of reproduction. The designs are a diamond, a 4-pointed star, Card 1 (the 'flags') from the WMS Form A, and card 3 from the WMS Form B. The scoring system is a combination of the WMS criteria and made-up examples or 'models' of the various designs at different levels of accuracy of

reproduction, rather like those used in such tests as Graham & Kendall's (1960) Memory for Designs test. The test resembles The WMS 'Visual Reproduction' subtest. The diamond and star are included to extend the range of difficulty of the test downwards. Card 3 of W.M.S. form B is used rather than that of form A because it looks rather less daunting to a subject. Card 2 of the WMS is omitted simply to keep the test reasonably short. In retrospect, the designs chosen are perhaps not ideal since they are nearly all quite easy to characterise verbally.

4. YES-NO PICTURE MEMORY The subject is shown a series of 10 meaningful colour photographs for 4 sec. each. He is then shown another series which has the same 10 mixed up with another 10 and has to say which he saw before. The raw score is simply the number of 'hits' (i.e. correct 'yes' answers, pictures correctly identified as having been seen before) minus the number of 'false alarms' (i.e. wrong 'yes' answers, pictures incorrectly identified as having been seen before). A signal detection approach to the analysis of this material will be described in Chapter 6. The test was simply made up.

5. FACE-NAME LEARNING Four colour photographs of different persons' faces (2 male, 2 female) are shown in turn for 5 sec. each while the tester clearly says the person's name twice. The order of the faces is then rearranged; each is shown in turn and the subject is asked for the person's name. If the answer is wrong or not forthcoming, the tester gives the Christian name and asks for the surname. The subject is always told the full correct name before proceeding to the next face. The faces are then reordered and tested again in the same way, and then reordered and tested a third

time. A fourth, delayed, test is given after a number of other intervening tests, on average about 30 minutes later. For each face tested, the subject scores 3 for the full name correct spontaneously, 2 for the Christian name or surname correct spontaneously, and 1 for the surname in response to the Christian name. Again, this test was simply made up. It is obviously a kind of association-learning task, though the nature of the associations (between faces and names, and between halves of a name) are somewhat complex. It was chosen in preference to common paired-associate word learning tests because it has greater face validity (and perhaps 'ecological' validity) to an elderly person when prefaced with an inquiry about how he gets on remembering people's names. It is in fact a fairly hard test, and in retrospect 3 faces would probably have been quite enough.

6. DIGIT SPAN Strings of digits of increasing length (from 2 to 7 digits) are read at a rate of one digit per second to the subject, who then has to repeat them in the same order. If a string of a given length is failed, a second one of the same length is given. The score equals the longest string of digits correctly reproduced, plus 0.4 if this is achieved on the first trial at a given length. This is the 'Digit span forward' subtest of the Wechsler Adult Intelligence Scale. The small scoring bonus for not requiring a second attempt at the span is used to try to increase the sensitivity of the measure. Backward digit span is not used because in piloting it became clear that the instructions for it are understood with considerable difficulty, or not at all, by many people with dementia; in these circumstances it may become a test of comprehension rather than backward span, with no way of judging

when this is the case.

7. SENTENCE REPETITION The subject is required to repeat, one at a time, 6 sentences of increasing linguistic complexity, scoring 1 point for each word correctly repeated in the proper order. The sentences come from the Stanford-Binet Intelligence Scale, years IV, XI, and XIII. This can be seen, like digit span, as a test of immediate memory. Lezak (1983) comments on the relative neglect of useful memory-for-sentences tests in this context. Of course, performance might also be affected by deficits in comprehension, production, or conduction of language. It could therefore just as easily be classified under the next heading.

Tests of Language Function.

1. TOKEN TEST The subject is presented with an array of 20 wooden pieces - circles and squares, large and small, in five different colours - and asked to carry out instructions of increasing length and linguistic complexity (ranging from 'Touch a circle' to things like 'Touch the large white circle and the small green square' and 'Put the green square away from the yellow square'). In most parts of the test he is allowed a second chance on items failed the first time. 1 point is awarded for correct performance at the first attempt, 0.5 at the second. The Token Test is probably the most sensitive existing clinical test for comprehension deficit, and the one least contaminated by other aphasic or neuropsychological deficits. The version used here is the 36-item one given by De Renzi (1979). It is preferred to Spellacy and Spreen's (1969) even shorter (16-item) version because its level of difficulty extends further downwards, and also because it allows second attempts: this is an advantage in dementia where concentration problems are so

common. (The relationship between the 36-item version and shorter versions will be referred to in Chapter 7.) Adequate performance on the first few very simple items ensures adequate colour perception for the rest of the test (and for the Weigl sorting test given later). In case of colour agnosia or colour blindness a 'Supplementary Dyphasia Test' is available where a Spoon, Key, and Pencil are laid out before the subject and he is aked to carry out 12 instructions ranging from 'Touch the pencil' to 'Touch the key after you've touched the spoon'. This little test is cobbled together from a number of brief comprehension tests appearing in standard aphasia batteries such as that of Goodglass & Kaplan (1972), though it would of course be affected by visual object agnosia. In practice, very early failure on the Token Test did not occur. (Out of interest, this Supplementary Dysphasia Test was given to some patients who were not necessarily particularly poor on the Token Test. The relationship between performances on the two tests is considered in Chapter 7.) The supplementary test also includes two questions to do with Right-Left discrimination - 'Show me your left hand' and 'Touch your left ear with your right hand': these two items were given to all subjects.

2. SENTENCE PRODUCTION The subject is required to make up 6 sentences, each containing a given word. Scoring is 1 point for each attempt which can reasonably be considered a sentence. This comes from the Minnesota battery for the Differential Diagnosis of Aphasia (Schuell, 1965). It is chosen as a test of expressive ability which is not dependent on speed of performance; many tests of expressive ability, particularly word-fluency tests, are 'paced' tasks and so perhaps confuse expressive ability with psychological

speed (Miller & Hague, 1975). In addition, Kaszniak et al (1978) found that performance on this particular test showed a highly significant relationship to subsequent longevity. The test is by no means a comprehensive indicator of expressive ability: no single test can be. Perhaps the most valuable or valid assessment of expressive aphasia is based on observation of the patient's conversation and spontaneous speech. Some such ratings were in fact made, and will be returned to later.

3. WRITING AND READING The subject is asked to write his name and then, to dictation, two short sentences and some numbers. He is then asked to read out a sentence (the famous 'quick brown fox jumps over the lazy dog') and then copy it down. Each of these is scored according to the number of correct words or characters produced.

4. AUTOMATIC SPEECH The subject is asked to count aloud from 1 to 20, and then recite the alphabet, as quickly as possible without making mistakes. Each is scored in terms of the number of errors made, with possible time bonuses where 1 error or less is made. This is the 'Mental Control' part of the CAPE, with slightly amended scoring to try to increase the test's sensitivity. It is here called 'automatic speech' since that is the common term in aphasia batteries for such tests.

5. OBJECT RECOGNITION AND NAMING The subject is shown a series of 23 line drawings of common objects and asked to name each one. With those he cannot name he is asked to describe or show how the object is used, or what it is for, to distinguish whether the problem is one of naming or of recognition. The recognition score is simply the number of objects recognised; a naming score is derived (as

described later) according to the proportion of objects recognised which are also named (since someone can hardly be expected to name an object if he does not know what it is). Response latency in naming each object (up to a limit of 15 sec.) is also recorded. The 23 objects were selected and ordered according to the word frequency of their names in the Thorndike & Lorge (1944) word count. The subject is also asked to name seven parts of the body indicated by the tester on himself, scoring 1 point for each correct answer. These aspects will be returned to in Chapter 7.

Tests of Visuospatial and Visuoconstructive Abilities.

1. BLOCK DESIGN The subject is required to construct a series of 6 two-dimensional patterns, presented on cards, using thick but flat square wooden blocks coloured either black, yellow, or half black and half yellow (divided along the diagonal) on one side only. The designs used all require only 4 blocks, in a 2 X 2 format, and the subject is provided with only those 4 blocks required for a given design. Scoring depends on correct reproduction of the model design, with time bonuses for quick completion in the later designs. Time limits are generous: piloting suggested that subjects not finishing successfully within them never do so regardless of time allowed. This is a modification of the common type of block design task, Kohs' Blocks, found in the Wechsler Intelligence Scales. The modifications are considerable, and designed to make the test easier: piloting suggested that many dementing patients are unable to make any score on the standard type of block design task. There seem to be at least two main types of reason for this failure: firstly an inability to select the correct faces of the 6-faced cubes (or indeed a failure to even

consider turning them over, so that the subject vainly tries to construct a design with only the faces which the tester happens to leave upwards) and only secondly an inability to put the faces together in the correct spatial relationships. Both of these are inherent components of the classic block design tests but it was here decided to concentrate on the latter, both to simplify interpretation and to give subjects a sporting chance, by using single-faced blocks and providing only those needed to make a given design. The use of black and yellow colouring is simply to provide strong colours with maximum contrast. The design 'shapes' used came from the Wechsler intelligence scales for adults and children. One 3 X 3 design was also made up and given for clinical interest to some patients who had little difficulty with any of the 2 X 2 designs; this is ignored in the analysis since almost no- one ever completed it. Early in the study another construction task was used where the subject is given substantial wooden sticks and required to copy model patterns made from similar sticks. After a time it became apparent that the relationship between performance on this and on Block Design was so strong that one test was reduntant. The Stick test was dropped because its range of difficulty is narrower.

2. COPYING DESIGNS The subject is simply asked to copy the 4 paper-and-pencil designs he was previously asked to draw from memory. Scoring principles are exactly as are used for the memorial attempts. The subject is also asked to draw spontaneously a circle, square, and triangle, being given an appropriate shape to copy for those he does not manage. These are simply scored according to how good the shapes are.

3. Some miscellaneous visuospatial tasks:

(a) SQUARE WITH STICKS The subject is provided with 4 substantial wooden sticks and asked to make a square with them. If he fails he is asked to copy with the sticks a model made from similar ones. Scoring depends on how long he takes and whether he requires the model. This test was part of the stick test mentioned above and is retained simply because it is similar to MacDonald's (1969) one of making a square with matches. Larger sticks are used to compensate for arthritic or fine motor problems.

(b) MULTIPLE-CHOICE BLOCK DESIGN The subject is shown a series of 3 cards, each with a target design on the left and a group of 3 choice deigns on the right - one identical to the target, one a rotation of the target, and one different from the target. In each case he is required to indicate which is identical to the target, scoring 1 for each correct choice. The target designs are 3 of the design shapes used in BLOCK DESIGN. This is included simply to give some indication of whether failure on BLOCK DESIGN reflects a constructional problem or more of a visouperceptual one.

(c) YERKES TEST The subject is shown 3 perspective drawings of piles of blocks or boxes (containing 3, 6, and 10 blocks respectively, though some are 'hidden') and is asked to say how many blocks would be in each pile, scoring 1 for each correct answer.

(d) TIME-TELLING The subject is shown 3 drawings of clock faces (one with the face numbered and two without) and asked to say what time is showing on each, scoring 1 for each correct answer.

(e) VISUAL PERCEPTION WITH INTERFERENCE The subject is shown cards with one of the drawings from the OBJECT RECOGNITION AND

NAMING test covered with criss-cross lines and three words from the introduction to the NART (described below) similarly defaced. He scores 1 for each item correctly identified.

Tests of Abstraction and Planning.

1. WEIGL TEST The subject is presented with 12 randomly scattered wooden pieces - circles, squares, and triangles in four different colours - and asked to sort the pieces into groups. The idea is to sort them by colour or by shape, though this is not pointed out. If no adequate grouping is forthcoming, the tester puts the circles together and asks the subject to group the other pieces in the same way. If a grouping is managed (with or without the help of a partial grouping), the subject is asked to group them again, in a different way (i.e. the other way from the one he used first). If this is failed, the tester makes one appropriate grouping and asks the subject to group the rest similarly. After any successful grouping, the subject is asked to identify verbally the basis for his grouping. Scoring is based on the number of groupings achieved, whether partial groupings are needed, and whether the bases of correct groupings are verbally identified. This is the Weigl-Goldstein-Scheerer Colour-Form Sorting Test originated by Weigl (1941). It is a test of abstraction and change of mental 'set', particularly sensitive to frontal lobe deterioration. Shallice (1982) reports a study in which this Weigl test was found to be a better detector of frontal lobe damage than the Wisconsin Card Sorting test. It is also of course comparatively brief and unstressful. Time bonuses are included here to allow differentiation between performances at the better end of the range. Piloting suggested that the largest contributor to slow

performances was not the movement time but rather the time taken to decide what to do with or how to try to sort the pieces. Some subjects would fiddle around with pieces for some time until some idea for a sorting came to mind, or would make false starts; it seemed sensible for such subjects to receive a lower score than those who clearly realised what to do almost immediately.

2. PORTEUS MAZES The subject is presented with two short paper-and-pencil mazes and asked to draw a line through from start to finish without crossing the 'walls' or going up 'dead ends'. Scoring is based on the number of such errors made, with time bonuses for speed in mazes where no more than one error is made. This requires a certain amount of foresight and planning. The mazes used are Years 5 and 6 from Porteus (1965), with the little rats of the originals replaced by arrows. Aspects of Porteus' scoring system, such as counting lifting the pencil as an error, are ignored as not being relevant in the present context. The use of time bonuses is to differentiate between people who get through the maze with few or no errors but take a very long time to decide on their route and those who do so rapidly with obvious speed of planning and decision- making. This consideration was considered to override the possible criticism that the use of time bonuses in this test (or in any test not specifically designed to measure psychomotor speed and accuracy) leads to confusion of mental and motor speed.

Tests of Psychomotor Speed and Accuracy.

1. SPIRAL MAZE 1 The subject is presented with a plain spiral maze and asked to begin at the central arrow and draw round on the white track until he comes out at the end, going as quickly as

possible without touching the black 'walls' of the maze. The time score is simply the number of seconds taken to complete the maze, and errors are counted according to how often and for what distances the subject touches or penetrates the maze walls. This is adapted from the well-known Gibson Spiral Maze (Gibson, 1965), as used in the CAPE; the outermost circuit is removed, to shorten the test without (it is hoped) greatly reducing its validity, and all the small circular obstacles are removed.

2. SPIRAL MAZE 2 A similar maze is presented, but this time the circular obstacles remain. The requirement is just as in the first maze except that the subject is told to miss the spots or obstacles as well as not touching the walls. This is again the Gibson Spiral Maze, this time unadulterated apart from the removal of the outermost circuit. In both mazes, Gibson's original repeated time-stresses are omitted (as in the CAPE) to avoid making the test stressful.

3. BOX-FILLING A 10 X 10 grid of 1 cm boxes with a mark like a 1 in each of the first three is presented. After using the remainder of the top row for practice, the subject is required simply to fill the boxes with marks like 1's for 30 sec, working as fast as possible along the rows. The score is simply the number of boxes so filled in the 30 sec. This is included as an extremely simple psychomotor test. A similar, though slightly less simple, task appears in the General Aptitude Test Battery (United States Employment Service, 1965); the box-filling test can be considered to lie at the end of a continuum of psychomotor/processing speed tests ranging from tasks such as the WAIS Digit Symbol subtest down through Kendrick et al's (1979) Digit Copying test and the GATB

test of filling boxes with a mark like an inverted pi symbol.

Other Tests.

1. ARITHMETIC The subject is asked 8 arithmetical questions in a 'problem-style' format, in other words in the form of a questioning sentence or two where the subject has to perform a calculation to answer. He is then asked to perform 8 'plain' calculations presented in as basic a way as possible, such as 'What's 27 and 8'. Scoring is simply 1 point for each correct answer. The first set of questions are the first 8 items from the WAIS Arithmetic subtest. The second 8 were simply made up; they are included because of the known effect of immediate memory or receptive dysphasic deficits on performance in problem-style arithmetic tests (Lezak, 1983). No paper-and-pencil tests of calculation are included, which in retrospect is unfortunate since such tasks greatly reduce the short-term memory load.

2. ESTIMATED PREMORBID I.Q. The subject is presented with a series of words written in large lettering on cards, and asked to read them out. The words used are the 50 irregular words of the National Adult Reading Test (Nelson, 1982) preceded by 6 words from the Burt (1921) reading test which, in combination with the sentence read in WRITING AND READING, provides a check on whether the subject is too dyslexic for the test to be valid. For those subjects where testing appears valid, words from the Schonell (1942) Graded Word Reading Test are used where necessary (as instructed in the NART manual), and estimated premorbid I.Q. is calculated from the tables in the manual. Those subjects who appear to be dyslexic even with regular words, and in whom the NART is therefore invalid, can be given the W.A.I.S. Vocabulary scale:

however many people for whom the NART is invalid are sufficiently dysphasic for the WAIS Vocabulary measure also to be an invalid estimator of premorbid I.Q, and so such data are not used here. The order in which the tests are actually given is as follows:

1. ORIENTATION
2. AUTOMATIC SPEECH
3. DIGIT SPAN
4. BOX-FILLING
5. TOKEN TEST (and Supplementary Dysphasia Test if used)
6. WRITING AND READING
7. ESTIMATED PREMORBID I.Q.
8. SENTENCE REPETITION
9. SENTENCE PRODUCTION
10. PARAGRAPH RECALL
11. ARITHMETIC
12. YES-NO PICTURE MEMORY
13. FACE-NAME LEARNING
14. SPIRAL MAZE 1
15. SPIRAL MAZE 2
16. PORTEUS MAZES
17. DRAWING SHAPES
18. MEMORY FOR DESIGNS
19. COPYING DESIGNS
20. SQUARE WITH STICKS
21. MULTIPLE-CHOICE BLOCK DESIGN
22. BLOCK DESIGN
23. YERKES TEST
24. TIME-TELLING
25. PARAGRAPH RECALL (Delayed)
26. FACE-NAME LEARNING (Delayed)
27. WEIGL TEST
28. OBJECT RECOGNITION AND NAMING
29. VISUAL PERCEPTION WITH INTERFERENCE
30. NAMING BODY PARTS

Testing took anywhere between 40 and 90 minutes, depending on the person being tested, and in slower patients was split into two or even three sessions. (The split was normally made after the SENTENCE PRODUCTION test, to allow consistent presentation of the memory and delayed memory tests with the same intervening tests. Occasionally, OBJECT RECOGNITION AND NAMING, WEIGL TEST, and NAMING BODY PARTS was given after SENTENCE PRODUCTION to end the first session.)

Comment should be made on three kinds of neuropsychological test which may appear conspicuous by their absence in view of other published material. (1) Auditory agnosia was not assessed here because subtle hearing impairments, particularly high frequency ones, might interfere with a subject's ability to identify certain sounds. Only a small proportion of subjects had undergone full audiometry and the author had neither the time nor the expertise to undertake such testing personally. It would therefore not be possible to attribute poor performance to a genuine agnosia with confidence. (In contrast, brief clinical assessment of the ability to hear spoken speech is relatively simple, as is a rough assessment of visual acuity and visual field defects.) (2) Tactile agnosia was excluded because, to attribute test failure to genuine tactile agnosia, detailed testing of simple tactile sensation would have been necessary; again this did not seem practicable. (3) Motor apraxia was not routinely tested because the author had found in previous assessment of people with dementia that the scoring of a patient's attempts is difficult and subjective and that the tests may seem silly to subjects. Also, performance on motor apraxia tests is likely to be considerably hindered by arthritis. Later in the study, after the appearance of a more suitable and scorable apraxia test (Kertesz & Hooper, 1982), samples of subjects were assessed. This will be reported in Chapter 8.

A Short form of the above battery is given to subjects who either (a) become unwilling to complete the Full battery or (b) turn out to be so impaired that the administration of the Full battery would be both unkind and not particularly illuminating in terms of patterns of performance, since scores would so often be

zero. (Piloting suggested that subjects who did badly on the tests included in Short testing could manage very little of most of the other tests included in Full testing.) The decision to use only Short testing with certain subjects was made on an impressionistic basis and the criteria were not operationally defined.

The tests are (in order of presentation):

1. ORIENTATION
2. AUTOMATIC SPEECH
3. DIGIT SPAN
4. BOX-FILLING
5. Supplementary Dysphasia Test
6. Writing Name & Reading Sentence
7. SQUARE WITH STICKS
8. SENTENCE REPETITION
9. SPIRAL MAZE 1
10. SPIRAL MAZE 2
11. PORTEUS MAZES

These tests are identical to those of the same name in the Full battery. In the Supplementary Dysphasia Test the subject is also asked to identify the three objects (using a similar procedure to that used in OBJECT RECOGNITION AND NAMING). Part of WRITING AND READING is omitted, so that the subject only has to write his name and read the 'quick brown fox' sentence. The reason for using the Supplementary Dysphasia Test rather than the Token Test here as an indicator of comprehension deficit is that the former is shorter and so less taxing. In retrospect this is nonsensical, since many subjects would have failed the criterion for continuing the Token Test quite early and so the test would not have been long.

Ratings made of various aspects of behaviour during testing and conversation will be described in Chapter 4. All testing referred to in the thesis was conducted by the author.

Subjects.

The subjects tested came from a variety of sources in and around Edinburgh - three psychogeriatric day hospitals (one in a large psychiatric hospital, one in a rehabilitation hospital, and one in a geriatric hospital); in-patient psychogeriatric assessment facilities in a large psychiatric hospital; in- and out-patient facilities of neurology departments in two general hospitals; a medium-size local authority Part IV home for the elderly; and a few other in- and out-patient psychiatric facilities.

Diagnosis took into account Consultant's diagnosis, score on the Hachinski index, and results of appropriate laboratory tests (and special investigations such as CT scan where available). All DAT and MID subjects had at least a 6 month history of signs of dementia. It was hoped that this would exclude from these groups any subjects with an acute or treatable condition (i.e. not DAT or MID) which had simply not been discovered or diagnosed at an early stage. The records of all subjects were also reviewed at least 6 months after initial testing to check that no information had emerged which would require revision of the diagnosis. (These are minimum periods: in most subjects the history was longer and in most a check was made more than 6 months later.) Old diagnosis of 'senile dementia' by unidentified staff was if necessary over-ridden by the recent opinion of a Consultant and score on the Hachinski index because of the possibilities for confusion in those not familiar with recent developments in diagnostic terminology in dementia.

An attempt was made to avoid assessing patients where it seemed likely (from prior consideration of their clinical records) that

confident classification would not be possible, though a few people were tested before all appropriate investigations (e.g. neurological examination) were complete. Any subject who later turned out to have an intermediate Hachinski score, a mixed or uncertain diagnosis, or a different diagnosis (i.e. not DAT or MID) was placed in the Other group. (In this context it is worth mentioning that authors sometimes use an intermediate Hachinski score as an indication of mixed DAT and MID whereas it is really no more than an indication of uncertain diagnosis: a high Hachinski score indicates the presence of MID, but co-existing Alzheimer-type change is not excluded unless the onset was indeed fairly sudden and recent. Strictly speaking, then, the MID group here are really a group with <u>at least</u> MID.) This consideration of diagnostic features before testing is why the Other group contains so few mixed or uncertain diagnoses.

So, no patients were included in the DAT and MID groups where there was eventual doubt about the diagnosis. The possibility of mis-classification cannot be ruled out in the absence of post-mortem data; as discussed in the introduction, the proportion of subjects mis-classified is likely to be very small.

Subjects were excluded who had uncorrected impairments of sight or hearing: they were required to read typewritten letters and to repeat softly spoken speech. In rare cases where dysphasic or dyslexic problems led to errors with these methods, testing was prolonged until it seemed clear whether there was an additional sensory impairment or not. Also excluded were subjects who were too physically disabled (whether through arthritis, stroke in the MID group, or neurological disease in the Other group) for valid all-

round testing to be possible. Subjects with current or recent illness or infection such as pneumonia or urinary tract infection were also excluded because of the likelihood of such illness impairing mental functioning. Subjects with any history of alcohol abuse, psychiatric problems, ECT, or head injury were excluded from the DAT and MID groups. Subjects on psycho-active medication were not excluded since drug-free patients may represent a special sub-group of patients with dementia.

The patients who received full testing were:

1. DAT: 58 patients (18 Male, 40 Female; Mean age 75.3 years, SD 9.0 yrs, Range 53-91).
2. MID: 58 patients (33 M, 25 F; Mean age 69.5 years, SD 12.8, Range 37-92).
3. Other: 58 patients (30 M, 28 F; Mean age 55.6 years, SD 12.0, Range 24-83). The diagnoses were:

Suspected mixed DAT/MID (1), Suspected Pick's Disease (2), Huntington's Chorea (5), Senile Chorea (3), Parkinson's Disease (4), Steele-Richardson Syndrome (1), Multiple Sclerosis (4), Late Onset Demyelination (1), Hydrocephalus (3), Benign Intracranial Hypertension (1), Alcoholic Dementia (1), Mixed features of Korsakov's & Alcoholic Dementia (1), Dementia Pugilistica ('Punch-drunk' syndrome) (1), General Paresis of the Insane (infection 20 yrs previously) (1), Steroid-Induced Dementia (1), Binswanger Encephalopathy (1), Encephalitis Lethargica (1), Cerebellar Ataxia of unknown origin (1), Brain Tumour (4) (Respectively, Craniopharyngioma; Left frontal glioma; Right frontal glioma; Left temporal glioma), Metastatic spread to brain from extra-cranial malignancy (3), Single left fronto-parietal stroke (1), Left

parietal head injury (1), Frontal head injury with evidence of added dementing process (1), Left temporal lobectomy for chronic epilepsy (1), Depression (6), 'Functional' condition (Hysterical or pseudo-dementia) (2), Chronic schizophrenia (1), Mental Handicap (1), Normal with peripheral hand tremor (1), No diagnosis reached (3).

Roughly equal numbers of subjects in the three groups arose by accident rather than design; a few more subjects were then assessed to equalise the numbers for convenience.

Further small groups of patients completed the Short testing as described in the previous section. These comprised 16 DAT patients (4 M, 12 F, Mean age 81.7 yrs, Range 70-93), 16 MID patients (7 M, 9 F, Mean age 76.8 yrs, Range 48-88), and 9 patients with other diagnoses (4 M, 5 F, Mean age 76.6 yrs, Range 58-93): Suspected mixed DAT/MID (4), Alcoholic dementia (2), Single stroke (1), Mental handicap (1), No diagnosis reached (1). Adding these onto the groups who received Full testing (which of course incorporates Short), we have the following groups:

1. 74 DAT patients (22 M, 52 F; Mean age 76.7 yrs, SD 9.1, Range 53--93).
2. 74 MID patients (40 M, 34 F; Mean age 71.1 yrs, SD 12.5, Range 37-92).
3. 67 patients with Other diagnoses (34 M, 33 F; Mean age 58.4 yrs, SD 13.8, Range 24-93).

Most subjects receiving only Short testing did so because they seemed too impaired to cope with Full. Almost none were unwilling to complete testing once it had been started. The numbers of people who refused outright to attempt either version of testing were also

extremely small and not confined to particular diagnostic or age groups.

Subject characteristics relevant to analyses to be presented later can be summarised in a series of tables. Tables 1 to 3 show numbers of subjects receiving Full testing broken down according to age group, sex, location, and drug use. Tables 4 and 5 repeat this for subjects completing at least Short testing (omitting data on location since the distribution of locations was so similar to that in subjects receiving Full testing).

Table 1 Numbers of patients who received Full testing, broken down by diagnostic category, age group, and sex (Male, Female in parentheses).

	DAT	MID	Other
-69	14 (7, 7)	24 (18, 6)	53 (28,25)
70-79	23 (6,17)	18 (10, 8)	4 (1, 3)
80+	21 (5,16)	16 (5,11)	1 (1, 0)
Total	58 (18,40)	58 (33,25)	58 (30,28)

Table 2 Numbers of patients who received Full testing, broken down by location, diagnostic category, and age group.

		PsDay	PsIn	NeIn	NeOut	LAiv	Other
DAT	-69	2	0	7	1	1	3
	70-79	14	4	1	0	4	0
	80+	12	3	0	0	6	0
tot		(28)	(7)	(8)	(1)	(11)	(3)
MID	-69	2	0	16	5	1	0
	70-79	13	2	1	0	2	0
	80+	10	1	1	0	4	0
tot		(25)	(3)	(18)	(5)	(7)	(0)
Other	-69	2	4	34	7	2	4
	70-79	4	0	0	0	0	0
	80+	0	0	0	0	1	0
tot		(6)	(4)	(34)	(7)	(3)	(4)
Total		59	14	60	13	21	7

PsDay-Psychogeriatric Day Patient, PsIn-Psychogeriatric In Patient, NeIn-Neurology In Patient, NeOut-Neurology Out Patient, LAiv-Local Authority Part IV Resident, Other-Other category.

Table 3 Numbers of patients who received Full testing, broken down by drug use, diagnostic category, and age group.

		Hypn	MinT	MajT	ADep	Two	Other	tot	None
DAT	-69	2	0	1	2	0	0	(5)	9
	70-79	1	1	5	0	0	0	(7)	16
	80+	2	0	3	0	0	0	(5)	16
tot		(5)	(1)	(9)	(2)	(0)	(0)	(17)	(41)
MID	-69	2	1	1	0	2	2	(8)	16
	70-79	1	0	5	2	0	2	(10)	8
	80+	1	0	1	1	1	0	(4)	12
tot		(4)	(1)	(7)	(3)	(3)	(4)	(22)	(36)
Other	-69	1	2	6	3	2	11	(25)	28
	70-79	0	0	0	2	0	1	(3)	1
	80+	0	0	0	0	0	0	(0)	1
tot		(1)	(2)	(6)	(5)	(2)	(12)	(28)	(30)
Total		10	4	22	10	5	16	67	107

Hypn-Hypnotic, MinT-Minor Tranquilliser, MajT-Major Tranquilliser, Adep-Antidepressant, Two-Two of the preceding categories, Other-Other psychoactive medication.

Table 4 Numbers of patients who received Short testing, broken down by diagnostic category, age group, and sex (Male, Female in parentheses).

	DAT	MID	Other
-69	14 (7, 7)	26 (19, 7)	55 (29,26)
70-79	31 (9,22)	24 (12,12)	7 (4, 3)
80+	29 (6,23)	24 (9,15)	5 (1, 4)
Total	74 (22,52)	74 (40,34)	67 (34,33)

Table 5 Numbers of patients who received Short testing, broken down by drug use, diagnostic category, and age group.

		Hypn	MinT	MajT	ADep	Two	Other	tot	None
DAT	-69	2	0	1	2	0	0	(5)	9
	70-79	2	1	8	0	1	0	(12)	19
	80+	2	0	4	0	0	1	(7)	22
tot		(6)	(1)	(13)	(2)	(1)	(1)	(24)	(50)
MID	-69	2	2	1	0	2	3	(10)	16
	70-79	1	0	6	3	0	3	(13)	11
	80+	2	0	3	1	1	0	(7)	17
tot		(5)	(2)	(10)	(4)	(3)	(6)	(30)	(44)
Other	-69	1	2	6	3	2	12	(26)	29
	70-79	0	0	1	2	0	2	(5)	2
	80+	1	0	1	0	0	0	(2)	3
tot		(2)	(2)	(8)	(5)	(2)	(14)	(33)	(34)
Total		13	5	31	11	6	21	87	128

Hypn-Hypnotic, MinT-Minor Tranquilliser, MajT-Major Tranquilliser, Adep-Antidepressant, Two-Two of the preceding categories, Other-Other psychoactive medication.

Some other recorded subject characteristics can be summarised without the tabulation of figures. Data on these variables were analysed broken down by diagnostic group and by age groups within diagnostic groups. The comments apply to both subjects completing Full testing and those completing only Short. (1) Marital status: There were no remarkable trends here; subjects in the Other group were more likely to be divorced or separated than were subjects in the DAT and MID group. Generally, older subjects were not

surprisingly more likely to be widowed than were younger ones. (2) Social class: There were no notable group differences of any description here. (3) Education: Here again there were no differences of note; in fact few subjects in any group had received formal education past the minimum school leaving age. (4) Scores on the Hachinski index: Here of course there were marked differences between diagnostic groups, but no differences between age groups within each diagnostic group. There were also no differences between groups within the main diagnostic groups when broken down by Orientation score or by overall Full score (where appropriate). There is therefore no reason to suppose that any differences found between groups within diagnoses might be attributable to less 'clear-cut' diagnoses in one group compared to another.

Results: patterns of neuropsychological performance.

Performance on a collection of tests such as that described above could be analysed in many ways. Initially, to provide a broad picture of patterns of performance, a factor analytic approach was used. To make the analysis more manageable, certain test scores were first grouped together. This was done on a priori grounds - that the tests seemed to measure related abilities - though in all cases the statistical relationships between the lumped variables were sufficiently strong for such combination to be reasonable. Scores lumped together were:

(1) The two halves of AUTOMATIC SPEECH.

(2) All four parts of WRITING AND READING.

(3) Immediate and delayed parts of PARAGRAPH RECALL.

(4) The two halves of ARITHMETIC.

(5) All four trials on FACE-NAME LEARNING (1,2,3, & delayed).

(6) The time and error scores on SPIRAL MAZE 1, using the formula Compound score = time in sec. x Log.(no. of errors + 10). (Hence a person making no errors scores simply the number of seconds he took to complete the maze.)

(7) The time and error scores on SPIRAL MAZE 2, in exactly the same way as above.

(8) SQUARE WITH STICKS, MULTIPLE-CHOICE BLOCK DESIGN, YERKES TEST, TIME-TELLING, and VISUAL PERCEPTION WITH INTERFERENCE. This combination will be known as 'MISC. VISUO. TASKS'.

(9) OBJECT RECOGNITION AND NAMING using the formulae

Recognition score = no. of pictured objects recognised / 23

Naming score = (no. of objects + no. of body parts named) / (no. of objects recognised + 7)

(The use of this last '7' rather than 'no. of body parts recognised' was because no subject was seen who actually proved unable to identify on himself which body parts the examiner was indicating; deliberate testing of this was therefore discontinued about half way through the study.) These derived scores are referred to as OBJECT RECOGNITION and NOMINAL ABILITY respectively.

This leaves 22 main scores, not counting ESTIMATED PREMORBID I.Q. (which will be returned to later). Each of these 22 scores was converted to a percentage score: this was to avoid excessive contributions to 'total score' coming from those tests which happen to have a large numerical range of scores. For most tests, this percentage score was simply calculated by dividing the obtained score by the maximum possible score on that test or combination of tests (i.e. dividing the score on ORIENTATION by 13, the TOKEN TEST

by 36, and so on). In a few, slightly different procedures were necessary: In SPIRAL MAZE 1, the Compound score was divided into 20, since the compound score is higher with poorer performance and a score of 20 (i.e. no errors in 20 sec) is approximately the best score anyone achieved. In SPIRAL MAZE 2, the Compound score was divided into 40, for similar reasons. Scores on OBJECT RECOGNITION and NOMINAL ABILITY as calculated above are already in percentage form. Overall 'Full score' is the mean of these 22 percentage measures. Scores on Short testing were treated in the same way, and an overall 'Short score' similarly derived.

A principal components factor analysis (type PA1 in the S.P.S.S. system) without iteration was carried out on the test scores of all the 174 subjects who completed Full testing. The use of all 174 subjects at this stage was to give a general idea of the relationships between tests and to provide a consistent framework within which to compare different sub-groups of patients. With 22 variables, it also provides a respectable number of cases with which to use a factor analytic procedure. The analysis produced four factors. Loadings (after varimax rotation) of each test on each of the factors are shown in Table 6. (Inter-test correlations for all the 174 patients, as well as for each of the 3 main groups of patients, appear in Appendix 3.)

As the table shows, the highest loadings on Factor 1 are for ORIENTATION, PARAGRAPH RECALL, MEMORY FOR DESIGNS, YES-NO PICTURE MEMORY, FACE-NAME LEARNING, PORTEUS MAZES, BLOCK DESIGN, and the WEIGL TEST. It seems to represent a combination of memory (other than immediate memory) and performance on tests particularly sensitive to frontal lobe dysfunction (i.e. PORTEUS MAZES, WEIGL,

Table 6 Varimax Rotated Factor Matrix of analysis of 22 test scores in all 174 subjects who completed Full testing. (Figures appropriately rounded and loadings higher than .55 underlined.)

	Factor 1	Factor 2	Factor 3	Factor 4
ORIENTATION	.717	.293	.132	.444
PARAGRAPH RECALL	.754	.119	.217	.382
MEMORY FOR DESIGNS	.747	.415	.278	-.023
YES-NO PICTURE MEMORY	.704	.224	.121	.160
FACE-NAME LEARNING	.792	.054	.174	.304
PORTEUS MAZES	.618	.349	.447	.027
BLOCK DESIGN	.607	.475	.429	-.039
WEIGL TEST	.599	.499	.320	.019
DIGIT SPAN	.172	.732	.108	.065
SENTENCE REPETITION	.338	.702	.052	.383
TOKEN TEST	.391	.719	.136	.329
SENTENCE PRODUCTION	.250	.622	.141	.366
WRITING AND READING	.037	.556	.324	.222
AUTOMATIC SPEECH	.090	.684	.261	.333
COPYING DESIGNS	.549	.627	.247	-.058
MISC. VISUO. TASKS	.517	.621	.325	.238
ARITHMETIC	.503	.666	.139	.204
SPIRAL MAZE 1	.188	.216	.879	.060
SPIRAL MAZE 2	.265	.139	.891	.046
BOX-FILLING	.257	.198	.801	.103
OBJECT RECOGNITION	.318	.246	.123	.700
NOMINAL ABILITY	.116	.373	-.017	.783
Percentage of the variance explained	24.3	23.4	15.0	10.2
				(Total 73.0)

and BLOCK DESIGN. This last test is generally more affected by parietal dysfunction, and it may be that the use of a system of time credits has edged it into this factor rather than the second one). It can reasonably be thought of as a memory and 'frontal lobe' factor, and will be referred to as 'memory'.

Factor 2 has highest loadings for DIGIT SPAN, SENTENCE REPETITION, TOKEN TEST, SENTENCE PRODUCTION, WRITING AND READING, AUTOMATIC SPEECH, COPYING DESIGNS, MISC VISUO TASKS, and ARITHMETIC. It seems to represent a combination of language function and performance on tests considered sensitive to parietal or parieto-occipital dysfunction. It can be called, roughly

speaking, a 'parietal' factor.

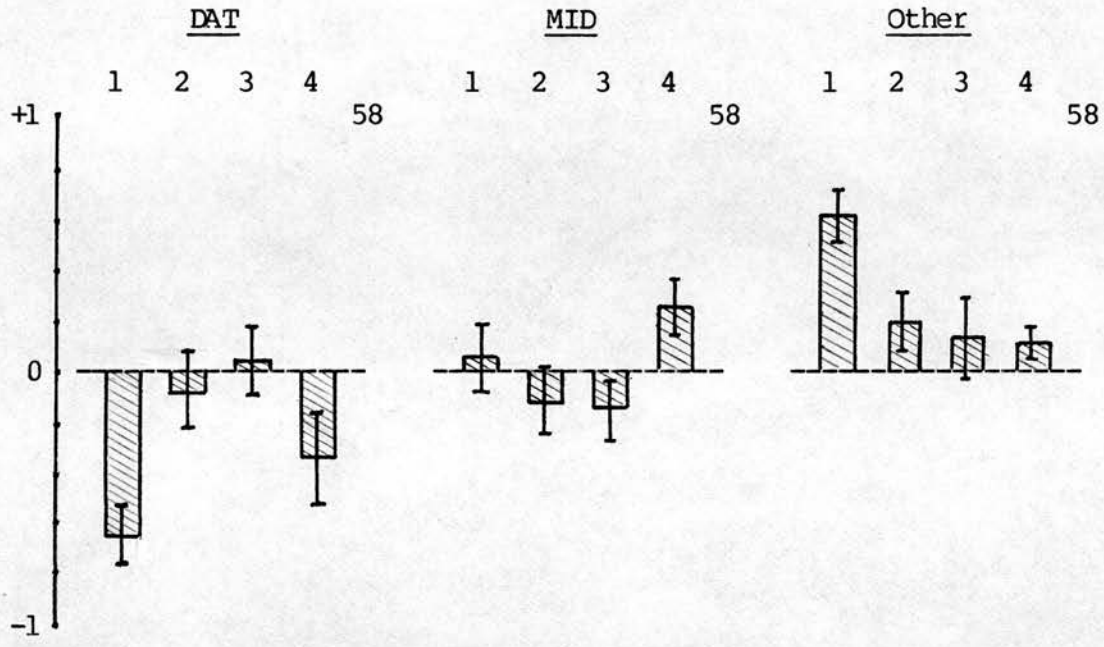
Factor 3 has high loadings only for SPIRAL MAZE 1 and 2 and for BOX-FILLING, and is a clear psychomotor speed and accuracy factor.

Factor 4 has high loadings only for OBJECT RECOGNITION and NOMINAL ABILITY. It is perhaps a little worrying that these variables come out in isolation on the same factor, since neuropsychologically speaking there is no reason to expect that visual agnosia and nominal dysphasia should be particularly strongly associated with each other; this raises the suspicion that the method of distinguishing nominal from recognition failure was not as successful as the author believed it to be. This will be considered further in Chapter 7.

Factor scores were then calculated for each subject on each factor, according to the formula given in the manual of the Statistical Package for the Social Sciences, i.e. scores for how well each person did on each of these four factors or aspects of test performance. Mean factor scores in various groupings of patients (grouped according to diagnosis, age, level of impairment, and so on) will now be considered.

The mean factor scores of the three main patient divisions (DAT, MID, Other) are shown in Figure 1. This does seem to indicate a different pattern of performance in the DAT and MID groups overall (probably not simply explicable, as subsequent figures will suggest, in terms of the MID group being rather better than the DAT group). In the DAT group Factor 1 (Memory etc.) and Factor 4 (Recognition and nominal ability) are particularly poor compared to the other two factors (Parietal and Psychomotor respectively). In the MID group the pattern is virtually reversed, with performance

Figure 1 Mean factor scores of patients who completed full testing, broken down by diagnostic category (n of each group at top right).

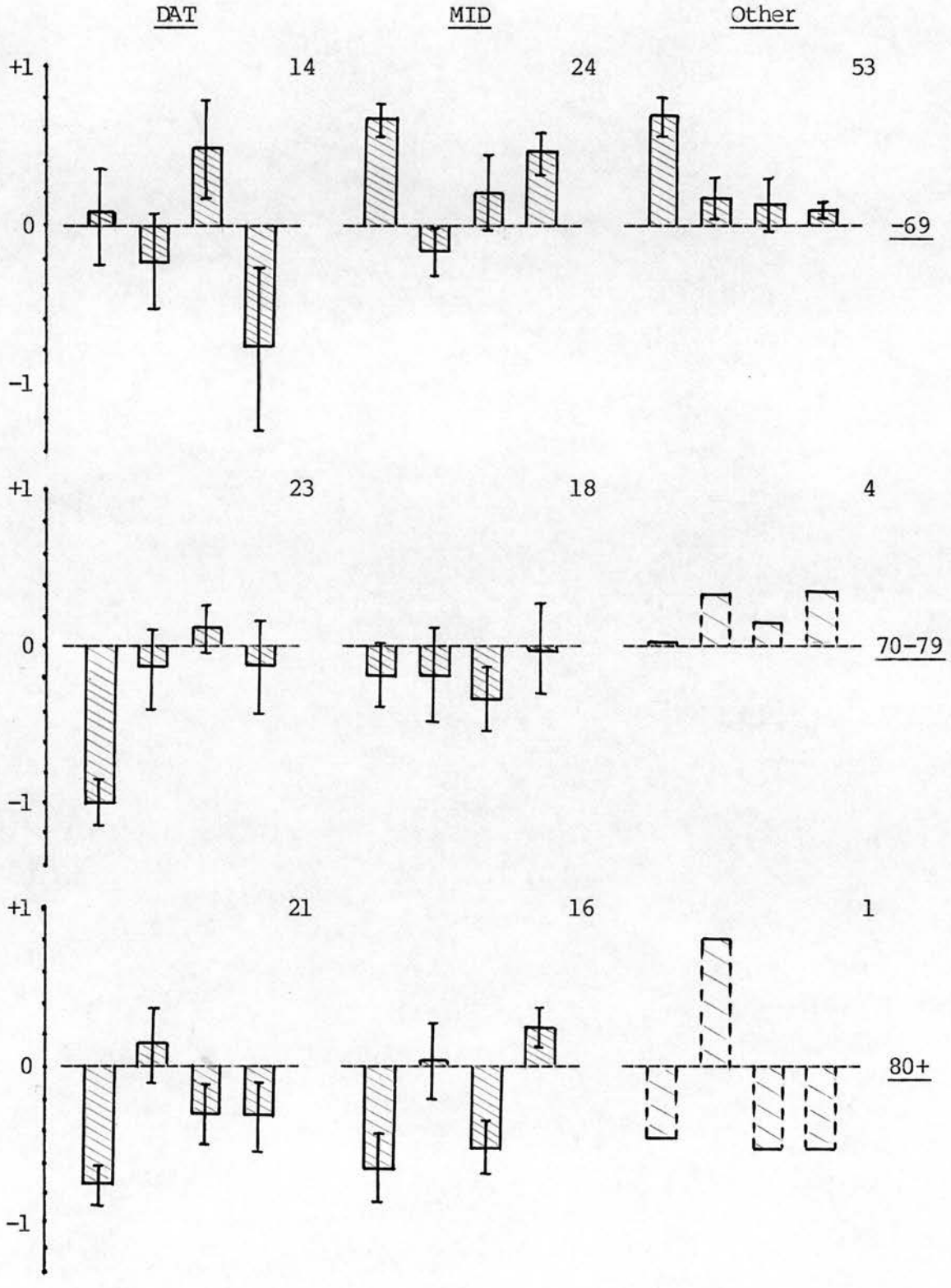


Factors are arranged in numerical order (from 1 to 4) from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

on Factor 4 and Factor 1 being rather better than on the other two factors. In other words, poor memory and visual agnosia/nominal dysphasia, relative to other factors (and in absolute terms here as it happens), are more likely in DAT than in MID. No claims are being made that such average differences in patterns would have diagnostic power in individual cases. The pattern of scores in the Other group is relatively uninteresting except for the fact that poor memory (i.e. Factor 1) performance (compared to other factors) is not a characteristic feature. This Other group is of course at a higher overall level than the other two groups. An ANOVA on Full scores showed a significant difference between the three groups, with the Other group having the highest level, the MID group the next highest, and the DAT group the lowest. (ANOVAs on each of the four factors alone, broken down by diagnostic group, showed significant differences for Factors 1 and 4 but not for Factors 2 and 3.)

Figure 2 shows the three groups broken down by age at first testing: under 69 at the top, 70 to 79 in the middle, and over 80 at the bottom. Age at first testing is used simply because a reliable dating of onset is not generally possible in DAT, which is by definition a condition with an insidious onset. The DAT groups will be considered first. The most interesting features of Figure 2 concern the relative levels of performance on Factor 1 (Memory) compared to Factor 2 (Parietal). In the under-69 DAT group, Parietal performance is worse than Memory performance. In the two older DAT groups, however, Parietal scores are considerably higher than Memory scores. In fact the mean level of Parietal performance is about the same in each group; but in the younger group this

Figure 2 Mean factor scores of patients who completed full testing, broken down by diagnostic category and age group (n of each group at top right).



Factors 1 to 4 are arranged from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

level of Parietal impairment has been reached while the Memory impairment is still relatively mild; in the older groups the same level of Parietal impairment is seen in people with much worse Memory impairment. Younger DAT patients do seem to show more 'parietal' symptoms, relatively speaking, than do older ones. As regards Factor 3 (Psychomotor speed and accuracy), performance worsens with age; this is no great surprise, though it is interesting that such a psychomotor slowing with advancing age continues to show through despite the presence of dementia. Performance on Factor 4 (Recognition and nominal ability) again suggests particular impairment (relative to the other factors) in the young group: an ANOVA on this factor broken down by age group produced a non- significant result owing to the large degree of variability in the youngest group. (Similar ANOVAs produced significant results for Factors 1 and 3 but not for 2.) An ANOVA on Full score broken down by age group showed no significant difference between the three DAT groups: in other words age is not confounded by level.

In the MID groups, the pattern of performance on Factor 1 compared to Factor 2 is quite similar to that seen in the DAT groups, with Parietal performance at about the same level in each group but with Memory performance being worse in the older groups. On Factor 3, again, the older groups are unsurprisingly poorer. Unlike in the DAT groups, however, there is no sign of the youngest MID subjects being particularly poor on Factor 4. As in the DAT groups, ANOVAs on factor scores broken down by age group indicated significant differences in Factors 1 and 3 but not 2 and 4. An ANOVA on Full score broken down by age group gave a significant

result, indicating that age and overall level of performance are confounded in the MID subjects. More caution is therefore required in interpreting the patterns. Numbers of subjects are too small to allow analysis with simultaneous breakdown by age and level of performance. It is interesting to note that the youngest and oldest DAT and MID groups conform to the impression gained from Figure 1 that object agnosia/nominal dysphasia (Factor 4) is more likely in DAT, but no such trend appears in the middle age groups.

The Other group cannot be broken down by age in a meaningful way because of the small numbers of older subjects. (In this and following figures, subject groups where the number of cases is very small are shown by broken bars: they are included for the sake of completeness.)

The scores on the parietal factor are at about the same level in all the subject groups shown, while the levels of the other factors can be quite different in the different groups. One possible reason for this might be that the tests contributing to this parietal factor are simply insensitive, or crude, or pitched at the wrong level. Data will be presented in the next chapter suggesting that this is not the case.

The fact that overall Full score does not differ significantly by ANOVA between the three DAT age groups (along with the breakdowns by Full score and Orientation score to be presented in Figures 3 and 4) indicates that the different age-related patterns in DAT cannot be explained in terms of differing levels of severity of dementia.

Another relevant point concerns the ageing process itself. Unless one thinks of DAT purely in terms of accelerated ageing, the

cognitive changes found must be a combination of those attributable to ageing itself and those attributable to the DAT. The older groups would have more of the changes attributable purely to ageing, and so it might be the case that the different patterns found reflect the superimposition of identical DAT-related impairments and different age-related impairments. This cannot be assessed on the basis of the data available, but age-related changes would have to be large to entirely account for the findings: as noted in the introduction, this is not generally thought to be true. Specific aspects of the patterns also do not fit well with such a hypothesis: the trend for Factor 1 to decline with age itself might be expected, but one would also expect a comparable decline in Factor 2 with age since it includes a number of new, unfamiliar tasks requiring 'fluid' intelligence to perform. The findings as regards Factor 4 are also very much against such an explanation.

Another possibility is that the different patterns simply reflect a cohort effect. Again this cannot be assessed on the basis of the data available, but any cohort effect would have to be large to explain the findings completely; and, again, specific aspects of the patterns do not fit well with such a hypothesis.

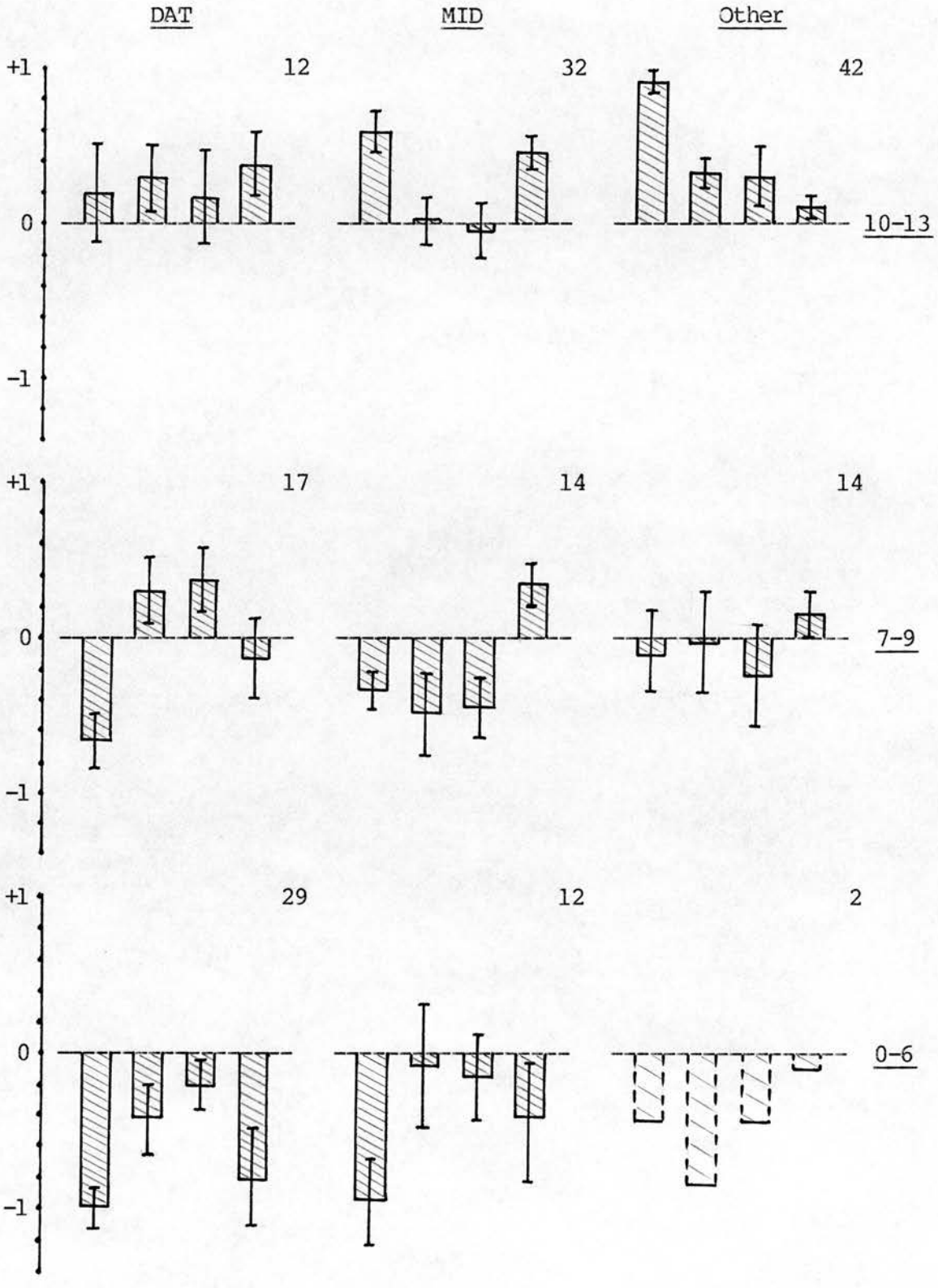
Still another possible influence concerns sampling. An effort was made to cover a reasonably broad range of subject sources, to avoid the possibility of having an over-representation of especially 'interesting' early or young cases of dementia (such as those with focal features, who might more easily find their way into, say, neurological clinics than patients showing only memory problems or a more global deterioration). However a further bias

remains possible concerning whether potential subjects come into contact with any outside agency or services whatsoever: younger subjects are more likely than older ones to have a spouse to look after them and hence may have to be more severely impaired or disabled if the help of outside agencies is to be sought. It is conceivable that parietal or Factor 2 impairments are in some way more disabling or disruptive of everyday life than are memory or Factor 1 impairments, leading to a finding of disproportionate parietal impairment in those young subjects who are actually seen.

The fact that roughly similar age-related patterns occur (at least with regard to Factors 1 and 2) in the MID subjects does suggest that some of the above factors may be relevant to some extent in DAT. (It does not seem plausible that many MID subjects had coexisting DAT.) Nevertheless, the most plausible hypothesis is that at least a large part of the age-related differences in patterns in DAT is related to the differing neuropathological and neurochemical characteristics of early-onset and late-onset DAT subjects as described in the introduction.

Comments on the importance of staging of severity of dementia appeared in the introduction. This is attempted here using ORIENTATION performance and then overall test score as indices. Figure 3 shows the diagnostic groups broken down by stage of dementia according to performance on ORIENTATION: 10 to 13 at the top, 7 to 9 in the middle, and 6 or less at the bottom. In each diagnostic group, performance on Factor 1 falls from the top graph to the bottom one, as one might expect since ORIENTATION contibutes most strongly to this factor. In the DAT group, Factor 4 also shows a more or less orderly decline from top to bottom

Figure 3 Mean factor scores of patients who completed full testing, broken down by diagnostic category and ORIENTATION score (n of each group at top right).



Factors 1 to 4 are arranged from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

graphs. Factors 2 and 3 however only drop once the poorest of the three levels of orientation is reached. One might therefore conclude, for example, that Factor 3 (psychomotor speed and accuracy) shows a clearer pattern of decline with age than with stage of dementia as defined by ORIENTATION level. ANOVAs on each factor broken down by level all produced significant results. In the MID group, Factor 4 again shows a fairly orderly decline from top to bottom. Factors 2 and 3 are not significantly different (by ANOVA) in the three groups, though there is a tendency for the mid-stage group to be worse than the lowest one: it is not clear why this should be so. In the Other group, Factors 2 and 3 are lower in the middle group than in the top one, but neither of the differences are significant. Factor 4 is no lower.

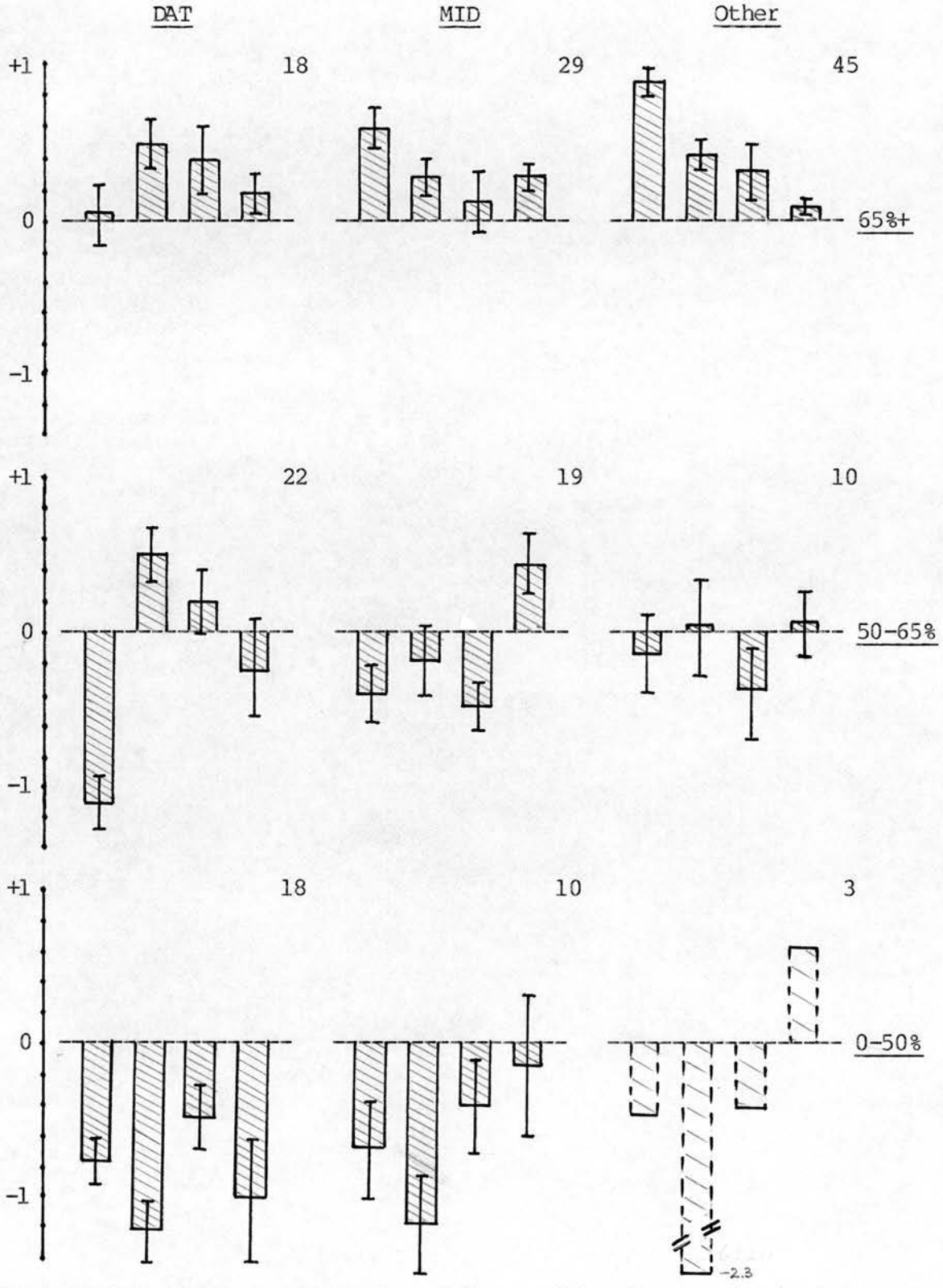
The only inter-diagnostic differences of note occur in the middle DAT and MID groups. In the DAT group Factors 2 and 3 are high while 4 is low; in the MID group Factors 2 and 3 are low while 4 is high. Perhaps parietal and psychomotor abilities 'go' relatively early in MID while recognition and nominal ability are well preserved until a late stage; whereas in DAT parietal and psychomotor abilities are better preserved while recognition/nominal ability 'goes' early.

As might be expected from the comments about Figure 2 concerning Full scores in the groups broken down by age: Age is not significantly different in the three DAT groups when broken down by ORIENTATION score, but is significantly different in the three MID groups (with a mean of 63 years in the 10-13 group, 76 in the 7-9 group, and 81 in the 0-6 group). Again, the confounding of age with level (as defined by ORIENTATION score) in the MID groups invites

caution in interpreting the patterns. In the Other groups, the poorer groups are also older than the best group, though this is less important. As previously noted, the numbers of subjects in the diagnostic categories are too small to allow a breakdown by age and level simultaneously. Not surprisingly, of course, Full scores are significantly different in the three groups within each diagnosis.

Performance on ORIENTATION may not be the best way of staging the level of dementia. It may be unrepresentative of the stage of dementia in some subjects at least. In this study data are available on a much broader sample of test performance, and it is possible to stage the level of dementia on the basis of the overall test score. Figure 4 shows the groups broken down by stage of dementia according to Full score: over 65% at the top, 50-65 in the middle, and under 50 at the bottom. Patterns of performance are very similar to those seen in the breakdown by ORIENTATION score except that the fall in scores on each factor from top graph to bottom is if anything more orderly here. This probably reflects the fact that ORIENTATION score correlates highly with Full score, and that using Full score as the criterion variable inevitably puts limitations on how far the constituent factor scores are free to vary. As with the ORIENTATION breakdown, age is confounded with level of performance in the MID and Other groups but not in the DAT group. Similar comments therefore apply here as were made concerning the breakdown by ORIENTATION score. Partial correlation analyses involving factor scores, age, overall Full score, and Orientation score in various combinations did not lead to any great clarification or reinterpretation of patterns, and so will not be

Figure 4 Mean factor scores of patients who completed full testing, broken down by diagnostic category and mean % correct score (n of each group at top right).



Factors 1 to 4 are arranged from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

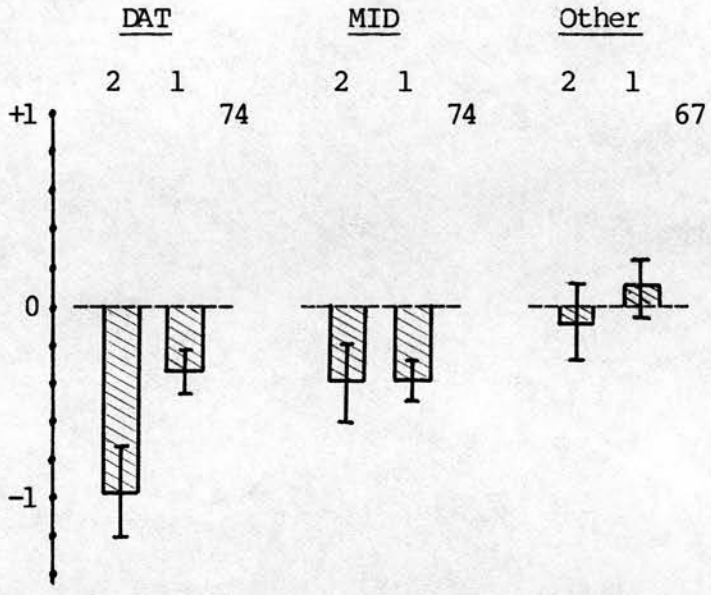
reported.

To bring into the picture those subjects who only received Short testing, a second principal components factor analysis (without iteration) was carried out, this time on the scores of only those tests included in the Short version. The patients included were only those 174 who had completed the Full version, to allow some comparability between the graphs which follow and those which have already been presented. This analysis produced just two factors; loadings (after varimax rotation) on each factor for the various tests are shown in Table 7.

Here Short factor 2 loads highly on ORIENTATION, DIGIT SPAN, SENTENCE REPETITION, AUTOMATIC SPEECH, and Identifying the three objects. In the context of this small group of tests, it might reasonably be considered a 'verbal' factor. Short factor 1 loads highly on SPIRAL MAZE 1 and 2, BOX-FILLING, PORTEUS MAZES, and constructing a SQUARE WITH STICKS. This can here be considered a 'performance' factor. Writing Name/Reading sentence does not load very highly on either factor, perhaps because it combines two different types of task, with both verbal and performance components. (The spread of scores in each half of the test alone is too narrow to permit their being analysed here as separate variables.)

Factor scores were then calculated for all patients tested, including those who only received Short testing. Figure 5 shows the mean factor scores of the three main diagnostic groups. (Short factor 2 is presented to the left of Sh.factor 1 because the former is most comparable to the first factor in previous graphs and the

Figure 5 Mean factor scores of patients who completed short testing, broken down by diagnostic category (n of each group at top right).



Factors arranged in the order 2,1 from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

Table 7. Varimax Rotated Factor Matrix of analysis of 11 test scores in all 174 subjects who completed full testing (Figures appropriately rounded and loadings higher than .50 underlined)

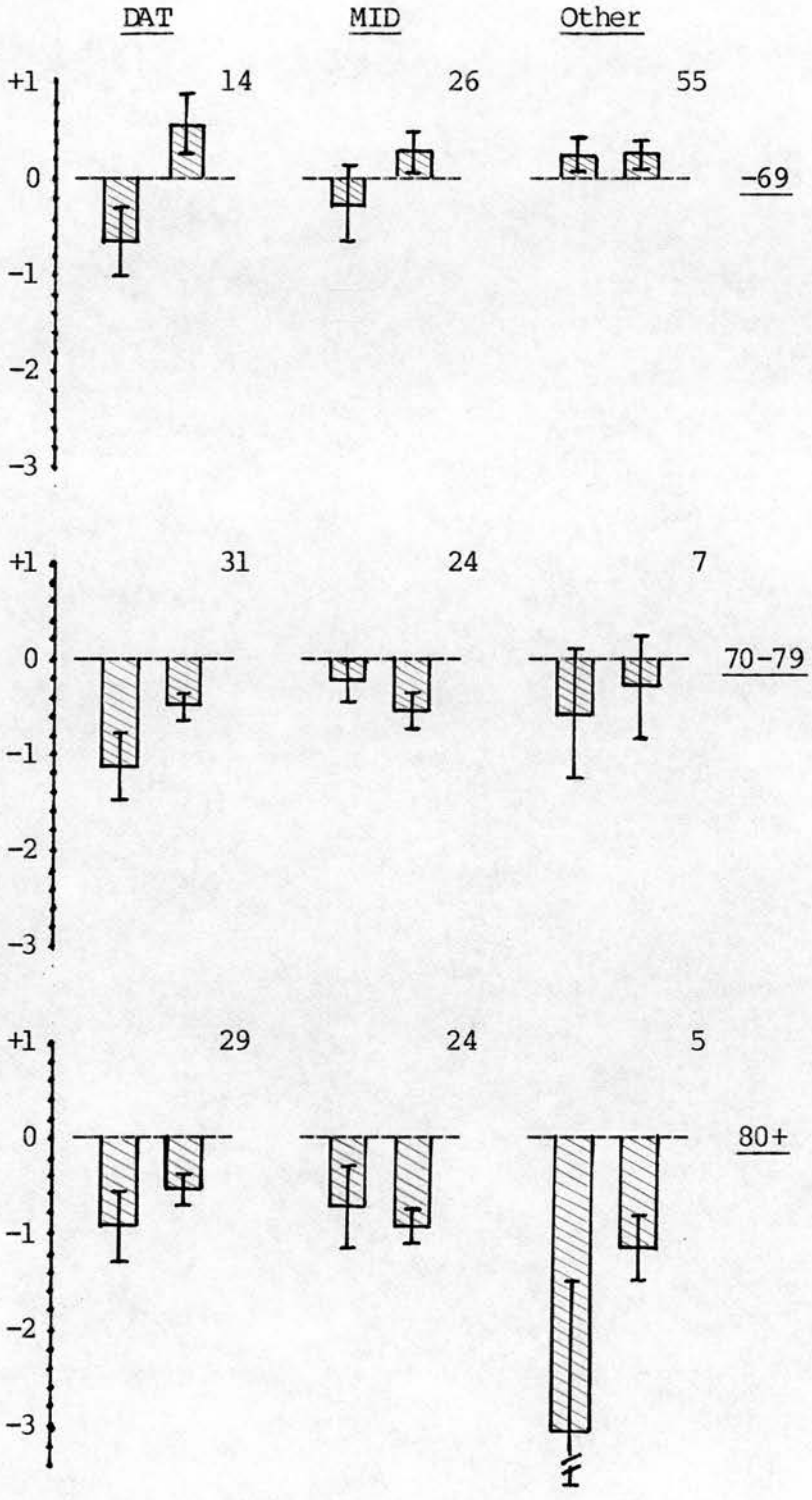
	Factor 1	Factor 2
ORIENTATION	.362	.682
DIGIT SPAN	.247	.669
SENTENCE REPETITION	.233	.814
AUTOMATIC SPEECH	.297	.750
Identifying SKP*	-.124	.636
Writing Name/Reading Sentence	.325	.317
SPIRAL MAZE 1	.912	.088
SPIRAL MAZE 2	.928	.057
BOX-FILLING	.816	.213
PORTEUS MAZES .	.665	.440
SQUARE WITH STICKS	.508	.361
Percentage of the variance explained	32.0	27.4
	Total	(59.4)

* from the Supplementary Dysphasia Test; the score on the Supplementary test itself is not included since only a proportion of subjects received it, while data were available for all subjects regarding whether a spoon, key, and pencil could be recognised and named.

latter is most comparable to the other factors in previous graphs.) The only point of note is that Sh.factor 2 is worse than Sh.factor 1 in the DAT group while in the MID and Other groups Sh.factors 2 and 1 are at about the same level as each other. An ANOVA on overall Short scores broken down by diagnostic group produced a significant result: in particular, Short score is lower in the DAT group than in the MID group, clearly as a result of the Sh.factor 2 difference. (ANOVAs on each Sh.factor alone broken down by diagnostic group were both significant.)

Figure 6 shows the groups broken down by age, as in Figure 2. In the DAT groups, Sh.factor 2 is not significantly different (by ANOVA) in the three age groups, while Sh.factor 1 is significantly higher in the youngest group than in the other two. In the young

Figure 6 Mean factor scores of patients who completed short testing, broken down by diagnostic category and age group (n of each group at top right).



Factors 2, 1 from left to right on each abscissa; ordinate shows factor score; bars show standard errors.

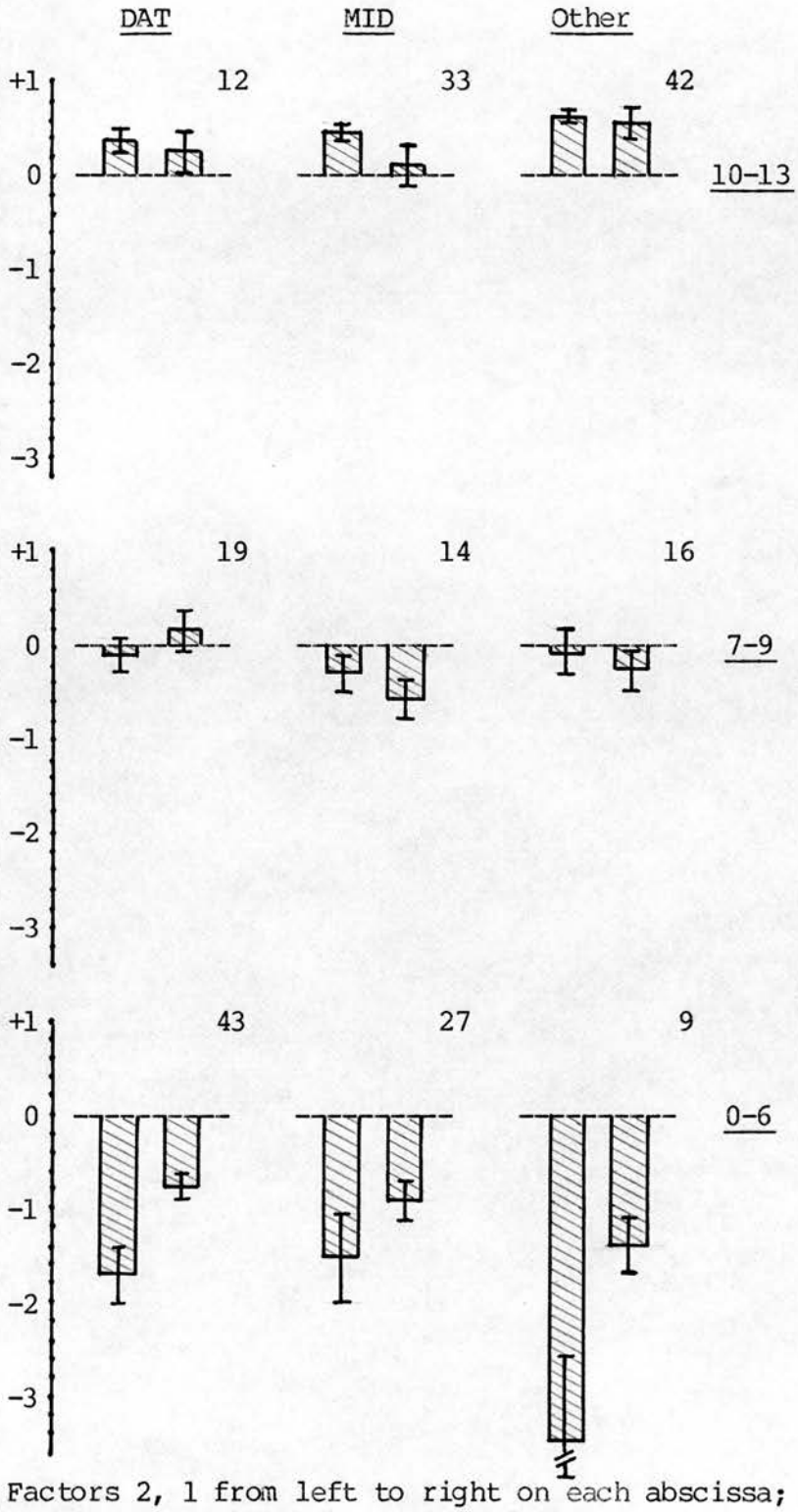
group Sh.factor 1 is considerably higher than Sh.factor 2, while this difference decreases in the older groups. However, these differences are not striking. In the MID group, patterns are similar to those in the DAT group, though Sh.factor 1 is little better than Sh.factor 2 even in the youngest group. Again Sh.factor 2 does not differ significantly between the three age groups, while Sh.factor 1 does. There are no surprises in the Other groups. The very poor performance in the small oldest group is attributable simply to the inclusion of some very impaired subjects in that group. Appropriate ANOVAs showed that age and overall Short score are confounded in all three diagnostic categories.

Figure 7 shows the groups broken down by ORIENTATION score, as in Figure 3. Both Sh.factors drop significantly from top to bottom in every diagnostic category, and there seem to be no particularly interesting pattern differences. Age is confounded with stage (as defined by ORIENTATION score) in MID and Other but not in DAT.

Differences in patterns on Short testing, then, are generally not striking and do not clearly parallel the results from Full testing. This may reflect the inclusion here of more severely impaired subjects or the limited range of tests used (and factors derived).

The factors lump a number of tests or abilities together: this is why they were used. The 22 test scores can be considered separately. To present patterns in a meaningful way, it is necessary to transform the scores in some way to take account of the fact that some tests are harder than others. The percentage scores do not allow for this: clearly it is easier to get 'full marks' i.e. 100% on NAMING OBJECTS than on FACE-NAME LEARNING, if

Figure 7 Mean factor scores of patients who completed short testing, broken down by diagnostic category and ORIENTATION score (n of each group at top right).



Factors 2, 1 from left to right on each abscissa; ordinate shows factor score; bars show standard errors.

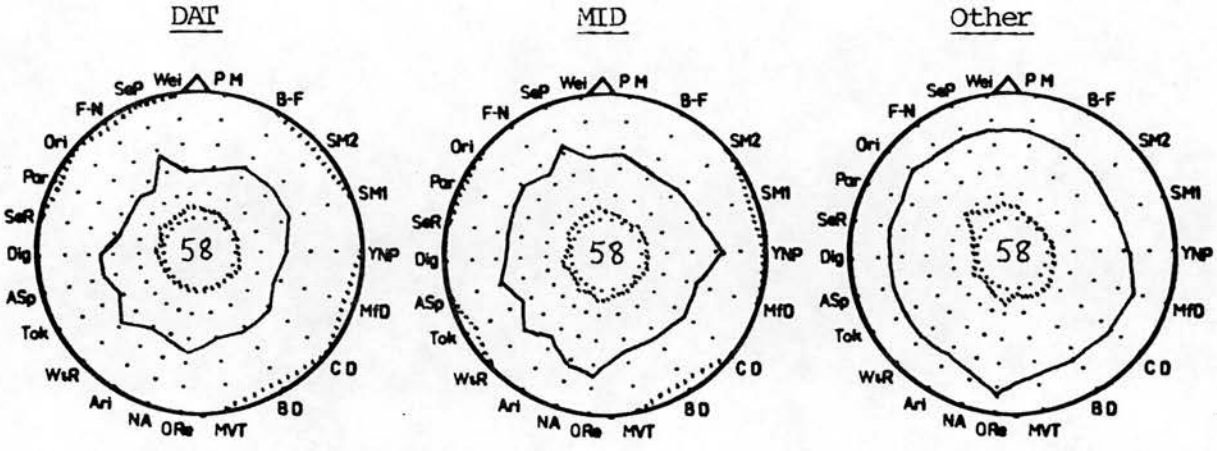
only because more people do. One option is to convert each of the 22 scores to standard scores such as z scores or some derivative such as t scores. Standard scores of this sort have been used by Rosen & Mohs (1982) to look at individual variability in DAT. For meaningful presentation of patterns, however, this method depends on the distributions of scores in each test being very close to normal. In the present study, many of the test score distributions do conform to this requirement reasonably well (and all well enough to justify the use of such techniques as factor analysis), but some, such as NAMING OBJECTS, do show somewhat skewed distributions. Another option is to convert scores on each test to percentile scores, so that the best performers receive scores close to 100 and the worst close to 0. This is a 'safer' transformation in that it makes no assumptions about the underlying distributions of scores, and of course it still takes account of the fact that some tests are inherently easier than others. Percentile scores (based on the distributions of scores of all 174 subjects who completed Full testing) were calculated.

Presentation of patterns in a comprehensible and reasonably brief form is difficult. Circular graphs are used where the innermost circle represents 0 or worst performance and the solid outer circumference represents 100 or best performance. The positioning of tests around the graphs was decided on the basis of firstly the known association between poor performance on certain tests and dysfunction of certain brain areas and secondly the sizes of the observed intercorrelations between tests. Such a view of association between tests and brain areas is clearly simplistic, but the tests have to be ordered somehow. The graphs are hence

convenient sketches or schematic representations of test performance: they are not supposed to be pictures of actual brains, and the author harbours no fantasies that areas of greatest atrophy or damage in subjects' brains would correspond to the biggest kinks in the graphs. Figures 8 to 14 use the circular graphs to present data on individual tests in subject groups broken down in the same way as in figures 1 to 7 (i.e. by diagnosis and then also by age, Orientation score, and Full score for subjects who completed Full testing; and by diagnosis and then also by age and Orientation score in those who completed Short testing), So Figure 8 corresponds to Figure 1, 9 to 2, 10 to 3, 11 to 4, 12 to 5, 13 to 6, and 14 to 7. Means are shown in solid line (except where the number of subjects is very small) with the range of scores in dotted line. (In the graphs relating to Short testing, score on Supplementary Dysphasia has had to be estimated from score on the Token Test in subjects who received the latter only. This was done on the basis of plotting Supplementary Dysphasia scores against Token Test scores for those subjects who completed both. Such estimates have not been used in any computations of Short scores, Factor scores, or any other calculations or analysis: they appear only in these circular graphs.)

Some of the performances on particular tests will be considered in later chapters concerning further analysis of certain abilities. These graphs serve general puposes. They show that there are no subgroups of subjects (when categorised according to diagnosis, age, or level of performance) where the patterns of factor scores 'hide' large deviations or unusual patterns of performance amongst those individual tests contributing to the factor scores. This

Figure 8 Mean percentile scores of patients who completed full testing, broken down by diagnostic category (n of each group in centre).



Dotted lines indicate ranges of scores.

Key:

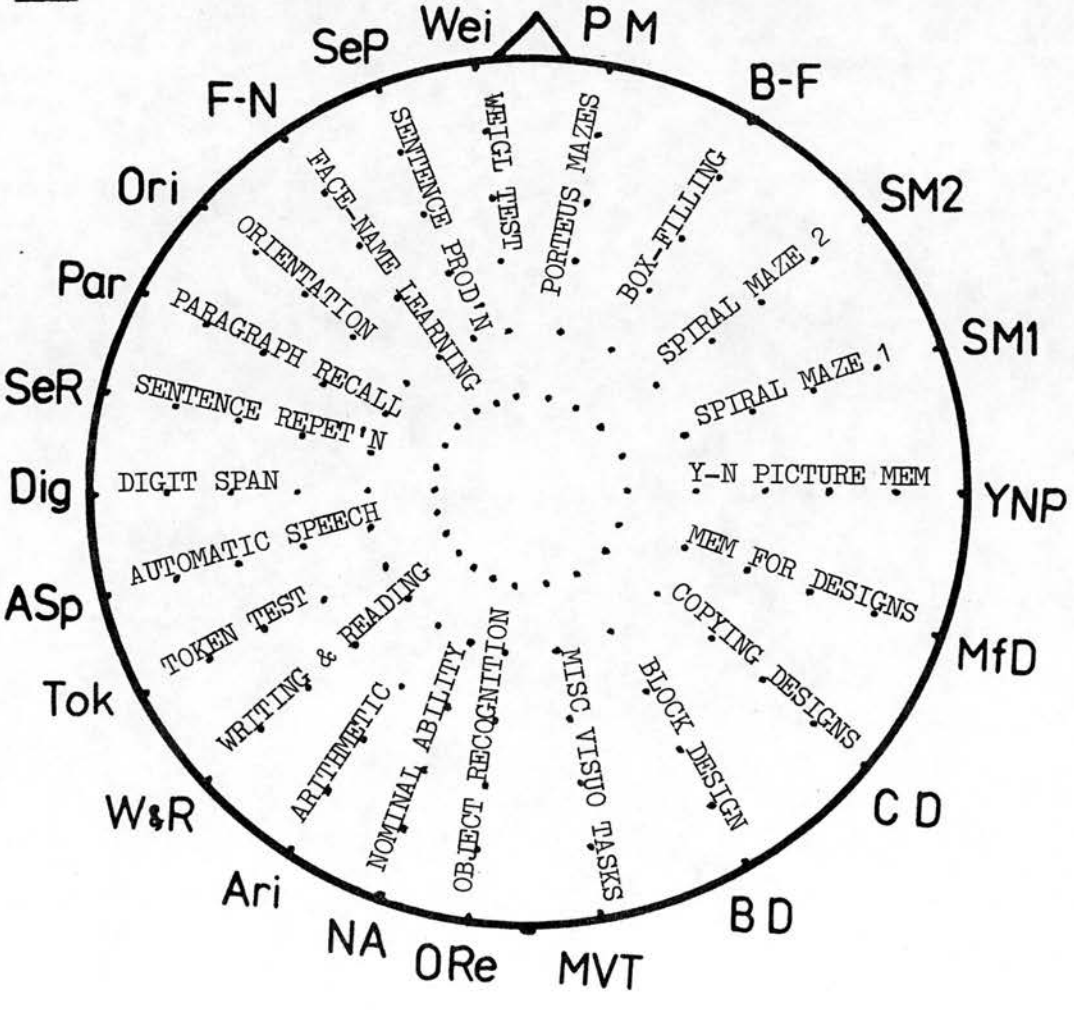
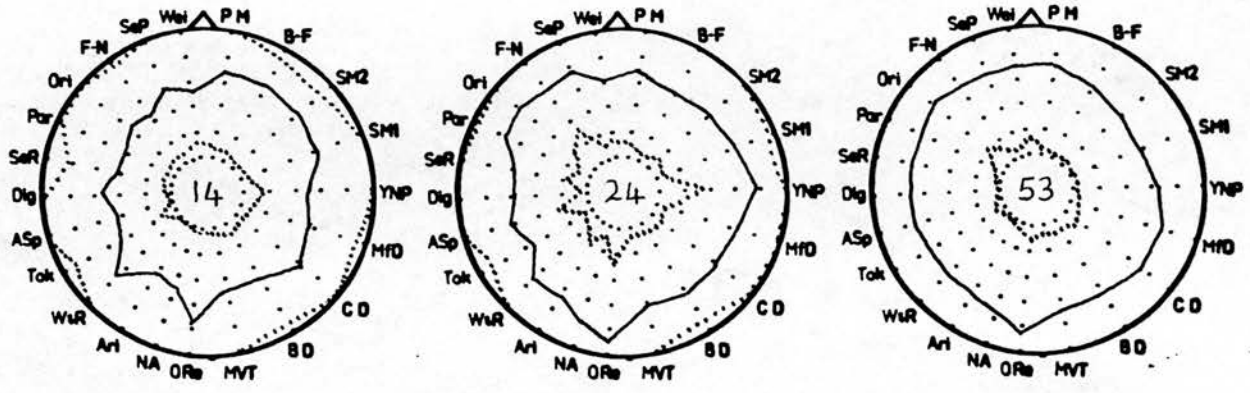


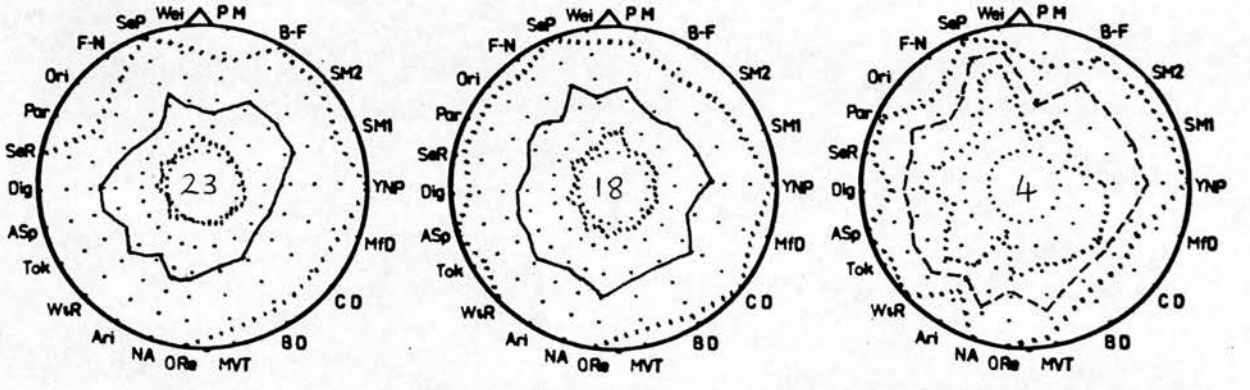
Figure 9 Mean percentile scores of patients who completed full testing, broken down by diagnostic category and age group (n of each group in centre).

DAT MID Other

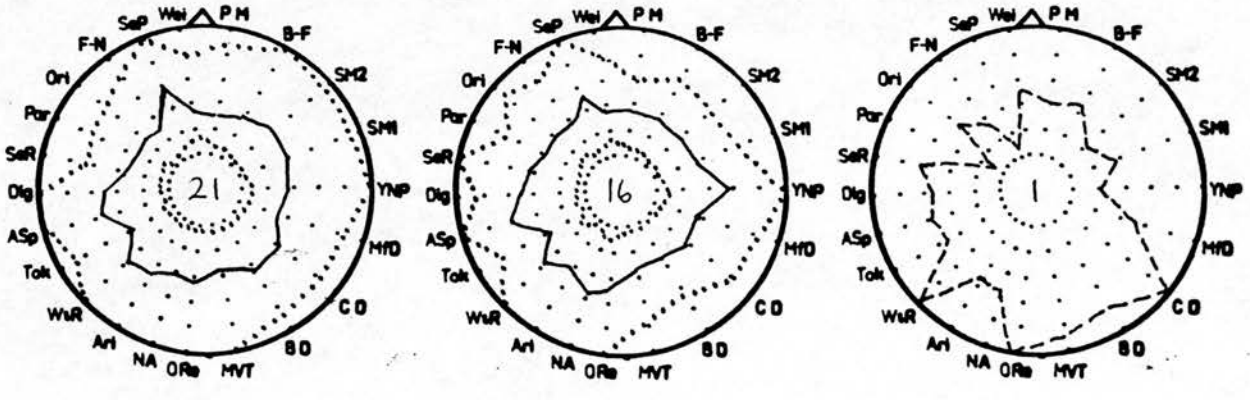
-69



70-79

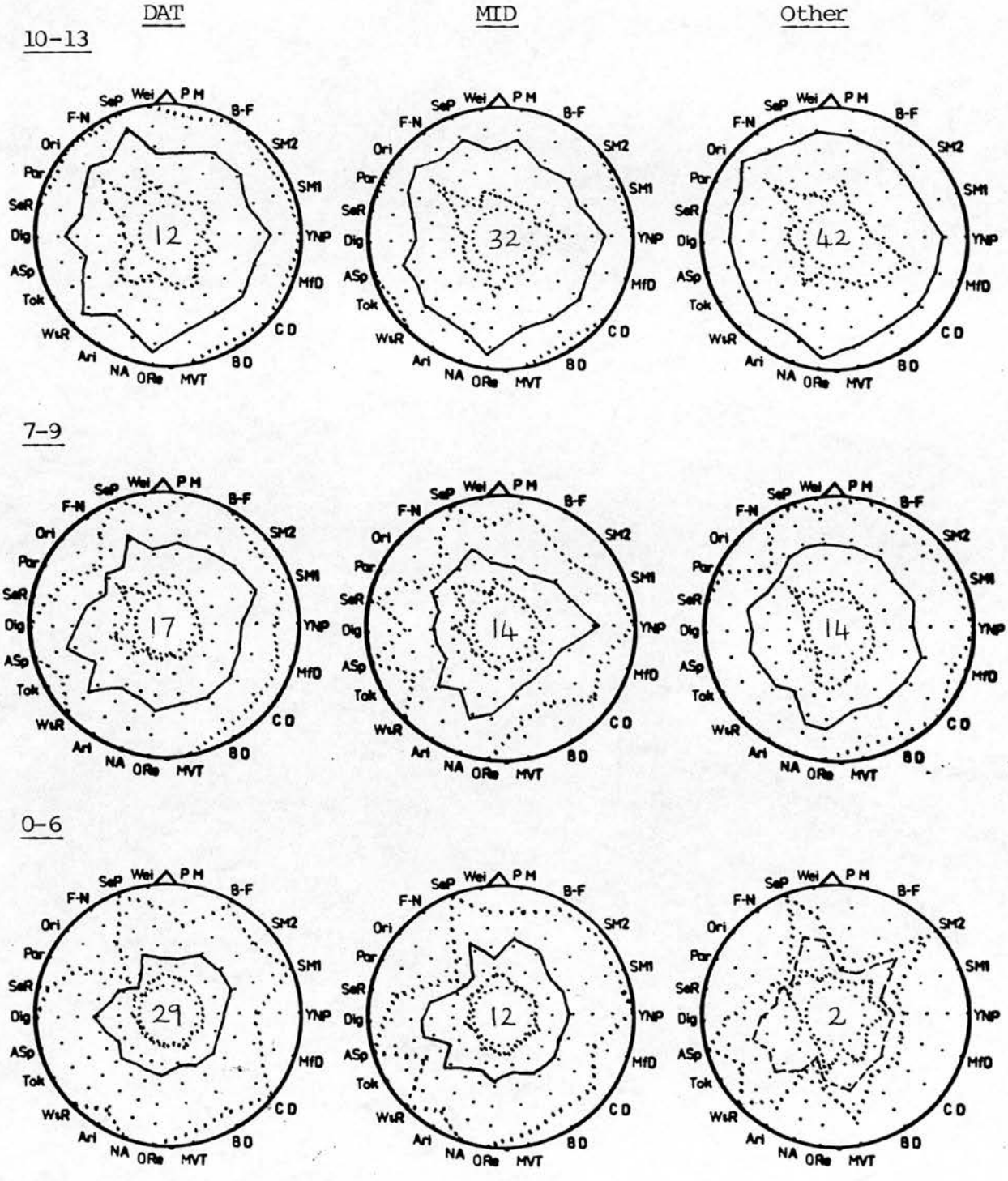


80+



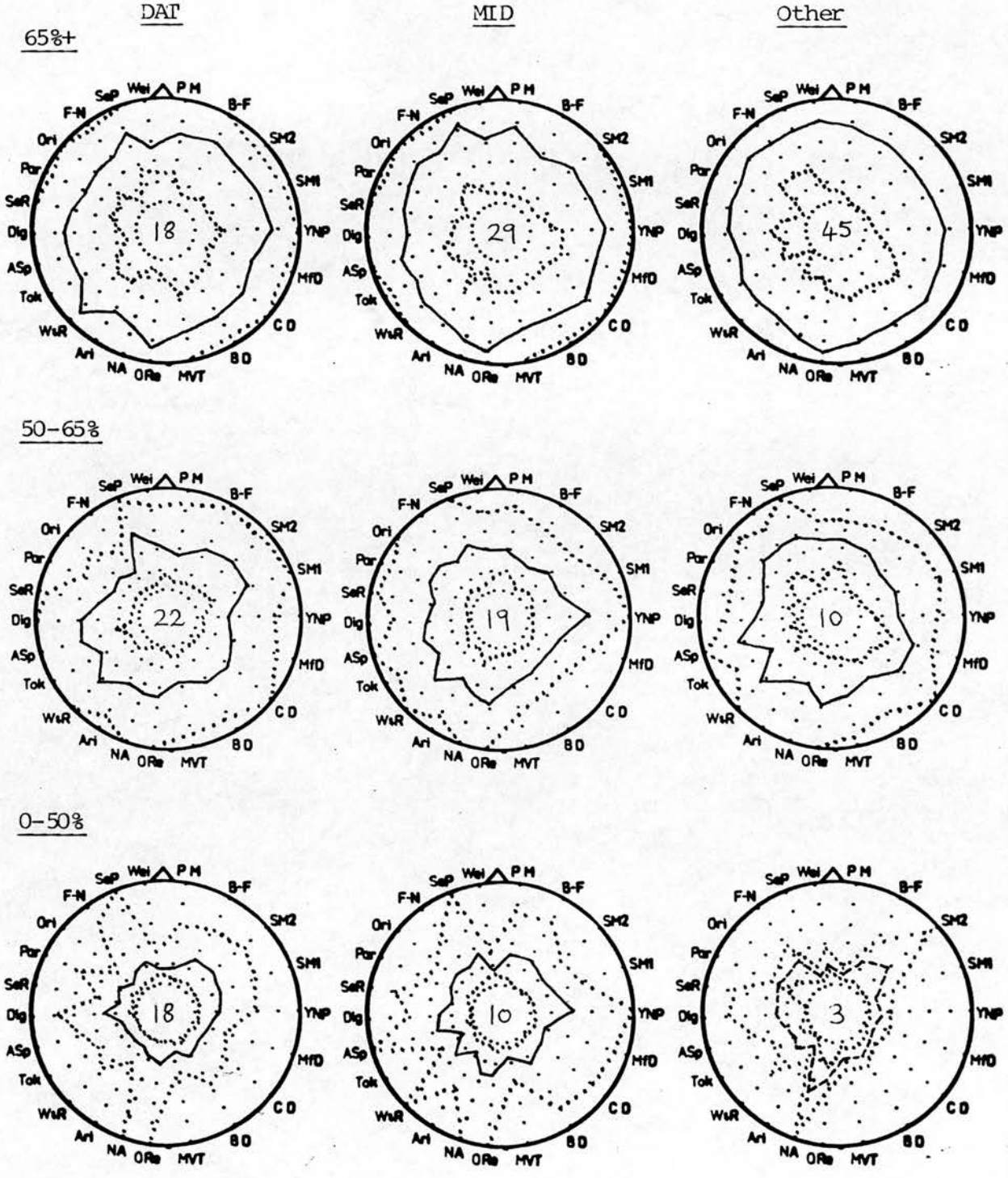
Dotted lines indicate ranges of scores.

Figure 10 Mean percentile scores of patients who completed full testing, broken down by diagnostic category and ORIENTATION score (n of each group in centre).



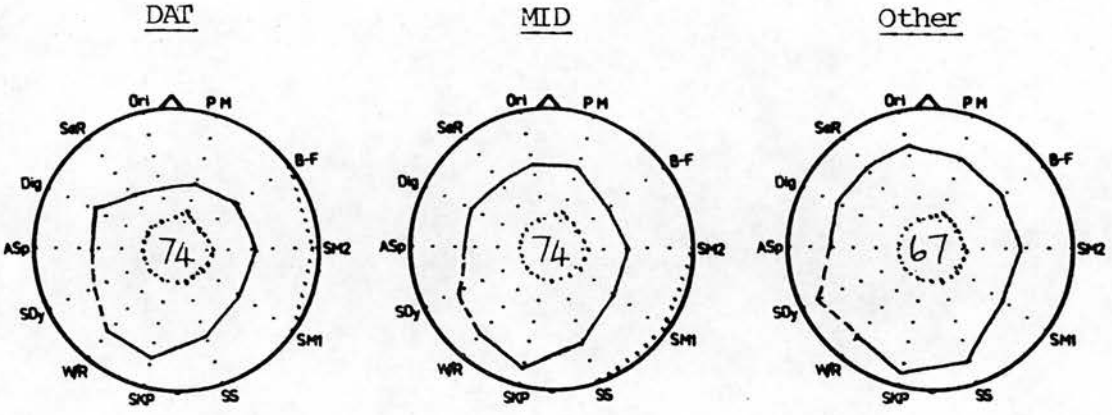
Dotted lines indicate ranges of scores.

Figure 11 Mean percentile scores of patients who completed full testing, broken down by diagnostic category and mean % correct score (n of each group in centre).



Dotted lines indicate ranges of scores.

Figure 12 Mean percentile scores of patients who completed short testing, broken down by diagnostic category (n of each group in centre).



Dotted lines indicate ranges of scores.

Key:

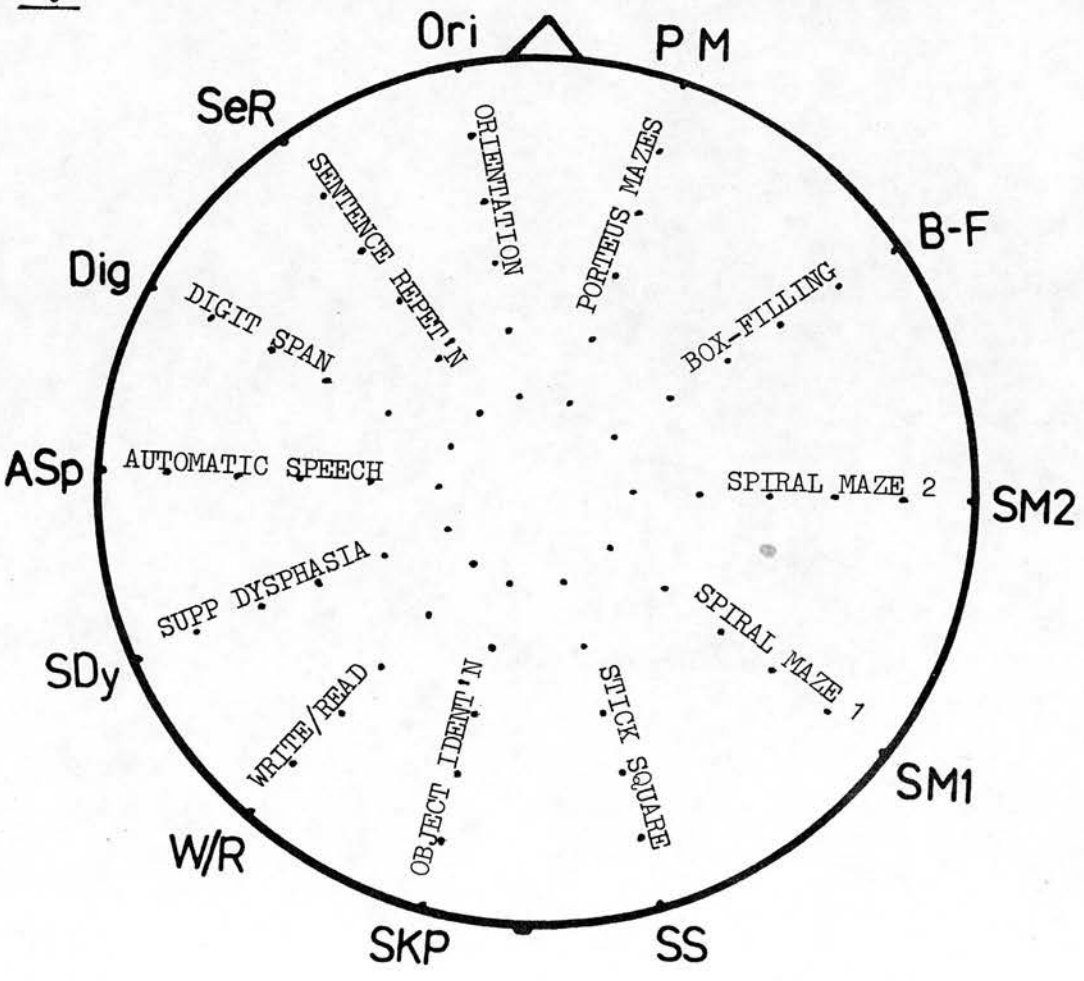
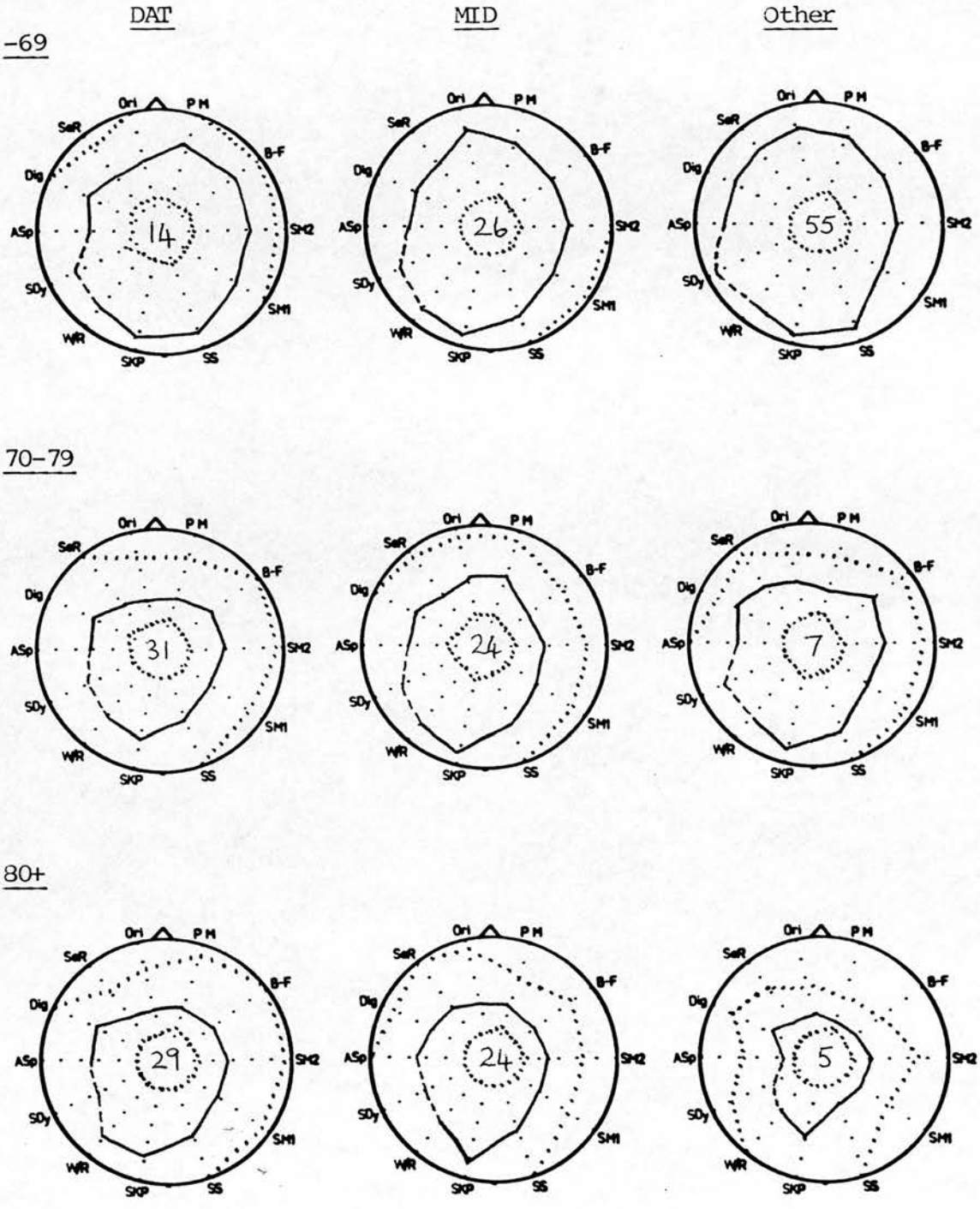
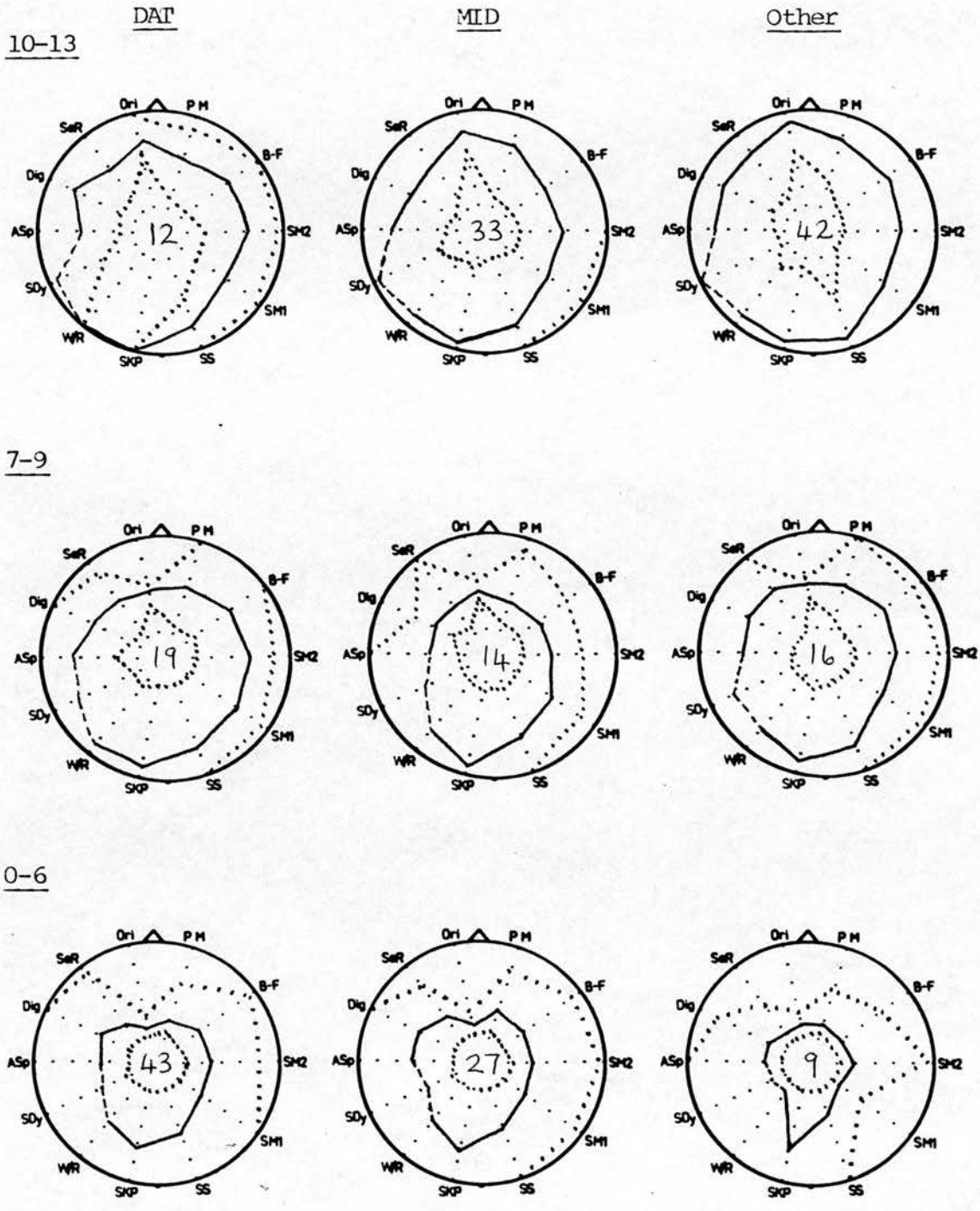


Figure 13 Mean percentile scores of patients who completed short testing, broken down by diagnostic category and age group (n of each group in centre).



Dotted lines indicate ranges of scores.

Figure 14 Mean percentile scores of patients who completed short testing, broken down by diagnostic category and ORIENTATION score (n of each group in centre).



Dotted lines indicate ranges of scores.

might have been the case if two tests which had correlated highly together in all subjects tested, and hence had become part of the same factor, showed great dissociation in some particular subgroup of subjects. Hence the factor patterns can be taken as reasonable or representative descriptions of the patterns on the individual tests.

The shapes of the performance 'circles' seem fairly similar in many subject subgroups, but of course these are averages. The graphs also show the ranges of scores in each subgroup, and perhaps illustrate the degree of spread or variability more clearly than do the standard error bars on the factor graphs. For example, Figure 1 indicated that the MID subjects in general showed less visual agnosia and nominal dysphasia (i.e. Factor 4) than the DAT subjects did. This is also apparent from Figure 8; but here it is evident that the range of performance in each group stretched from close to zero to close to 100, emphasising that the presence or absence of these particular impairments can tell us nothing about the likely diagnosis in an individual case; or, to put it another way, a firm diagnosis of DAT or MID cannot tell us whether such impairments will be present or not. This tendency for the ranges of scores to be very broad is similarly apparent in Figures 9 and 13, where the subjects are broken down by age. In Figures 10 and 14, where subjects are categorised according to their score on ORIENTATION, the ranges on "Ori" are inevitably small: but the ranges on most other tests are again very broad. Only in parts of Figure 11, where subjects are categorised according to their Full score, do the ranges begin to show much curtailment, and this of course is partly a reflection of the inherent restrictions produced by the method of

categorisation.

Focality of performance.

The percentile scores can be used in another interesting way. Considering the neuropathological bases of DAT and MID, MID subjects might be expected to show more 'focal' patterns of impairment than would DAT subjects: not necessarily any particular or consistent focal deficits (as would have shown up on previous graphs), just any focal pattern as opposed to a global or non-specific or undifferentiated pattern of cognitive impairment. Since the percentile scores were based on a fairly large number (174) of impaired subjects, it seems reasonable to suppose that a pattern of performance appearing on the circular graphs as a perfect circle (with a percentile score of, say, 60 on every test) represents a global or non-focal impairment of cognitive functioning. Focal impairment, on the other hand, might be represented by a very 'spiky', star-shaped pattern or by any D-shaped asymmetrical pattern where some tests were performed well and others badly. A crude numerical measure of how 'focal' a subject's performance is can be calculated from his percentile scores: the sum of the absolute values of the differences between each test percentile score and the mean percentile score for that subject on all tests, divided by the number of tests, i.e.:

'Focality' = (Abs(p1-Mp) + Abs(p2-Mp) + ... + Abs(pn-Mp)) / n

where p1 to pn are a subject's percentile scores on the n tests and Mp is the mean of all his percentile scores. By this method, a perfect circle on a circular graph (equivalent to a flat horizontal line on a conventional graph) yields a focality score of zero, and any jagged or asymmetrical pattern gives a higher focality score.

Such scores were calculated for every subject's performance on Full testing (with n of tests = 22) and on Short testing (with n of tests = 11). Table 8 shows the focality scores for subjects who completed Full testing, broken down by Diagnosis and then by age group as well. Parametric statistics are applicable since the distributions of the focality scores appears roughly normal despite the flat distribution of the underlying percentile scores.

Table 8 Focality scores on Full testing broken down by Diagnosis and age group. (Standard deviations and n of cases in parentheses.)

	DAT		MID		Other	
	mean	(S.D; n)	mean	(S.D; n)	mean	(S.D; n)
All	17.0	(5.4, 58)	17.6	(5.1, 58)	16.0	(5.3, 58)
-69	17.9	(6.2, 14)	17.5	(5.4, 24)	15.7	(5.2, 53)
70-79	16.3	(5.0, 23)	18.0	(5.8, 18)	19.6	(---, 4)
80+	17.2	(5.6, 21)	17.2	(3.8, 16)	20.7	(---, 1)

It can be seen that the figures are all very similar. There are no significant statistical differences between any of the groups (where numbers are large enough to allow such testing).

It seems that DAT subjects - regardless of age - are just as likely to show focal patterns of impairment as are MID subjects. Or, in more practical terms, one might say that a focal pattern of impairment on testing is in itself (i.e. without consideration of its nature) no more indicative of MID (or even perhaps of other neurological conditions) than of DAT. This might be seen as unsurprising since some people with DAT are said to show focal patterns of impairment: but such DAT subjects are generally held to be those with onset at a relatively young age. The breakdown by age as well as diagnosis still shows no differences between

subgroups of subjects.

Although the accuracy of diagnosis can never be 100% in this kind of study, it is not very plausible to explain the finding entirely in terms of misclassification of patients. The remaining possibilities are that focal patterns of performance are common in DAT at all ages; that focal patterns are not particularly common in MID; or that the focality measure is invalid. Table 9 shows relevant correlations between the focality score on Full testing and age and Full score.

Table 9 Correlations between the focality score on Full testing and Age and Full score. (Pearson product-moment correlations; n=58; ** significant at .01 level or better, 2-tailed test.)

	Age	Full
DAT	-.02	.47**
MID	-.02	-.12
Other	.37**	-.54**

The positive correlation between focality and Full scores in DAT, indicating that better subjects showed more focal performance, might suggest a floor effect in poorer subjects where focality scores could not be high because all percentile scores were low. Such an effect would act against a trend for DAT subjects to show just as high focality scores as MID ones, and so cannot account for the result. There is no suggestion of artefact in the MID group. The negative correlation in the Other group might indicate a ceiling effect in better subjects where focality scores could not be high because all scores were high. As with any system of test scaling, it is debatable how comparable the difference in 'actual' performance between say the 50th and 55th percentiles is to that between the 90th and 95th. This may have contributed to the

findings.

Table 10 shows the focality scores based on Short testing, broken down by diagnosis and age. These are perhaps less interesting than those based on Full testing in view of the smaller number and more restricted range of tests involved. However, the figures generally seem to show the same pattern as seen with the Full measure. Again there are no significant differences between groups. Table 11 shows relevant correlations.

Table 10 Focality scores on Short testing broken down by Diagnosis and age. (Standard deviations and n of cases in parentheses.)

	DAT		MID		Other	
	mean	(S.D; n)	mean	(S.D; n)	mean	(S.D; n)
All	20.8	(8.8, 74)	21.1	(7.9, 74)	18.5	(8.4, 67)
Age						
-69	21.8	(7.4, 14)	18.9	(9.2, 26)	18.3	(8.0, 55)
70-79	19.3	(9.6, 31)	22.1	(6.4, 24)	23.4	(7.2, 7)
80+	21.8	(8.5, 29)	22.7	(7.3, 24)	13.8	(12.5, 5)

Table 11 Correlations between the focality score on Short testing and age and Short score (and focality score on Full testing). (Pearson product-moment correlations; ** significant at .01 level or better, 2-tailed test.)

	Age	Short	Full focality score (n=58)
DAT (n=74)	.04	.55**	.53**
MID (n=74)	.15	.19	.78**
Other (n=67)	.20	-.06	.85**

The pattern of correlations in the first two columns is roughly similar to that found with the Full focality measure, and can be interpreted similarly. The correlations between the Full focality measure and the Short focality measure suggest that the two

measures do have some degree of consistency or correspondence, though of course this might simply reflect a similar degree of artefact in the calculation of both.

It can be concluded that there is no good evidence of differences between groups in focality of performance, but that the possibility of artefact does not rule out the possible existence of some such differences.

Another way of considering whether one diagnostic group shows more 'focal' performance than another is to look at mean inter-test correlations in the groups. This was done for each of the three main diagnostic groups, though not for any subgroups within diagnostic categories. Table 12 shows mean inter-test correlations for subjects completing Full testing and for those completing Short testing. (The actual inter-test correlations appear in Appendix 3.)

Table 12 Mean inter-test correlations (Pearson product-moment) for subjects completing Full and Short testing. (Standard deviations in parentheses.)

	x Full		y Short	
DAT	.47	(.16)	.60	(.10)
MID	.44	(.18)	.58	(.13)
Other	.48	(.15)	.65	(.14)

x number of subjects = 58, but number of correlations on which each mean is based = 231, i.e. all intercorrelations between 22 tests excluding redundant duplicate correlations.

y number of subjects = 74 for DAT and MID and 67 for Other, but number of correlations on which each mean is based = 55, i.e. all intercorrelations between 11 tests excluding redundant duplicate correlations.

By independent t-test, none of the 3 diagnostic groups differs significantly from either of the others regarding the Full intercorrelations. As regards the Short ones, the DAT and MID group

means are both significantly lower than the Other mean (at .02 and .01 levels respectively, 2-tailed test), while the DAT and MID means are not significantly different from each other. In other words there is some indication from the Short test intercorrelations that the DAT and MID groups show more focal performance than the Other group, but with no difference between the DAT and MID groups themselves. The results are compatible with the focality results presented above and can be seen as confirmation of the lack of measurable differences between diagnostic groups as regards focality of performance.

Inter-individual variability.

Rosen & Mohs (1982) comment on the inter-individual variability in patterns of performance within their samples of DAT subjects, and other writers have noted variation in characteristics of dementing subjects within particular categories of dementia. Perhaps the most interesting impression from this part of the analysis concerns this within-group range or variation. Circular graphs of individual subjects' performances appear in Appendix 5. The graphs are arranged in order of increasing age within each diagnostic category; scores at initial testing are shown in solid dark line, with retest scores where available in grey or yellow. A glance through this appendix gives some idea of the great variation in patterns of impairment between subjects even within particular diagnostic and age categories.

Woods & Britton (1985, p105) speculate as to whether this variability is a result of the subjects assessed being at different stages of the condition with different levels of impairment, of them having different patterns of premorbid abilities, of them

having different ages of onset, or to actual differences in the nature of DAT in different subjects. Roth (1971) found considerable individual variation in the relative number of senile plaques present in different lobes of the brain, and even in adjacent gyri, in patients with DAT (though no part of the neocortex was found to be entirely spared). Perry (1986) similarly reports that the cortical distribution of both tangles and plaques may be surpisingly focal and variable, particularly in subjects dying older than 75 or 80 years. Whatever the underlying reason, it is fair to say that this study shows that the kind of focal impairments people show, or whether they show them at all, varies greatly within each of the groups.

The variability found here was such that any attempt at deriving a 'cognitive taxonomy' in dementia to supplement traditional diagnostic categorisation seemed unlikely to succeed. Types within DAT and MID (other than those based on the age-related differences already described) were looked for by examining the individual circular graphs and graphs of individuals' factor scores, but nothing of note emerged. Statistical analysis was not attempted. The only comments worth making about the diagnostic groups within the Other group (where there are enough subjects to judge) seem to be that subjects do not generally show severe impairment, there is considerable inter-individual variation within diagnostic groups, and a characteristic pattern does not seem to be present in any diagnostic group. Obviously the numbers are far too small to treat these as worthwhile conclusions.

Discriminant function analyses could be undertaken to determine the best discriminators between the groups on testing for purposes

of differential diagnosis. This was not done for reasons outlined in the introduction in comments on previous work by Perez et al (1975). Also, the groups had already been diagnosed on the basis of other (briefer and simpler) clinical criteria. Considering the way the groups were defined and selected in the first place, no test discriminators could be 'better' than the Hachinski index (and related clinical variables). Such analyses could only have been worthwhile had post-mortem information been available to prove the clinical diagnoses or had a 'doubtful' diagnostic group been included and subsequently followed up to post-mortem examination or some other convincing determination of diagnosis. In fact, given the great variation in patterns of performance described above (both within and between diagnostic groups), it seems highly unlikely that general statements could be derived which might have clinical value in establishing diagnosis in an individual case.

Some potentially important influences on patterns of performance will now be considered.

The Influence of Premorbid IQ.

Of the 174 subjects who completed Full testing, twenty-three (9 DAT, 9 MID, and 5 Other) were considered too be too dyslexic or dysphasic for a valid estimate of premorbid IQ to be obtained using the NART. This judgement was based on subjects' ability to read the "quick brown fox..." sentence and the Burt words given as a prelude to the NART. Hence 151 subjects (49 DAT, 49 MID, and 53 Other) did have premorbid IQ estimated. Judging by these samples, the three diagnostic groups did not differ significantly by analysis of variance on estimated premorbid IQ. The means were 102.3 (S.D. 8.1) for DAT subjects, 102.7 (S.D. 9.9) for MID subjects, and 104.7

(S.D. 9.9) for Other subjects. Therefore any overall differences previously described between the three groups in patterns of performance are unlikely to be attributable to the influence of premorbid IQ.

It is however notable that scores on the great majority of tests used in this study correlate significantly with estimated premorbid IQ, sometimes quite highly so. Correlations appear in Appendix 3. There seem to be two broad lines of argument to account for this. One is that intelligence continues to 'shine through' or show its influence despite the ravages of a dementing condition, i.e. that intelligence may in some way 'buffer' a person against impairment produced by a dementing process. The other is that this particular estimate of premorbid IQ is no such thing in such a sample, but simply yet another index of the severity of dementia. If scores on the test are lowered by dementia then the correlations are not surprising since virtually all other tests correlate significantly with each other as well. The NART was designed for use as a comparison indicator in subjects where cognitive impairment is in doubt: this is clearly not the case with subjects in the the present study. Other relatively robust or impairment- resistant tests, such as Vocabulary, are known to show decline in dementia eventually, and the NART may well also do so. Quite possibly the criteria for accepting a subject's performance on the NART were too lax in the present study. For example, the longest of the regular words used to screen reading ability was not as long as some of the NART words: it is possible that as dementia progresses there is especial difficulty in synthesising longer words, which might outwardly appear just as if the subject had never known the word in

question (leading of course to a lower IQ estimate).

It is notable that the correlation between estimated premorbid IQ and estimated duration of condition was -.35 in the 40 DAT subjects where both measures were available. This strongly suggests that the IQ estimates have been influenced by the severity of dementia. (The correlation in 40 MID subjects was negligible, only .03, but here severity of dementia is hardly related to duration of condition, as will be described below, so this does not constitute a contradiction. The correlation in the Other group is irrelevant.) There were also weak negative correlations between age and estimated premorbid IQ, which would not be expected. Further circumstantial evidence that NART performance has been affected by severity of dementia appears in Table 13, which shows correlations between Full score and estimated premorbid IQ at different levels of severity of dementia.

Table 13 Correlations of N.A.R.T. premorbid IQ estimate with Full score, broken down by stage of dementia as defined by Orientation score, in each diagnostic group and in the three combined. (Pearson product-moment correlations; n of cases in parentheses.)

Orient. score	DAT r	(n)	MID r	(n)	Other r	(n)	All r	(n)
0-13	.70	(49)	.56	(49)	.58	(53)	.58	(151)
10-13	.78	(12)	.57	(28)	.62	(39)	.58	(79)
7-9	.60	(17)	.26	(12)	.35	(13)	.40	(42)
0-6	.72	(20)	.49	(9)	—	(1)	.63	(30)

Unless the buffering or 'shining through' of premorbid intelligence continues even at quite severe levels of dementia, it seems likely that NART estimates are affected by dementia and hence that the overall test performances previously described are not simply a reflection of how clever the subjects used to be. It is a

pity that the NART was not repeated on at least a sub-sample of subjects after 10 months at the second testing, in which case a much clearer picture of the effects of progressing dementia on premorbid IQ estimates might have emerged. Because of the possible confounding of degree of dementia and estimated premorbid IQ, no attempt is made in this thesis to use any sort of deterioration measure or index based on the comparison of estimated premorbid IQ with other test performance measures.

The Influence of Drugs.

In an ethically remarkable study involving the stopping and starting of medication in schizophrenic and depressed patients, Killian et al (1984) concluded that psychotropic medication did not affect performance on a variety of neuropsychological tests in any manner. They thought that results of studies using such tests with medicated patients could therefore not be attributed to drug effects. Such surprising findings unfortunately do not generalise easily to the subject groups used in the present study, and some psychotropic medications are thought to have an adverse effect on performance in elderly and cognitively impaired people. The data available here cannot test this possibility adequately. Subjects on medication may be a special group with different impairment characteristics regardless of the effects of the medication. Table 14 shows Full, Short, and Factor scores at initial testing of subjects on psychoactive medication and those on none in each diagnostic group and in the three combined.

Mean

Table 14 Full, Short, and Factor scores of subjects on psychoactive medication and those on none in each diagnostic group and in the three combined. (Standard deviations in parentheses; + significant at .10 level, * .05 level, ** .01 level by independent t-test, 2-tailed test.) (NB See Appendix 7.)

	DAT		MID		Other		All	
	Drugs	None	Drugs	None	Drugs	None	Drugs	None
n =	17	41	22	36	28	30	67	107
Full	.58 (.17)	.55 (.16)	.67 (.12)	.63 (.17)	.72 (.14)	.79 + (.15)	.67 (.15)	.65 (.19)
Factor 1	-.38 (1.11)	-.77 (.73)	.16 (.85)	-.03 (.94)	.42 (.89)	.79 + (.67)	.13 (.98)	-.08 (1.01)
Factor 2	-.12 (1.07)	-.05 (1.13)	-.09 (.88)	-.14 (1.03)	.27 (1.01)	.13 (.82)	.05 (.99)	-.03 (1.01)
Factor 3	-.25 (.95)	.16 (.86)	-.12 (.76)	-.19 (1.01)	-.13 (.99)	.37 + (1.26)	-.16 (.90)	.10 (1.05)
Factor 4	.06 (1.03)	-.54 (1.50)	.27 (1.06)	.23 (.64)	.08 (.54)	.14 (.38)	.13 (.86)	-.09 (1.07)
n =	24	50	30	44	33	34	87	128
Short	.59 (.23)	.59 (.22)	.62 (.25)	.66 (.19)	.68 (.26)	.78 + (.21)	.63 (.25)	.66 (.22)
Sh Fact 2	-.76 (1.84)	-1.09 (1.86)	-.67 (2.26)	-.24 (1.27)	-.35 (2.19)	.14 (1.17)	-.57 (2.11)	-.47 (1.58)
Sh Fact 1	-.56 (1.00)	-.24 (.95)	-.42 (.95)	-.39 (1.11)	-.15 (.98)	.32 + (1.30)	-.36 (.98)	-.15 (1.14)

In the DAT and MID groups there are no significant differences. There is some trend for subjects on drugs to score better than those on none: presumably this is because less impaired people are more likely to be given psychoactive drugs rather than because drugs make people better at tests. In the Other group, subjects on drugs tend to be worse than those on none, but this is of little importance considering the mixed nature of this group in terms of both diagnosis and types of medication received. In each diagnostic group, the average age of those on drugs was compared to that of

those on none (for subjects completing Full testing and then for those completing at least Short). There were no significant differences, and so there is little likelihood of an important pattern of drug influence being masked by a confounding with age.

The Influence of Sex.

It is no surprise to find that the distributions of the sexes are uneven in the various diagnostic and age categories, as shown in previous tables, with a surplus of relatively old females with DAT and relatively young males with MID. Unfortunately this leads to a possible confounding of the age group variable and any sex differences in the patterns of performance previously shown in different age groups in each diagnostic category. The numbers of subjects are generally insufficient to allow a breakdown by sex within each age group in each diagnostic category. Collapsing the age groups, i.e. looking simply at male versus female in each diagnostic category, leads to confounding of age and gender but gives some idea of whether there are any striking overall sex differences. Table 15 shows Full, Short, and Factor scores at initial testing broken down by sex in each diagnostic group and in the three combined.

Allowing for the confounding of age and sex, there is no clear evidence that sex differences have been important influences on the patterns previously described. Traditionally in research into dementia, age differences have been studied more (and considered more important) than sex differences. In the absence of good evidence to the contrary, this study assumes that the age group rather than the sex differences are of primary importance.

Table 15 Full, Short, Factor scores, and age at initial testing broken down by sex in each diagnostic group and in the three combined. (Standard deviations in parentheses; + significant at .10 level, * .05 level, ** .01 level by independent t-test, 2-tailed test.)

	DAT		MID		Other		All	
	Male	Female	Male	Female	Male	Female	Male	Female
n =	18	40	33	25	30	28	81	93
Full	.60 (.17)	.55 (.16)	.69 (.16)	.59 * (.13)	.77 (.16)	.74 (.14)	.70 (.17)	.62 ** (.17)
Factor 1	-.45 (.99)	-.75 (.81)	.40 (.63)	-.43 ** (1.01)	.78 (.68)	.43 (.89)	.35 (.86)	-.31 ** (1.02)
Factor 2	-.27 (1.15)	.02 (1.08)	-.17 (1.05)	-.06 (.87)	.23 (.88)	.16 (.95)	-.04 (1.03)	.04 (.98)
Factor 3	.18 (.89)	-.03 (.91)	-.09 (1.00)	-.27 (.81)	.13 (1.22)	.12 (1.11)	.05 (1.06)	-.05 (.95)
Factor 4	.13 (.86)	-.58 + (1.54)	.22 (.97)	.28 (.56)	.00 (.47)	.23 * (.43)	.12 (.79)	-.11 (1.15)
Age	71.3 (9.7)	77.1 * (8.1)	66.1 (11.2)	73.9 * (13.5)	55.0 (11.3)	56.3 (12.8)	63.1 (12.7)	70.0 ** (14.4)
n =	22	52	40	34	34	33	96	119
Short	.61 (.25)	.58 (.21)	.70 (.21)	.58 * (.21)	.76 (.23)	.70 (.25)	.70 (.23)	.61 ** (.23)
Sh Fact 2	-1.08 (2.01)	-.94 (1.79)	-.14 (1.67)	-.74 (1.79)	.09 (1.62)	-.30 (1.89)	-.27 (1.78)	-.71 + (1.82)
Sh Fact 1	-.12 (.95)	-.44 (.98)	-.14 (1.02)	-.71 * (.99)	.18 (1.16)	-.01 (1.19)	-.02 (1.06)	-.40 * (1.07)
Age	72.5 (9.9)	78.5 ** (8.2)	67.7 (11.9)	75.0 * (12.2)	56.7 (11.9)	60.2 (15.5)	64.9 (13.0)	72.4 ** (14.0)

The Influence of Duration of Condition.

For as many subjects as possible, an estimate of the duration of their dementing condition was recorded. This was based on information (usually coming originally from relatives or a General Practitioner) in the hospital or Local Authority records. Semple et al (1982) noted that different acquaintances' estimates of the duration of a given patient's condition could vary enormously (with up to 3 years' discrepancy between estimates from relatives and doctors, the relative's estimate always being the longer). This is hardly surprising in a condition such as DAT which by definition has an insidious onset. People may tend to date the onset to a particular life event or occurrence; the frequency of contact of the observer with the dementing person may influence when dementia is first noticed; and the death of a spouse or carer may 'unveil' a dementia previously unnoticed by others. Therefore little attention will be paid to the estimates of duration. Correlations between estimated duration (to the nearest 6 months, which is what was recorded) and Full and Short scores are shown in Table 16. The shortfall between the numbers of cases here and the total numbers of people who were tested reflects the absence of the relevant information in subjects where no reliable informant had been available.

Even though these samples may not be representative of the whole groups (since there may be something special about subjects where no reliable informant is available to estimate the date of onset), the correlations are much as one would expect. DAT subjects with longer durations tend to show poorer test performances (as well as tending to be older). The correlations are not particularly high in

Table 16 Correlations between estimated duration of condition, age, overall Full score and overall Short score in subjects completing Full and Short testing respectively. (Pearson product-moment correlations; + significant at .10 level, * .05 level, ** .01 level, 2-tailed test.)

Full testing	Full	Age	Short testing	Short	Age
DAT (n=48)	-.35 *	.34 *	DAT (n=53)	-.48 **	.26 +
MID (n=47)	-.05	.21	MID (n=62)	-.36 **	.28 *
Other (n=44)	-.47 **	-.10	Other (n=51)	-.15	-.15

absolute terms (i.e. in terms of variance explained): to know how long someone has been showing signs of DAT will predict little about how impaired they are now likely to be. The very low correlation between duration and Full score in the MID group presumably reflects the fact that here the degree of cognitive impairment is likely to depend on the number and severity of the multiple infarcts rather than on the length of time in which a relentlessly progressive disease has had to progress (as one would imagine is the case in DAT). The significant correlation with Short score probably reflects the inclusion in this group of some very impaired subjects with long histories. The significant correlation in the Other group simply reflects the fact that many of the subjects in this group were suffering from consistently progressive conditions.

Other possible influences.

The influence of motivation cannot be assessed, though the author's opinion was that subjects tested were sufficiently well motivated for their performances to be an adequate reflection of their abilities. Miller (1977, Ch5) suggests that motivational

influences are probably not usually major influences on test performance in dementia, particularly where a pattern of impairment in some areas rather than others can be demonstrated. There is no reason to suppose in the present study that motivation would have varied systematically depending on age group or severity of dementia. Similarly the influence of mild sensory impairments cannot be judged. O'Neill & Calhoun (1975) demonstrated a relationship between such impairments and cognitive performance: a third factor explanation in terms of a process of deterioration causing both sensory and cognitive impairment does not seem adequate. All that can be said is that subjects with clinically demonstrable sensory impairments which would have been likely to affect test performance were excluded.

The Factor Structure in the Three Groups Separately.

Factor analysis has so far been used to group test scores into a manageable number of variables. It can also be used to compare the structure of abilities in different groups (within the limitations imposed by the choice of tests).

Various factor structures concerning abilities in dementia have been reported, with numbers of factors ranging from one, for example in a study by Dixon (1965) using various brief cognitive tests, to many, for example in a study by Gustafson & Hagberg (1975) involving a large number of wide-ranging measures and a smaller number of PDAT subjects (thus making the analysis highly suspect). Differences between studies will in part reflect the number, nature, and range of the tests used, the severity of subjects' impairment, and the method of analysis; but some studies have suggested the possible existence of separate

neuropsychological dimensions in dementia. Studies generally do not compare one type of dementia with another. This section compares factor structures in DAT and MID, and also allows a check that factor structure derived from all subjects and used in the main analyses did not grossly misrepresent the factor structure in any single diagnostic group.

Principal components factor analyses (without iteration) were carried out for each group separately. As regards Full testing, the number of subjects in each group (58) is not very large compared to the number of tests (22), and so the results must be interpreted with reserve. All the relevant factor matrices (after Varimax rotation) appear in Appendix 4. A summary will be presented here.

With Full test scores, the analyses produced four factors in each diagnostic group (as with the three groups combined). To simplify interpretation, tests contributing to any factor at a given level (or higher) will be considered as 'part' of that factor: the most convenient criteria are .55 or higher in the rotated factor matrix for the DAT and MID groups (and for all three groups combined, as before) and .54 for the Other group. Using these criteria, Table 17 shows the factor structure in all the groups combined as before and in each group separately.

In the DAT analysis, Factor 3 emerges as a clear memory factor, though it also includes Porteus Mazes. Memory for Designs falls into another factor, number 1, presumably because in these subjects the memory component of the test has been swamped or over-ridden by the visuoperceptual or visuoconstructive one. Factor 1 appears to be a 'parietal' factor, containing tests concerning visuospatial abilities, language comprehension, reading and

Table 17 The factor structures on Full testing of each diagnostic group and of the three combined. Numbers indicate which factor each test is part of according to the criteria described in the text. (The presence of two digits indicates that the test contributed to more than one factor at the criterion level used, while a dash indicates that the test contributed to no factor at that level.)

	All (174)	DAT (58)	MID (58)	Other (58)
(Criterion level)	.55	.55	.55	.54
ORIENTATION	1	3	3	2
PARAGRAPH RECALL	1	3	3	2
MEMORY FOR DESIGNS	1	1	4	2
YES-NO PICTURE MEMORY	1	3	3	2
FACE-NAME LEARNING	1	3	3	2
PORTEUS MAZES	1	3	2	1
BLOCK DESIGN	1	1	4	1
WEIGL TEST	1	1	-	1
DIGIT SPAN	2	1	1	1 3
SENTENCE REPETITION	2	12	1	3
TOKEN TEST	2	1	1	1
SENTENCE PRODUCTION	2	2	1	3
WRITING AND READING	2	1	1	3
AUTOMATIC SPEECH	2	2	1	3
COPYING DESIGNS	2	1	4	1
MISC. VISUO. TASKS	2	1	4	1
ARITHMETIC	2	1	1	1
SPIRAL MAZE 1	3	4	2	4
SPIRAL MAZE 2	3	4	2	4
BOX-FILLING	3	4	2	4
OBJECT RECOGNITION	4	2	3	1
NOMINAL ABILITY	4	2	1	1
Percentage of the variance explained				
Factor 1	24.3	24.3	22.4	25.0
Factor 2	23.4	17.5	17.9	16.9
Factor 3	15.0	17.2	17.1	15.5
Factor 4	10.2	14.0	16.2	15.4
Total	73.0	73.0	73.7	72.8

writing, and arithmetic. The Weigl test also falls into this factor. Factor 2 has Object Recognition and Nominal Ability (like Factor 4 in the overall analysis) but also some other language tests, and seems largely to reflect expressive speech and word-

finding. Factor 4 is a psychomotor speed and accuracy factor corresponding to Factor 3 in the overall analysis.

In the MID analysis, Factor 3 is again a memory factor, though it also includes Object Recognition (but not Nominal Ability: this is the only analysis where the two tests fall in different factors). Memory for Designs appears in another factor, number 4, presumably for the same reasons as given in the DAT analysis. Factor 4 appears to be a clear visuospatial/visuoconstructive factor. Factor 1 emerges as a clear language factor, and includes Nominal Ability; Arithmetic unsurprisingly appears in the same factor. Factor 2 is the pschomotor speed and accuracy factor corresponding to Factor 3 in the overall analysis, but here Porteus Mazes are also included: the use of time bonuses for quick completion in this test may have been particularly relevant in MID subjects. The Weigl test does not contribute particularly highly to any test: its contribution is spread over a number of factors.

In the Other analysis, Factor 2 emerges as a very clear memory factor. This is the only group where Memory for Designs really appears to be assessing memory more than visuospatial abilities. Factor 1 emerges as a rather complex mixture of visuospatial, language, and 'frontal' tests, while Factor 3 contains the remaining language tests. The relatively complex mixture here may of course have something to do with the heterogeneous nature of the subject group. Factor 4 is the psychomotor speed and accuracy factor corresponding to Factor 3 in the overall analysis.

With the Short test scores there are no reservations about the numbers of subjects compared with the number of tests (11), but the picture is less interesting simply because of this restricted

range of tests. Principal component factor analyses (without iteration) were carried out on each group separately (and on all three combined) first for those who completed Full testing (since these were the subjects upon which the calculation of Short factor scores was based) and then for all subjects who completed at least Short testing. Table 18 shows the factor structures resulting from these analyses, presented in a similar manner to the factor structures based on Full scores. The criteria for allocation to a given factor (in terms of magnitude of contribution in the Varimax rotated factor matrix) are rather more variable than before, and are indicated in the table.

Part (a) shows the structures in subjects who completed Full testing. In the DAT and MID groups the picture is similar to that seen in the overall analysis except that Orientation contributes particularly highly to no factor, and that a third factor emerges which has large contributions from only a couple of tests. The picture in the Other group is not neatly interpretable: here even a fourth factor has emerged, though this may reflect statistical artefact since very few of these 58 Other subjects made less than the maximum score on the simple task of identifying and naming three common objects.

Part (b) shows the structures in all subjects tested. All the analyses show a similar picture to that seen in the original analysis of Short test scores (as shown in the first column of part a of the table) though the actual numerical labels of the factors are reversed. Here however Writing name/Reading sentence always appears in the 'verbal' factor (as opposed to neither factor) and making a Square with Sticks for some reason appears in the verbal

Table 18 The factor structures on Short testing of each diagnostic group and of the three combined. Numbers indicate which factor each test is part of according to the criteria described in the text.

(a) Subjects who completed Full testing

	All (174)	DAT (58)	MID (58)	Other (58)
(Criterion level)	.50	.52	.64	.58
ORIENTATION	2	-	-	1
DIGIT SPAN	2	2	2	1
SENTENCE REPETITION	2	23	2	1
AUTOMATIC SPEECH	2	2	2	1
Identifying SKP	2	3	3	4
Write Name/Read Sentence	-	2	3	3
SPIRAL MAZE 1	1	1	1	2
SPIRAL MAZE 2	1	1	1	2
BOX-FILLING	1	1	1	2
PORTEUS MAZES .	1	1	1	1
SQUARE WITH STICKS	1	1	1	3
Percentage of the variance explained				
Factor 1	31.9	30.4	33.8	26.8
Factor 2	27.5	21.8	21.1	25.4
Factor 3		18.2	14.6	14.1
Factor 4				10.6
Total	59.4	70.4	69.5	77.0

(b) Subjects who completed at least Short testing.

	All (215)	DAT (74)	MID (74)	Other (67)
(Criterion level)	.61	.58	.58	.59
ORIENTATION	1	2	1	1
DIGIT SPAN	1	1	1	1
SENTENCE REPETITION	1	1	1	1
AUTOMATIC SPEECH	1	1	1	1
Identifying SKP	1	1	1	1
Write Name/Read Sentence	1	1	1	1
SPIRAL MAZE 1	2	2	2	2
SPIRAL MAZE 2	2	2	2	2
BOX-FILLING	2	2	2	2
PORTEUS MAZES .	2	2	2	2
SQUARE WITH STICKS	1	1	1	1
Percentage of the variance explained				
Factor 1	41.5	40.0	39.7	47.1
Factor 2	34.7	34.6	36.1	33.1
Total	76.2	74.6	75.9	80.2

factor as opposed to the 'performance' one. In the MID group Orientation has just crept into the so-called performance factor rather than the verbal one. The reason behind the shift of Square with Sticks is unclear, but may be connected with the addition in this analysis of subjects who were very poor on all aspects of testing and consequent statistical bias.

The analyses suggest minor differences in factor structures in different groups, but these differences are not striking enough to suggest that the analyses of factor scores based on the original factor analyses are invalid or distorted.

CHAPTER 3

Decline Over Time.

Subjects

As many subjects as possible were retested 10 months after the first assessment. The various reasons for entirely failing to retest given subjects are detailed in Tables 1 and 2. Originally the planned test-retest interval had been a year: owing to delays in starting data collection attributable to the South Lothian Ethics Committee it was decided to reduce this interval, since the study was a time-limited one. Even with this reduction, the proportion of patients retested is disappointingly low. Further reduction of the interval was considered unwise in view of previous studies finding no significant decline in test performance in dementia over relatively short periods such as 6 months, as reviewed by Gilleard (1978). The same inclusion-exclusion criteria as at first testing were used.

Table 1 Reasons for failure to retest patients given Full testing at Test 1, broken down by diagnostic category and age group.

		Dead	Phys	Refus	Geog	Time	Other	tot
DAT	-69	0	0	0	3	3	0	(6)
	70-79	0	0	0	5	2	0	(7)
	80+	3	0	1	5	1	0	(10)
tot		(3)	(0)	(1)	(13)	(6)	(0)	23
MID	-69	0	1	1	6	10	0	(18)
	70-79	1	0	0	3	6	0	(10)
	80+	0	1	0	5	6	0	(12)
tot		(1)	(2)	(1)	(14)	(22)	(0)	40
Other	-69	2	2	2	7	21	5	(39)
	70-79	0	0	0	2	1	0	(3)
	80+	0	0	0	0	0	0	(0)
tot		(2)	(2)	(2)	(9)	(22)	(5)	42

Dead-dead, Phys-too physically disabled, sight-impaired, hearing-impaired, or otherwise ill or suffering from infection for valid retesting to be possible, Refus-refused (or occasionally failed to keep an out-patient appointment without making contact), Geog-geographically inaccessible (because now in a house, part IV facility, or hospital which could not practically have been visited), Time-initially seen too late for 10-month retesting to be possible. Other-e.g. untraceable, or no point in retesting in the circumstances.

Table 2 Reasons for failure to retest patients given Short testing at Test 1, broken down by diagnostic category and age group.

		Dead	Phys	Refus	Geog	Time	Other	tot
DAT	-69	0	0	0	3	3	0	(6)
	70-79	1	0	1	5	3	0	(10)
	80+	3	1	1	9	1	0	(15)
tot		(4)	(1)	(2)	(17)	(7)	(0)	31
MID	-69	2	1	1	6	10	0	(20)
	70-79	1	0	0	6	7	0	(14)
	80+	1	1	0	7	6	0	(15)
tot		(4)	(2)	(1)	(19)	(23)	(0)	49
Other	-69	2	2	2	9	21	5	(41)
	70-79	1	1	0	3	1	0	(6)
	80+	2	0	0	0	0	0	(2)
tot		(5)	(3)	(2)	(12)	(22)	(5)	49

These figures are inclusive of those in Table 1, since Full testing incorporates Short. Abbreviations are as for Table 1.

Subjects who received only Short testing at Test 1 were given Short testing again at Test 2 (or were not retested at all). Some of the subjects who received Full testing at Test 1 again received Full testing at Test 2, while others only received Short. The reasons for retesting with only the Short version were the same as described in the previous chapter concerning initial assessment. This complicates analysis and interpretation but could not be helped.

Hence three possible combinations of test-retest combination were possible (excluding those where no retesting was carried out): Full:Full, Full:Short, and Short:Short. The numbers of subjects undergoing these combinations in the various diagnostic categories are shown in Table 3. (As usual, the numbers are cumulative from left to right since Full testing incorporates Short.) The age structure of the groups is detailed in Table 4.

Table 3 Numbers of patients who were retested, broken down by form of test, diagnostic category, age group, and sex (Male, Female in parentheses).

	Full:Full	Full:Short	Short:Short
DAT -69	7 (4, 3)	8 (4, 4)	8 (4, 4)
70-79	13 (4, 9)	16 (5,11)	21 (7,14)
80+	7 (1, 6)	11 (2, 9)	14 (2,12)
tot	(27)(9,18)	(35)(11,24)	(43)(13,30)
MID -69	5 (5, 0)	6 (5, 1)	6 (5, 1)
70-79	7 (4, 3)	8 (4, 4)	10 (5, 5)
80+	2 (0, 2)	4 (1, 3)	9 (3, 6)
tot	(14)(9, 5)	(18)(10, 8)	(25)(13,12)
Other -69	11 (4, 7)	14 (5, 9)	14 (5, 9)
70-79	1 (0, 1)	1 (0, 1)	1 (0, 1)
80+	1 (1, 0)	1 (1, 0)	3 (1, 2)
tot	(13)(5, 8)	(16)(6,10)	(18)(6,12)
Total	54 (23,31)	69 (27,42)	86 (32,54)

Table 4 Age structure of those subjects retested, broken down by form of tests and diagnostic category.

		Full:Full	Full:Short	Short:Short
DAT.	mean	74.0	74.9	75.8
	SD	(9.1)	(8.5)	(8.9)
	range	54-89	54-89	54-93
MID.	mean	69.9	71.3	74.0
	SD	(10.4)	(9.9)	(9.8)
	range	53-86	53-86	53-87
Other.	mean	59.2	60.1	62.4
	SD	(12.8)	(11.7)	(12.9)
	range	39-83	39-83	39-83

Locations were usually the same as at initial testing. Wherever possible subjects were tested at roughly the same time of day as at initial testing. The diagnoses of the Other subjects who were retested were as follows:

Full testing on both occasions: Suspected Pick's Disease (1), Huntington's Chorea (3), Senile Chorea (2), Hydrocephalus, treated (1), Steroid-Induced Dementia (1), Binswanger Encephalopathy (1), Frontal head injury with evidence of added dementing process (1), Chronic schizophrenia (1), Normal with peripheral hand tremor (1), No diagnosis reached (1).

Full testing followed by Short: allthe above plus General Paresis of the Insane (infection 20 yrs previously) (1), 'Functional' condition (Hysterical or pseudo-dementia) (1), Mental Handicap (1).

Short testing on both occasions: all the above plus Suspected mixed Alzheimer/Multi-Infarct (2).

Tables 5, 6, and 7 show characteristics of drug use in retested subjects. A proportion of subjects has experienced a change in medication, though this is not confined to particular age groups within diagnoses.

Table 5 Numbers of patients who were retested (all test categories), broken down by drug use (Test 1, Test 2), diagnostic category, and age group.

	Hypn	MinT	MajT	ADep	Two	Other	tot	None
DAT -69	1,1	0,0	1,2	2,1	0,1	0,0	(4, 5)	4, 3
70-79	1,1	1,1	6,8	0,0	1,3	0,0	(9,13)	12, 8
80+	2,2	0,0	2,3	0,0	0,0	0,0	(4, 5)	10, 9
tot	(4,4)	(1,1)	(9,13)	(2,1)	(1,4)	(0,0)	(17,23)	26,20
MID -69	1,1	0,0	1,1	0,0	0,0	0,1	(2, 3)	4, 3
70-79	1,0	0,0	3,5	1,1	0,0	0,0	(5, 6)	5, 4
80+	0,0	0,0	2,4	0,0	0,0	0,0	(2, 4)	7, 5
tot	(2,1)	(0,0)	(6,10)	(1,1)	(0,0)	(0,1)	(9,13)	16,12
Other -69	1,2	0,0	3,2	1,0	1,3	1,1	(7, 8)	7, 6
70-79	0,0	0,0	0,0	0,0	0,0	0,0	(0, 0)	1, 1
80+	1,0	0,0	0,1	0,0	0,0	0,0	(1, 1)	2, 2
tot	(2,2)	(0,0)	(3,3)	(1,0)	(1,3)	(1,1)	(8, 9)	10, 9
Total	8,7	1,1	18,26	4,2	2,7	1,2	(34,45)	52,41

Hypn-Hypnotic, MinT-Minor Tranquilliser, MajT-Major Tranquilliser, Adep-Anti-Depressant, Two-Two of the preceding categories, Other-Other psychoactive medication.

Table 6 Numbers of patients retested (Full followed by Full or Short) broken down by change in psychoactive drug use from Test 1 to Test 2 and diagnostic category.

	None:Some	Some:None	Other	No Change
DAT	6	0	2	27
MID	2	0	1	15
Other	2	1	2	11

None:Some-No psychoactive medication at Test 1 but some at Test 2, Some:None-Psychoactive medication at Test 1 but none at Test 2, Other-Other change in drug use. Same-No change in drug use.

Table 7 Numbers of patients retested (all test categories), broken down by change in psychoactive drug use from Test 1 to Test 2 and diagnostic category. Labels as for Table 7.

	None:Some	Some:None	Other	No Change
DAT	7	1	3	32
MID	4	0	1	20
Other	2	1	3	12

The representativeness of retested subjects.

Before considering the results of the second testing, consideration will be given to the representativeness of the subjects who were retested. The actual proportions and percentages of patients retested in various categories are shown in Table 8.

Table 8 Proportions of subjects retested, broken down by form of test, diagnostic category and age group.

	Full:Full	Full:Short	Short:Short
DAT -69	7/14 50%	8/14 57%	8/14 57%
70-79	13/23 57%	16/23 70%	21/31 68%
80+	7/21 33%	11/21 52%	14/29 48%
tot	(27/58 47%)	(35/58 60%)	(43/74 58%)
MID -69	5/24 21%	6/24 25%	6/26 23%
70-79	7/18 39%	8/18 44%	10/24 42%
80+	2/16 13%	4/16 25%	9/24 38%
tot	(14/58 24%)	(18/58 31%)	(25/74 34%)
Other -69	11/53 19%	14/53 26%	14/55 25%
70-79	1/ 4 25%	1/ 4 25%	1/ 7 14%
80+	1/ 1 100%	1/ 1 100%	3/ 5 60%
tot	(13/58 22%)	(16/58 28%)	(18/67) 27%)

Tables 1 and 2 showed that a common reason for failing to retest subjects was simply that the initial assessment occurred too late. One would not expect this alone to bias the sample of subjects who were retested in any important way; but some of the other reasons might plausibly have led to the retested group being 'special' or unrepresentative in some way. Hence a variety of comparisons of aspects of performance at Test 1 were made between those who were retested and those who were not. (Some pictorial idea of the relationships between these groups can be gained by comparing Graphs 1 to 16 in the previous chapter with Test 1 scores on the corresponding graphs 1 to 16 in this one: even more graphs will not be presented on this point.) Statistical comparisons were made on the basis of t tests comparing subjects who were seen again with

those who were not (rather than comparing those who were seen again with the whole Test 1 pool: this would be a weaker test of whether the retested sample was atypical, though it is the only comparison possible by looking at the graphs as suggested above). These comparisons are summarised in Table 9. All factor scores referred to in this chapter are of course calculated using the same formulae used in the previous chapter, i.e. those derived from factor analysis on the performances of those 174 subjects who completed Full testing at Test 1: no new factor analysis has been performed.

Table 9 Comparison of age, overall test scores, and factor scores at initial assessment in subjects who were retested compared with those who were not, broken down by diagnosis and then by age, Orientation score, and overall test score as well, as in previous graphs. In all three parts of the table:

R means retested subjects had significantly higher mean score (or age) at initial testing.

N means non-retested subjects had significantly higher mean score (or age) at initial testing.

- not significant by independent t-test, 2-tailed
+ significant at .10 level
* significant at .05 level
** significant at .01 level

A blank space indicates that the n of cases in one or both groups was too small to permit statistical comparison. (Numbers of cases are not shown, but can be deduced from previous tables and figures.)

(1) Subjects completing Full testing at Test 1 and at Test 2.

	DAT						MID						Other					
	Age	Full	Factor				Age	Full	Factor				Age	Full	Factor			
		1	1	2	3	4		1	1	2	3	4		1	1	2	3	4
All Ss	-	-	-	-	-	-	-	-	-	R*	-	-	-	-	-	-	-	-
Age																		
-69	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
70-79	-	-	-	-	-	-	-	-	-	R*	N+	-						
80+	-	-	-	-	-	-												
Orient																		
10-13	-	-	-	-	-	-	-	-	-	R*	-	-	-	-	-	-	-	-
7-9	-	-	-	-	-	-							-	-	-	-	-	-
0-6	N+	-	-	-	-	-												
Full																		
65%+	-	-	R+	-	-	-	-	-	-	R+	-	-	-	-	-	-	-	-
50-65%	-	-	-	-	-	-	-	-	-	R*	N*	-						
0-50%	-	-	-	-	-	-												

(2) Subjects completing Full testing at Test 1 and Short testing at Test 2.

	DAT						MID						Other					
	Age	Full	Factor				Age	Full	Factor				Age	Full	Factor			
		1	1	2	3	4		1	1	2	3	4		1	1	2	3	4
All Ss	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N*	-	R+
Age																		
-69	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N**	-	R*
70-79	-	-	-	-	-	-	-	-	-	R+	-	-						
80+	-	-	-	-	-	R*	-	-	-	-	-	-						
Orient																		
10-13	-	-	-	-	-	-	R+	-	-	-	-	-	-	-	-	-	-	-
7-9	-	-	-	-	N*	-	-	-	-	-	-	-	-	-	-	N+	-	-
0-6	-	-	-	-	-	-												
Full																		
65%+	-	-	-	-	-	-	-	-	-	-	-	-	R+	-	-	-	-	-
50-65%	-	-	-	-	-	-	-	-	-	-	N*	-						
0-50%	-	-	-	-	-	-												

(3) Subjects completing Short testing at Test 1 and at Test 2.

	DAT				MID				Other			
	Age	Sht	factor		Age	Sht	factor		Age	Sht	factor	
			2	1			2	1			2	1
All Ss	-	-	-	-	-	-	-	N*	-	-	-	-
Age												
-69	-	-	-	-	-	-	-	-	-	-	-	-
70-79	-	-	-	-	-	-	-	-				
80+	-	-	-	-	-	N*	-	N*				
Orient												
10-13	-	-	-	-	R+	-	-	-	-	-	-	-
7-9	-	-	-	N*	-	-	-	-	-	-	N+	-
0-6	-	-	-	-	-	-	-	N+				

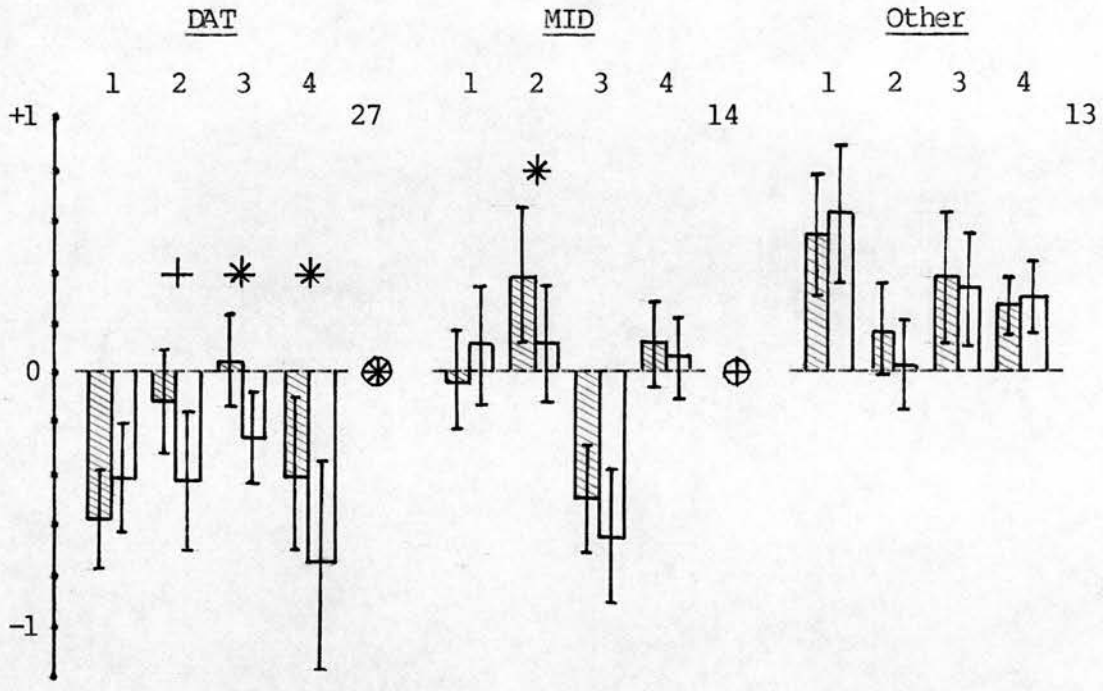
It can be seen from the tables that significant differences between subjects who were retested and those who were not are not very frequent in any of the breakdowns (by diagnosis, age, Orientation score, or overall score). The groups almost never differ on age or overall level of performance (as judged by Full or Short score). Occasionally they differ on individual factor scores: sometimes subjects who were retested have the higher initial score, sometimes those who were not. Hence patterns of performance in terms of relative levels of different factors in retested subjects may not be entirely representative of the group as a whole, but this unrepresentativeness is not marked. (In cases where subject numbers are too small for valid statistical comparison, visual inspection again generally suggests no gross differences between the groups.) It seems in general that subjects who were retested are not seriously unrepresentative of the whole subject groups. It is therefore reasonable to consider patterns and predictors of decline in these subjects.

Results and discussion.

In the following graphs, Test 1 scores are represented by hatched bars, Test 2 scores by clear bars beside the Test 1 bars. A correlated t-test was carried out for each pair of Factor scores, Test 1 versus Test 2: significant changes are shown above the Factor bars by * (.05 level of significance, 2-tailed test) or + (.10 level, 2-tailed test). A symbol encircled at the right-hand end of an abscissa indicates that the overall Full score (in Figures 1 to 4) or overall Short score (in Figures 5 to 7) for that particular group of subjects has changed significantly (again by correlated t-test, * .05 level, + .10 level, 2-tailed test).

All graphs concerning Full testing are likely to give a conservative picture of patterns of decline since very deteriorated subjects might only receive Short testing on the second occasion. Figure 1 shows mean factor scores at Test 1 and Test 2 of those subjects who completed Full testing twice, broken down by diagnostic category. The possibility was raised in the last chapter that the similar level of Factor 2 scores across groups might be an artefactual result of the tests contributing to it being insensitive, blunt instruments. The results shown in Figure 1 suggest that this is not the case. Factor 2, the so-called parietal factor, has declined significantly in the DAT and MID groups while Factor 1 (i.e. 'memory', the factor used as the main comparison factor with Factor 2 in the previous chapter) has not. This suggests that Factor 2's tests cannot be particularly insensitive. The absence of a decline in the 'memory' factor (allowing for the comments above) seems consistent with other findings of little short term decline in dementia if assessment is confined mainly to

Figure 1 Mean factor scores of patients who completed full testing twice, broken down by diagnostic category (Test 1 scores hatched, n of each group at top right).

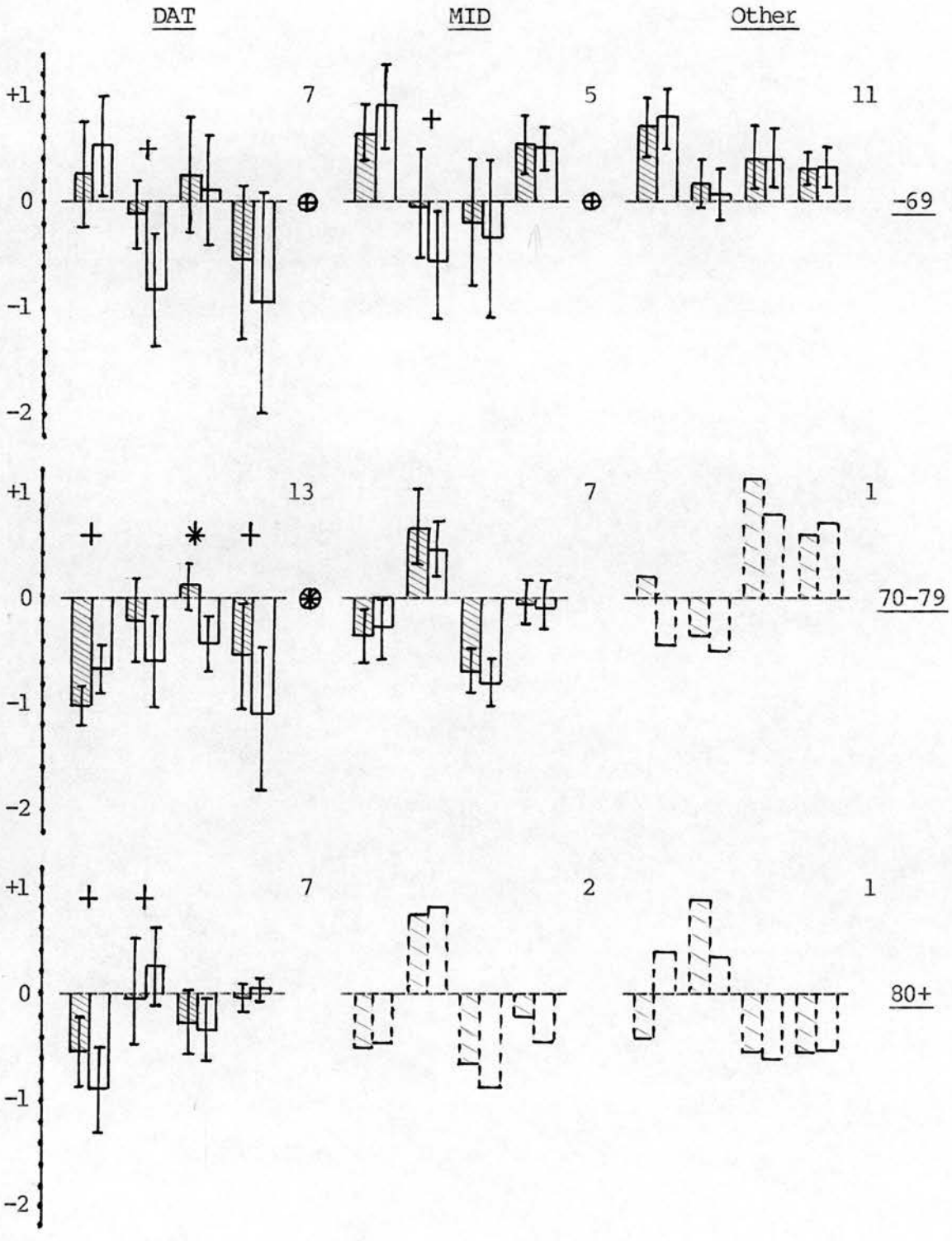


Factors are arranged in numerical order (from 1 to 4) from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

the area of memory. Here, the DAT subjects show some decline on Factors 2, 3, and 4 (i.e. parietal function, psychomotor speed and accuracy, and recognition and nominal ability), with a significant drop in overall Full score despite the lack of decline in the memory factor. The MID group show significant decline only on Factor 2 (though also on overall Full score at the .10 level). The Other group show no significant changes, which is not surprising in view of the number of cases and the variety of diagnostic conditions included.

Figure 2 shows the groups broken down by age. Here, changes are less likely to be significant because of the smaller subject numbers, but the graphs still indicate the mean magnitude of change. The youngest DAT group shows very similar changes to those seen in the DAT group as a whole, but with only the changes in Factor 2 and overall Full score being significant. The middle DAT group again shows a fairly similar pattern, but here only the drops in Factors 3 and 4 and in overall Full score are significant; Factor 1 has improved significantly. The oldest DAT group shows a different pattern: Factor 1 has declined significantly while Factor 2 has improved significantly; Factors 3 and 4 and overall Full score show no change. This might appear to suggest a different pattern of decline in younger DAT subjects compared with old, but there may be an artefact: in the oldest group a few people were not retested because they had died or were too physically ill (or were retested but managed only Short testing). These may well have been the subjects who would have shown the most decline on Full testing. Such problems were rarer in the younger groups. The oldest group may therefore be more self-selected than the younger ones and so

Figure 2 Mean factor scores of patients who completed full testing twice, broken down by diagnostic category and age group (Test 1 scores hatched, n of each group at top right).



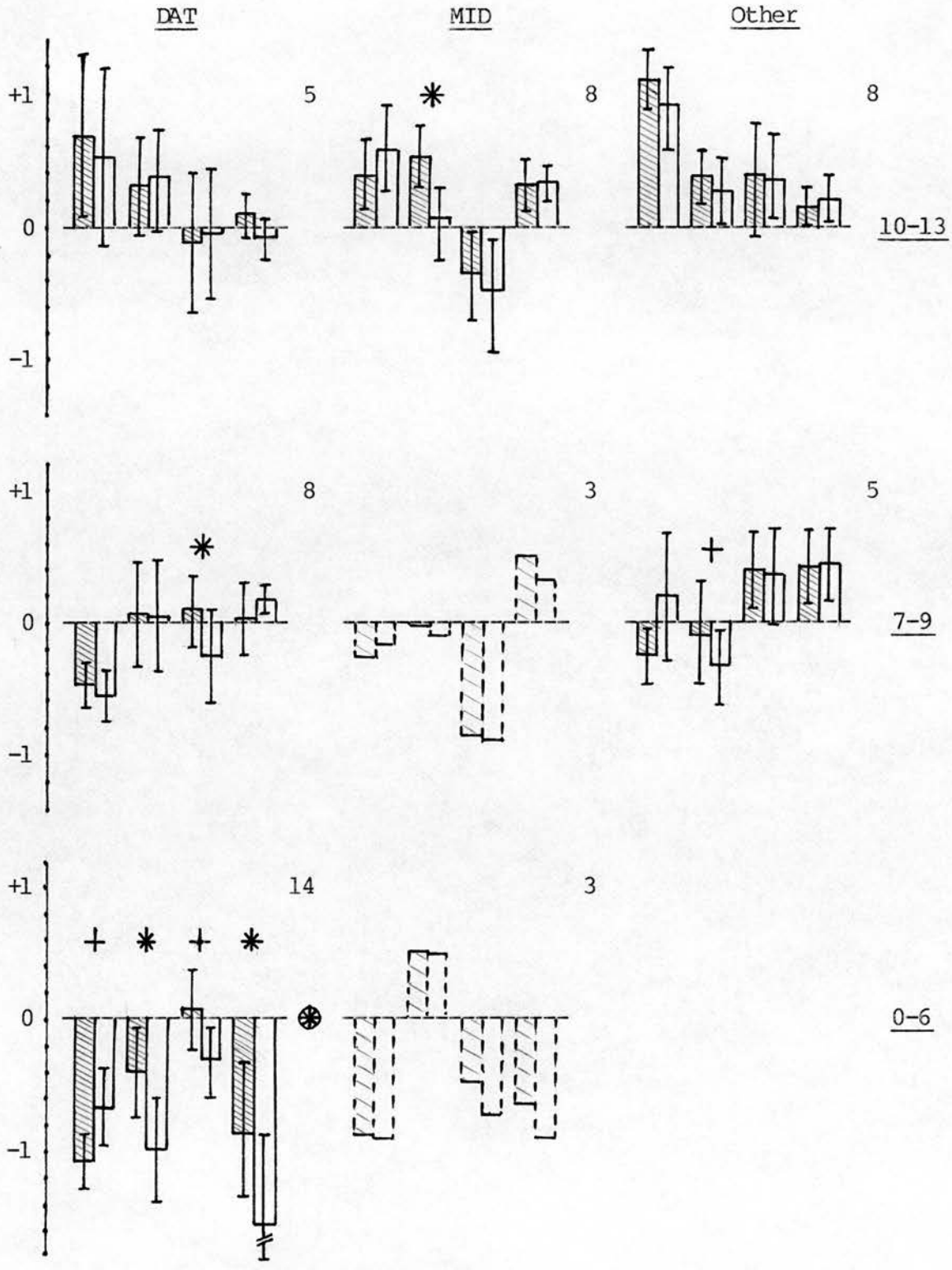
Factors 1 to 4 are arranged from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

show a misleading pattern of apparent decline. The young and middle MID groups show the same pattern of change as the MID group as a whole: slight improvement in Factor 1, slight decline in Factors 2 and 3, and no change in Factor 4. The oldest MID group is too small to base any conclusions on, and the breakdown of the Other group by age is similarly not of interest.

Figure 3 shows the groups broken down by Orientation score. The best DAT group shows no significant decline. The idea that this might be because of a ceiling effect at Test 1 is not very plausible, since the mean Test 1 scores are not very high on any factor, unless the mean scores are hiding ceiling effects in some individual subjects. The middle DAT group shows significant decline only on Factor 3. The poorest DAT group, however, shows decline on Factors 2, 3, and 4 and on overall Full score, with a significant improvement on Factor 1. Clearly no floor effect has occurred. It seems that the most impaired subjects (as gauged by Orientation score) deteriorate most strikingly. This might reflect some terminal drop phenomenon or some artefact of test scaling: there is no way of knowing whether a numerical drop of x points reflects the same degree of decline at different ends of the range. Only the best MID group is large enough to merit consideration: in it (not surprisingly as with the MID group as a whole) only Factor 2 shows significant decline. In the Other groups, only Factor 2 in the middle group declines significantly, and this is of no particular interest.

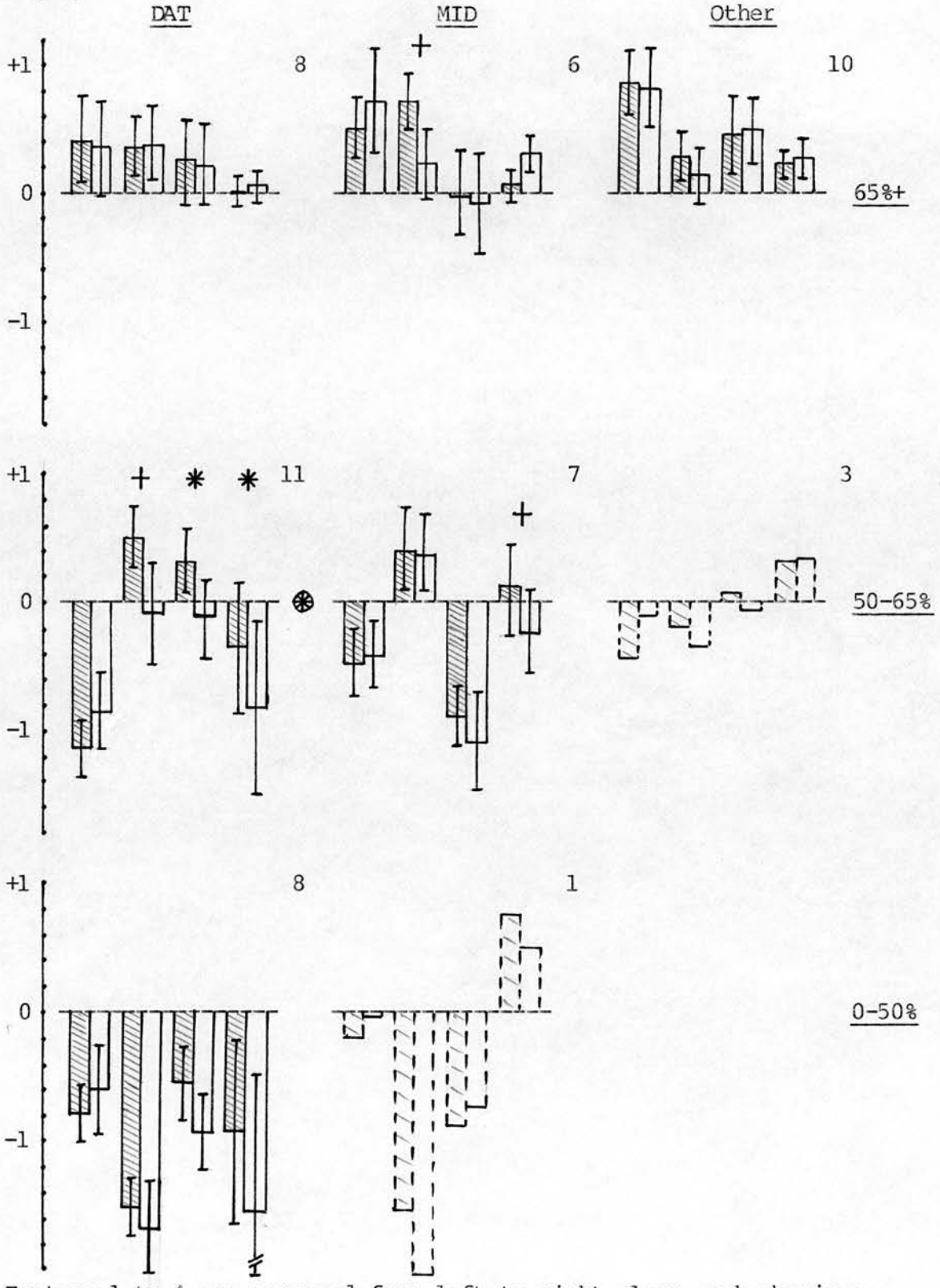
Figure 4 shows the groups broken down by overall Full score. Of the DAT groups, only the middle one shows significant drops (on Factors 2, 3, and 4 and on overall Full score itself). The best

Figure 3 Mean factor scores of patients who completed full testing twice, broken down by diagnostic category and ORIENTATION score (Test 1 scores hatched, n of each group at top right).



Factors 1 to 4 are arranged from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

Figure 4 Mean factor scores of patients who completed full testing twice, broken down by diagnostic category and mean % correct score (Test 1 scores hatched, n of each group at top right).



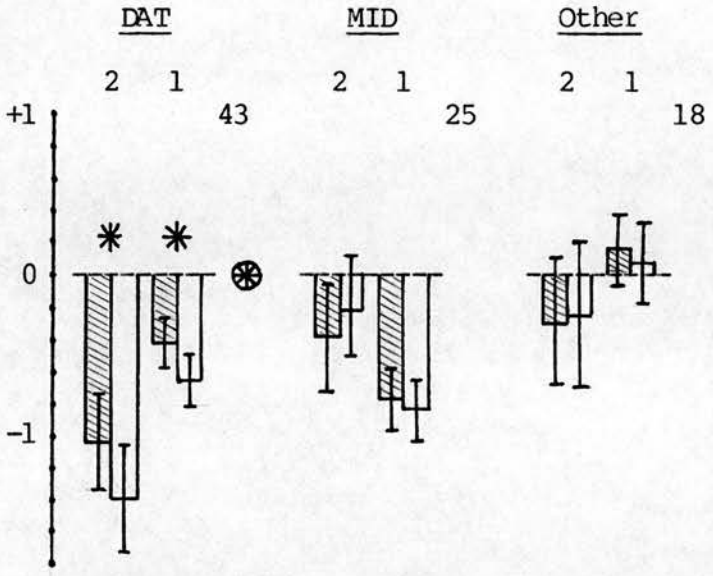
Factors 1 to 4 are arranged from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

group, as in the analysis according to Orientation score, show virtually no change; again this does not seem to reflect any ceiling effect at initial testing, but of course ceiling effects in individual subjects could be masked in the mean scores. The poorest group do show some decline in Factors 2, 3, and 4, but with no significant changes; this perhaps reflects a combination of some floor effect and the relatively small numbers of subjects in the group. The analysis of MID groups seems (as far as can be judged) to be roughly similar to the previous analysis by Orientation score. Significant drops only in Factor 2 in the best group and Factor 4 in the middle group. There is nothing of note to say about the Other group.

Figure 5 shows Short factor scores in those who received at least Short testing twice, broken down by diagnostic category. The numbers are considerably larger than in Figure 1 since a good many subjects received Full testing at Test 1 but only Short at Test 2. In the DAT group both Factor 2 (the 'verbal' factor) and Factor 1 (the 'performance' factor), and so also the overall Short score, show significant decline. Once again, decline over 10 months is shown to be demonstrable in DAT subjects. In the MID group there are no significant changes, though the verbal factor shows marginal improvement. In the Other group there are again no significant changes.

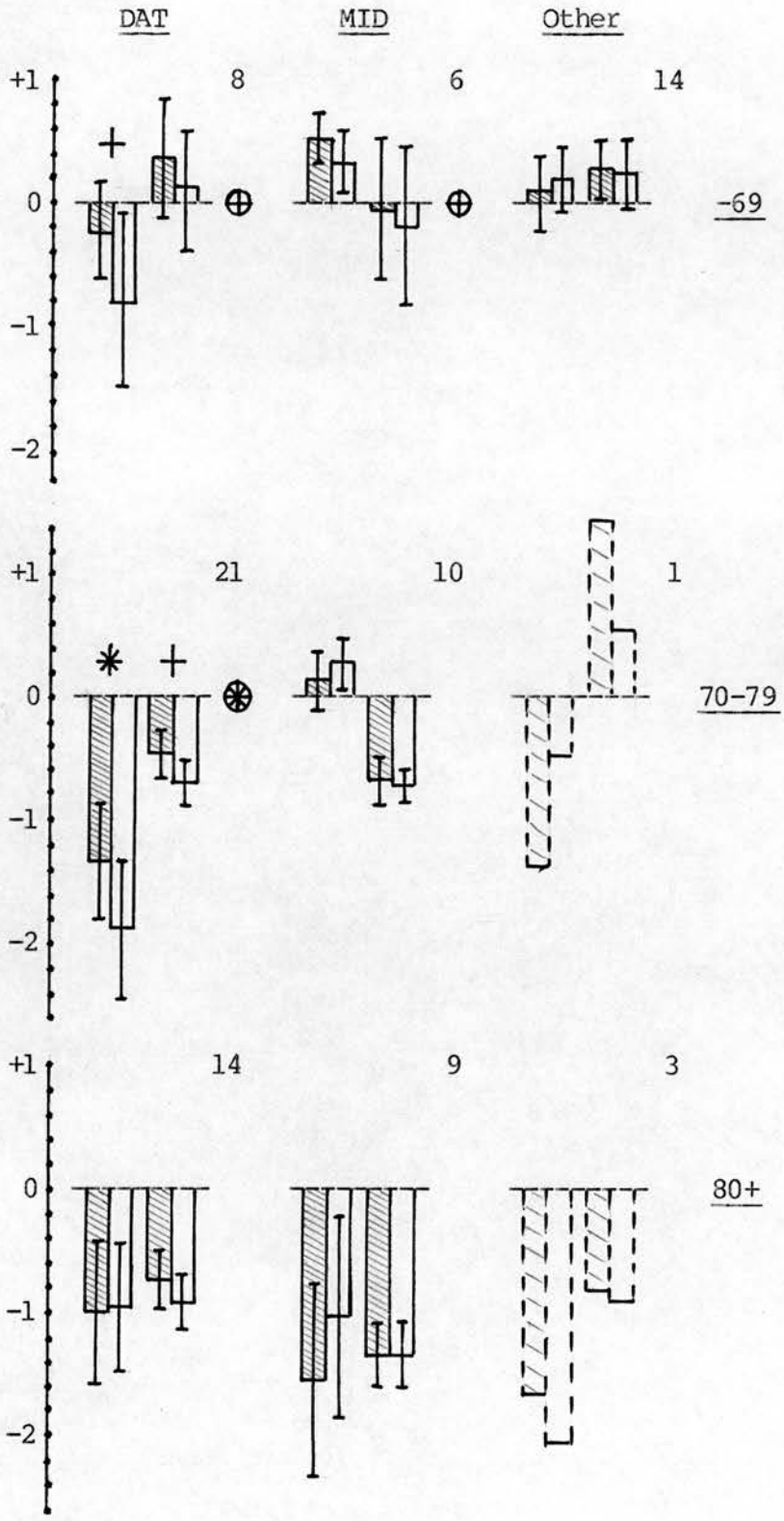
Figure 6 shows the groups broken down by age. In the youngest DAT group, only the verbal factor (and overall Short score) shows significant decline, though there is a slight fall in the performance factor. In the middle DAT group both factors show significant decline, and in the oldest group neither factor does.

Figure 5 Mean factor scores of patients who completed short testing twice, broken down by diagnostic category (Test 1 scores hatched, n of each group at top right).



Factors arranged in the order 2,1 from left to right along each abscissa; ordinate shows factor score; bars show standard errors.

Figure 6 Mean factor scores of patients who completed short testing twice, broken down by diagnostic category and age group (Test 1 scores hatched, n of each group at top right).



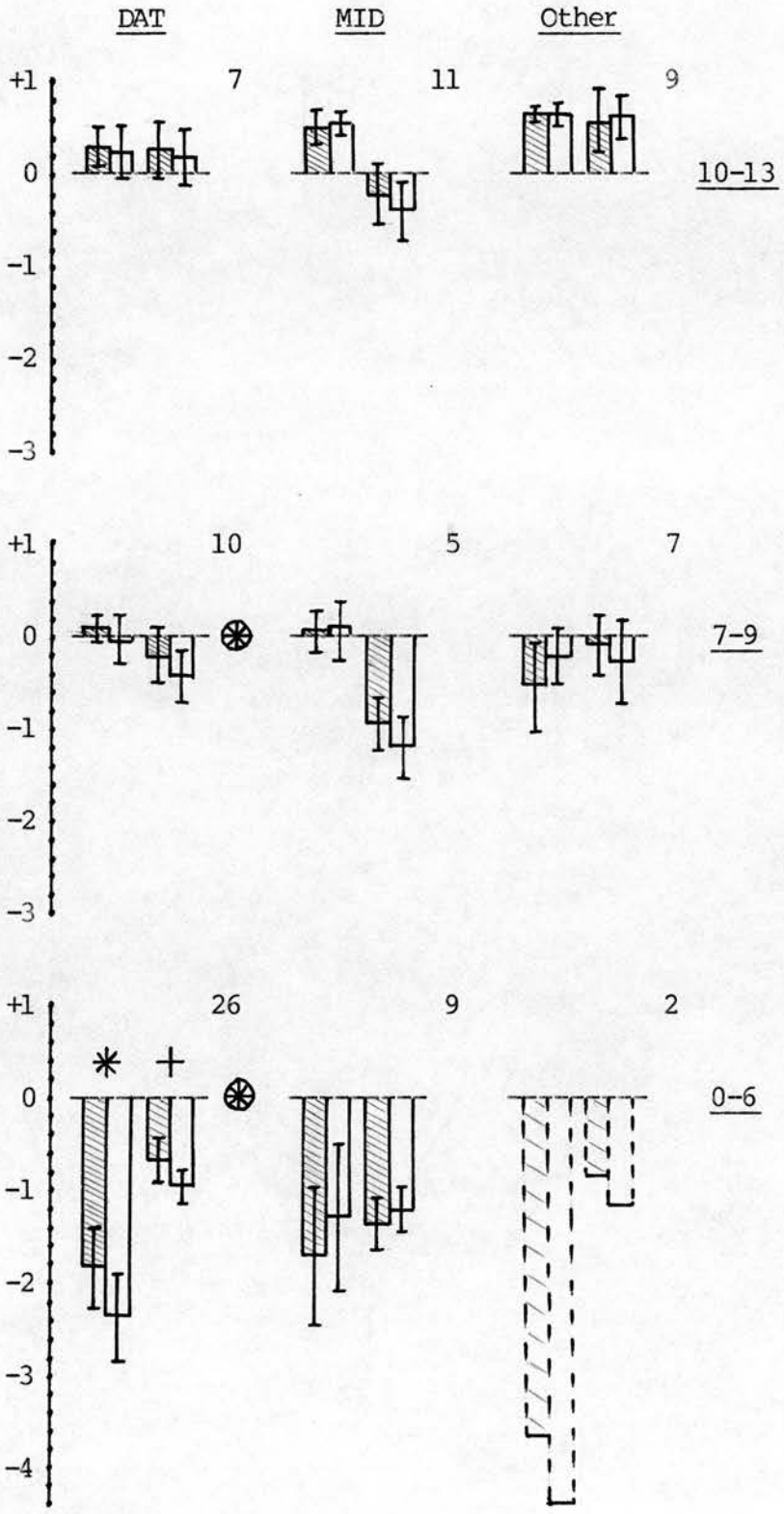
Factors 2, 1 from left to right on each abscissa; ordinate shows factor score; bars show standard errors.

This is somewhat similar to the findings with the Full factors, and the same comments apply concerning possible self-selection of subjects in the oldest group. The only significant change in the MID groups is the fall in overall Short score in the youngest group: no individual factor shows a significant change. The oldest group here is largely responsible for the slight rise in the verbal factor seen in the MID group as a whole in Figure 5. As with the DAT subjects, the oldest group here may be rather special in some way. Again the numbers in the Other group are too small to invite comment.

Figure 7 shows the diagnostic groups broken down by Orientation score. There are no significant changes in the best DAT group, again not apparently as a result of a Test 1 ceiling effect. Only overall Short score falls significantly in the middle DAT group. In the poorest group, both factors (and the overall Short score) show significant decline. Once again the tendency seems to be for the poorer subjects to show the greatest decline, perhaps reflecting terminal drop or scaling artefact as mentioned in relation to Figure 3. There are no significant changes whatever in the MID groups, nor in the two Other groups where the numbers of subjects are large enough to permit statistical test. The poorest Other group serves to illustrate that other factor graph bars are well within the possible 'limits' of test scores and not close to a floor.

As in the previous chapter, Figures 8 to 14 use circular graphs to present data on individual tests in subject groups broken down in the same way as in Figures 1 to 7, with Figure 8 corresponding to Figure 1, 9 to 2, 10 to 3, 11 to 4, 12 to 5, 13 to 6, and 14 to

Figure 7 Mean factor scores of patients who completed short testing twice, broken down by diagnostic category and ORIENTATION score (Test 1 scores hatched, n of each group at top right).

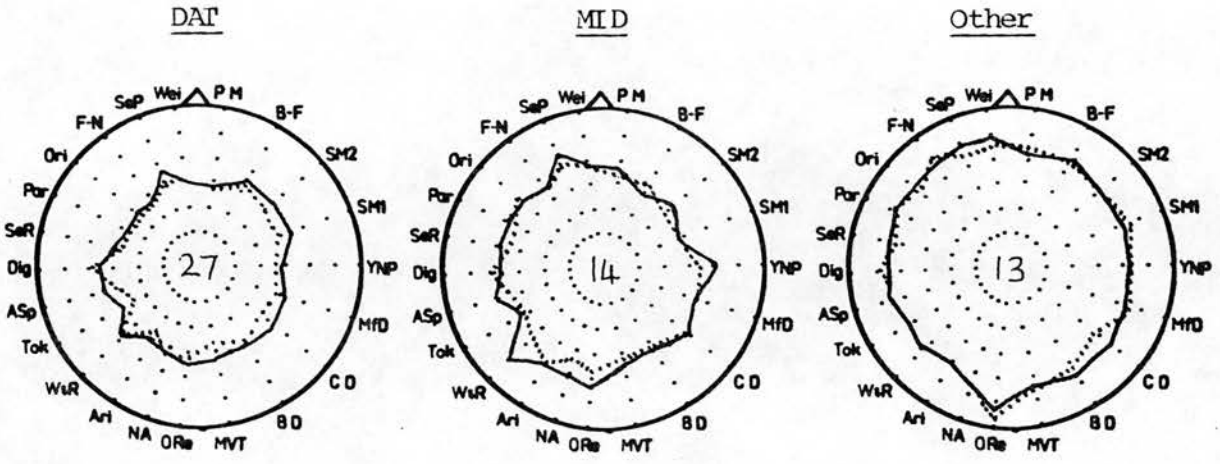


Factors 2, 1 from left to right on each abscissa; ordinate shows factor score; bars show standard errors.

7. Here, ranges of the percentile scores are ignored: mean Test 1 scores are shown in solid line, mean Test 2 scores in dotted line. (The percentile scores shown here are of course still based on the score distributions of the 174 subjects who completed Full testing at Test 1. The distributions of scores at Test 2, or at Test 1 in only those subjects who were later retested, were not taken into account. So if a score of x on a given test corresponded to a percentile score of y in the previous chapter, it still always corresponds to y here.) Again, some of the changes in scores on specific tests will be considered in later chapters concerning further analysis of certain abilities. Individual subjects' patterns of performance are shown in Appendix 5, with Test 1 scores in heavy black line and Test 2 scores, where available, in yellow or grey marking.

In general, the patterns of performance in the diagnostic groups and various subgroups appear very similar at Test 1 and Test 2. This consistency of the patterns suggests that the initial patterns of performance were reasonably reliably depicted (though this is not what this part of the study was intended to show, and is not necessarily the most desirable outcome in a study of change and decline over time). Any large changes from Test 1 to Test 2 on particular tests in any given subject group seem to have been reflected in significant changes in the factors to which the tests contribute most highly, as presented in Figures 1 to 7, and so Figures 8 to 14 will not be discussed in detail. They serve some of the same functions as did the circular percentile graphs in the previous chapter.

Figure 8 Mean percentile scores of patients who completed full testing twice, broken down by diagnostic category (Test 1 line solid, Test 2 dotted, n of each group in centre).



Key:

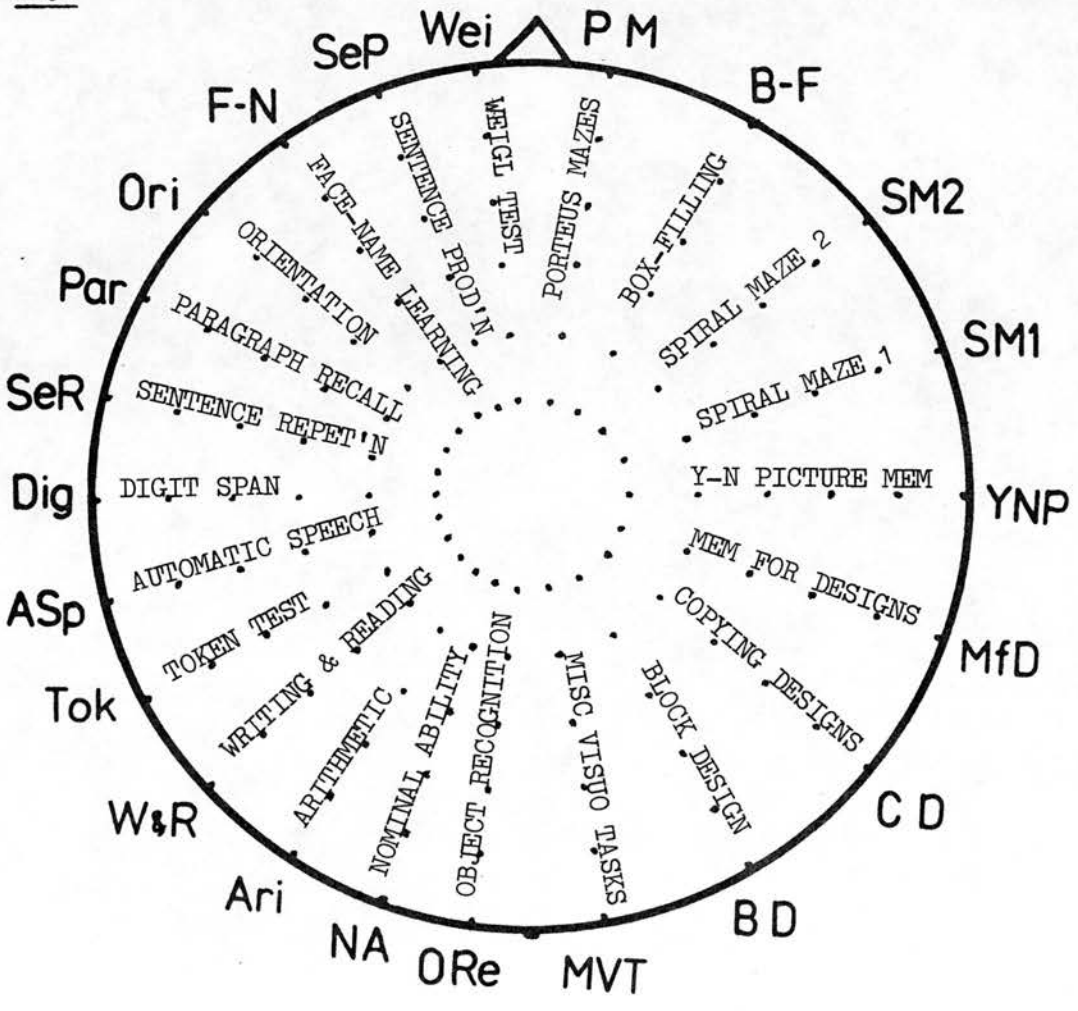


Figure 9 Mean percentile scores of patients who completed full testing twice, broken down by diagnostic category and age group (Test 1 line solid, test 2 dotted, n of each group in centre).

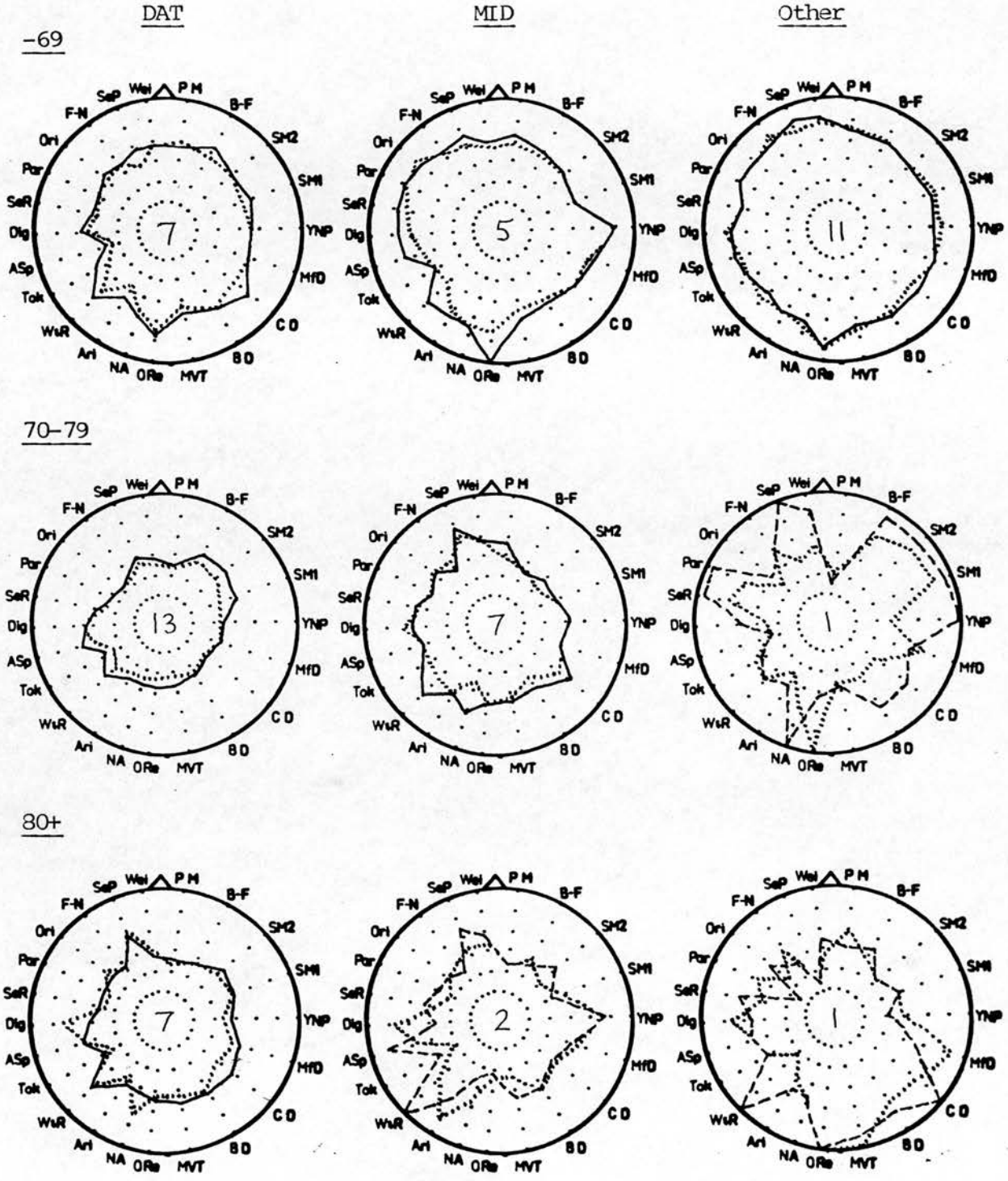


Figure 10 Mean percentile scores of patients who completed full testing twice, broken down by diagnostic category and ORIENTATION score (Test 1 line solid, Test 2 dotted, n of each group in centre).

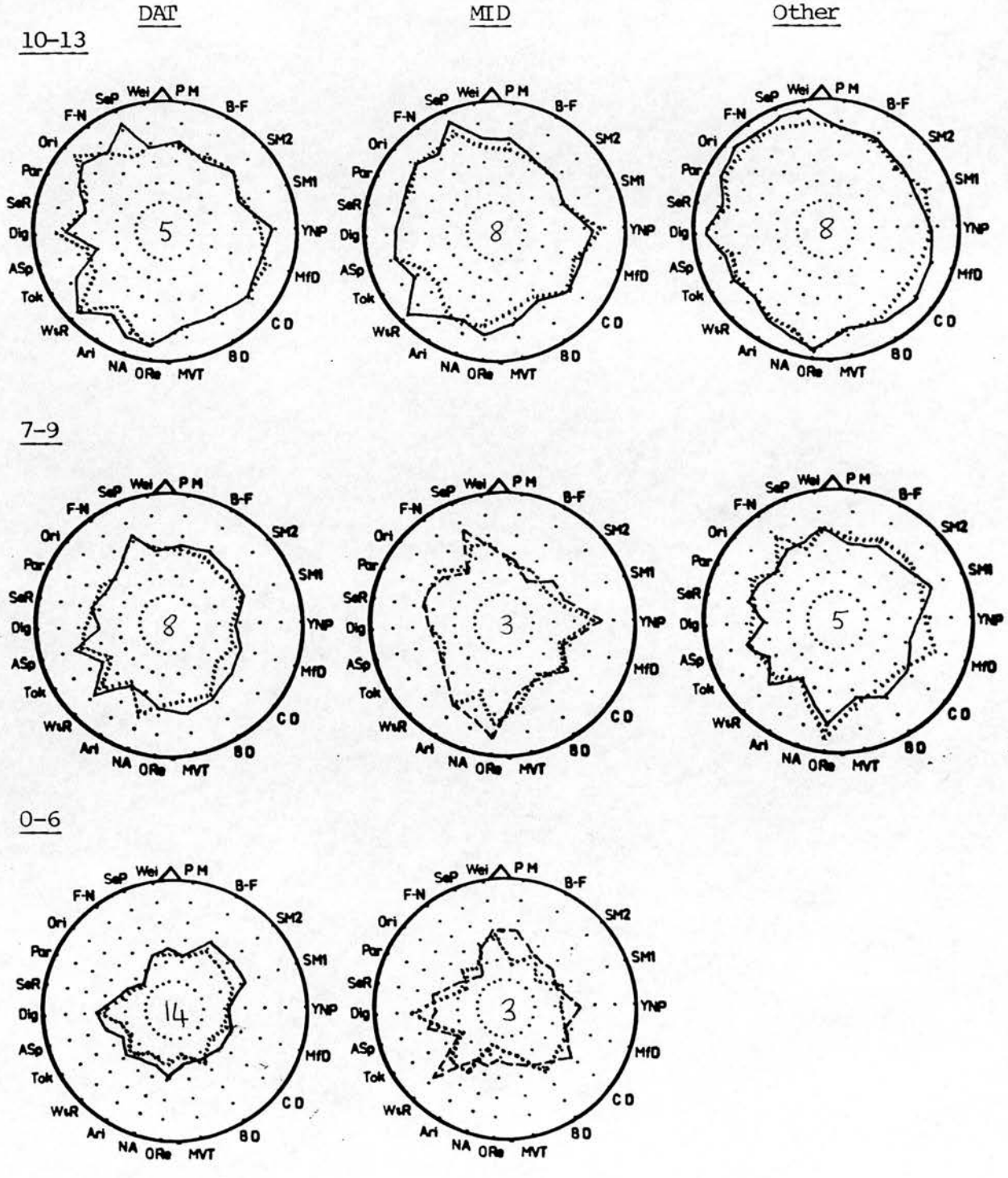


Figure 11 Mean percentile scores of patients who completed full testing twice, broken down by diagnostic category and mean % correct score (Test 1 line solid, Test 2 dotted, n of each group in centre).

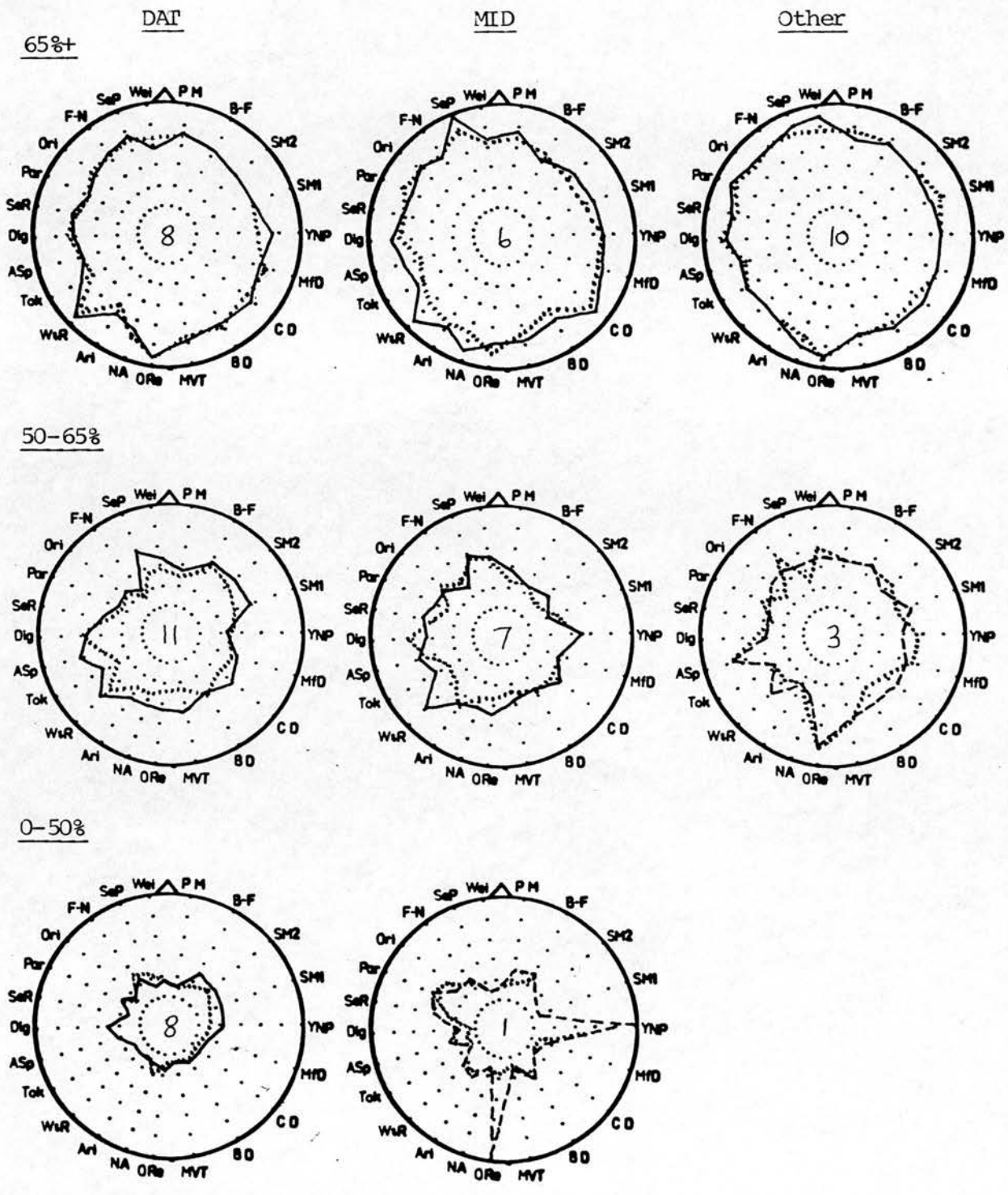
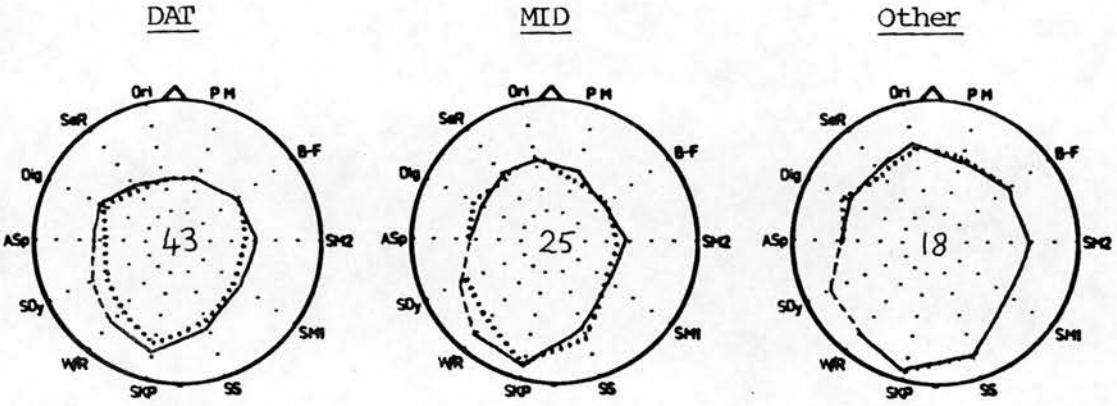


Figure 12 Mean percentile scores of patients who completed short testing twice, broken down by diagnostic category (Test 1 line solid, Test 2 dotted, n of each group in centre).



Key:

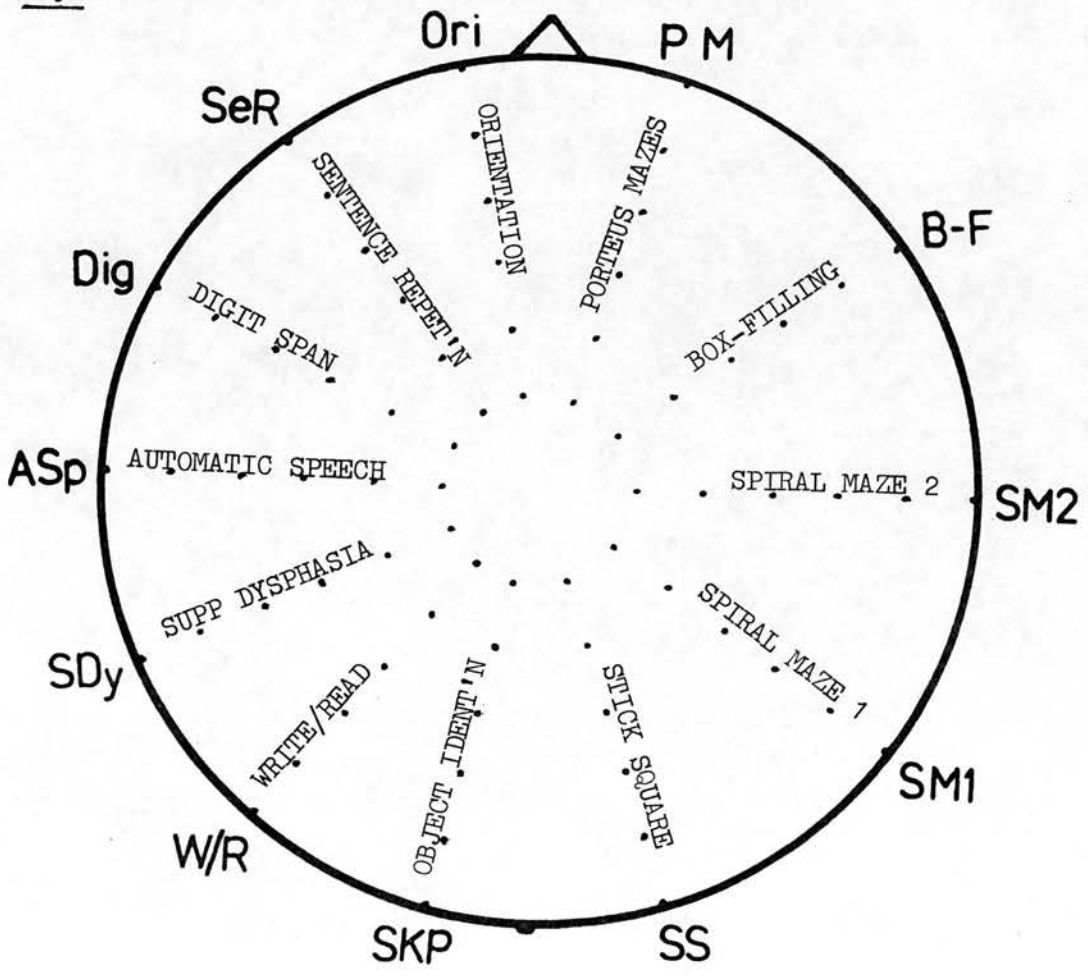


Figure 13 Mean percentile scores of patients who completed short testing twice, broken down by diagnostic category and age group (Test 1 line solid, Test 2 dotted, n of each group in centre).

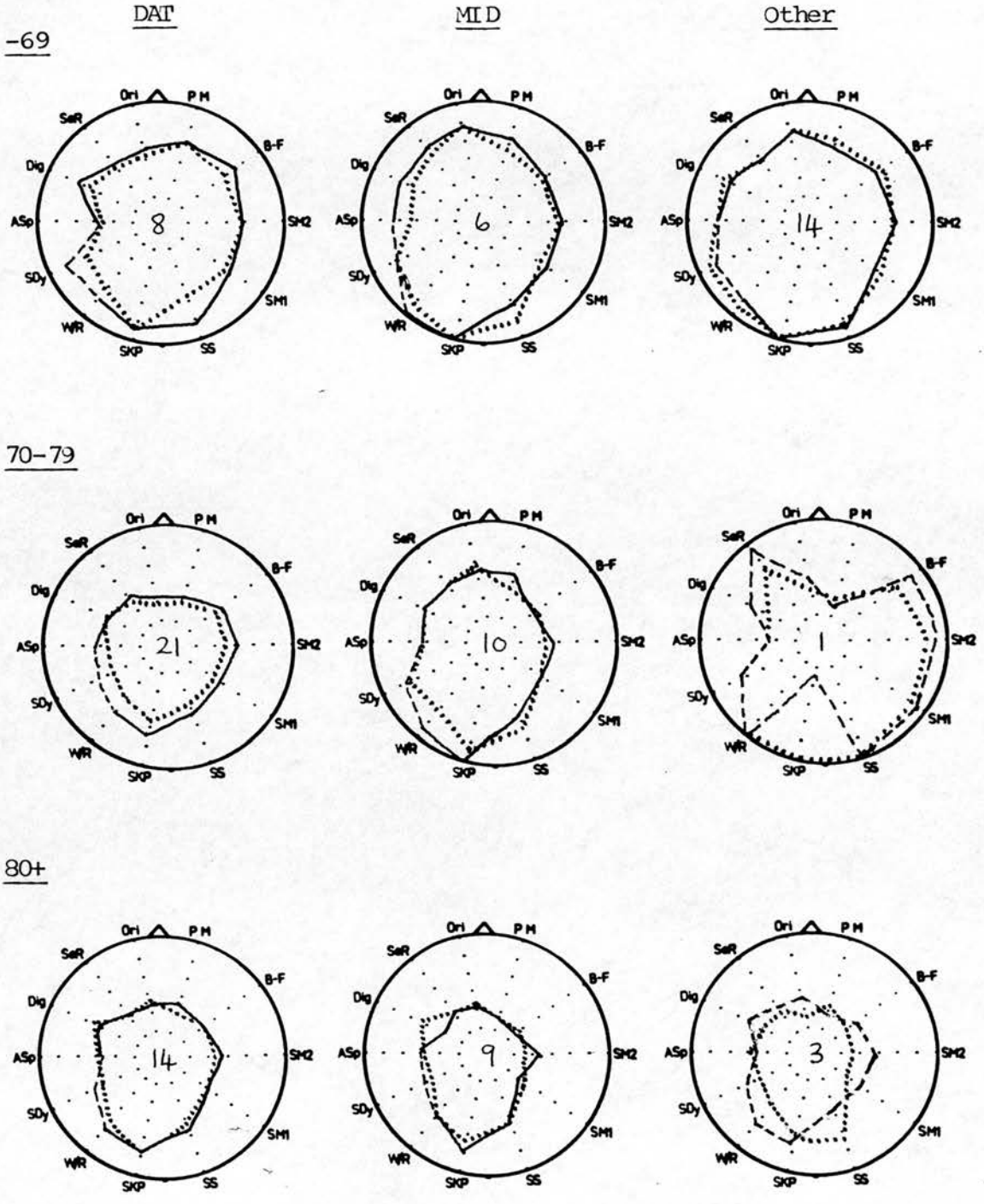
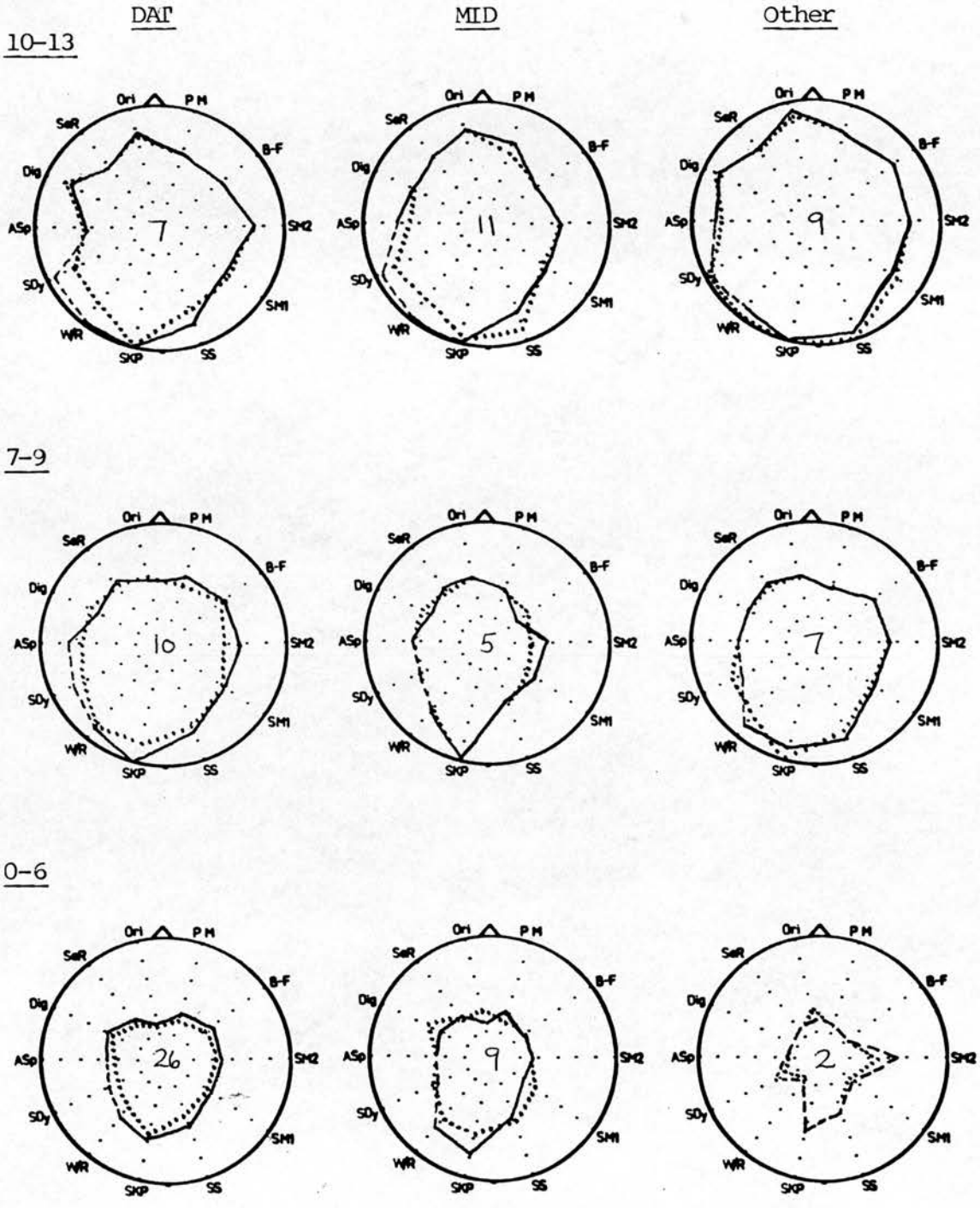


Figure 14 Mean percentile scores of patients who completed short testing twice, broken down by diagnostic category and ORIENTATION score (Test 1 line dotted, Test 2 dotted, n of each group in centre).



The prediction of decline.

It would clearly be both interesting and potentially useful if the extent of decline in test performance over the 10 month test-retest interval could be predicted from subject characteristics or aspects of performance at initial testing. Attempts to do this will now be described.

A single measure of decline in test performance over the 10 months was used for all subject groups: the overall Short score at the first testing minus the overall Short score at the second. This measure was chosen to maximise numbers of subjects (since some subjects completed Full testing followed by only Short). Full and Short scores correlate very highly with each other, and the decline in Short score correlates highly with the decline in Full score in subjects completing Full testing twice: 0.90 in 27 DAT subjects, 0.68 in 14 MID subjects, 0.74 in 13 Other subjects, and 0.83 in all 54 subjects lumped together.) Comparison of the degree of decline over time in subjects with widely varying overall levels of impairment is difficult. Problems of scaling are inherent in psychological tests and, as mentioned above, it is impossible to know whether a decline of 10 points on a given test in a high-scoring subject is 'equivalent' to a decline of 10 points in a low-scoring one. The use of proportionate decline (i.e. decline divided by initial level) would not clarify matters: it is if anything even more difficult to guess whether a 50 per cent drop from 100 to 50, i.e. 50 points, is in any way comparable to a 50 per cent drop from 10 to 5, i.e. 5 points, on some test. Such an analysis would here simply serve to magnify the apparent extent of decline in the more impaired groups and diminish it in the less impaired ones.

A variety of subject characteristics and aspects of performance at initial testing were examined in searching for predictors of the measure of decline. No meaningful relationships were expected in the Other, very mixed, subject group: analysis was carried out here since positive relationships in this Other group might suggest that similar positive relationships in the DAT or MID groups might be artefactual, as stated in the introduction. Analyses included:

Correlations with:

(a) Age, Duration of condition (where this was estimated), overall Full score, and overall Short score.

(b) Raw score on each of the tests comprising Full and Short tests.

(c) Deviation scores on each of the tests comprising Full and Short tests based on the percentile distributions of test scores as previously described in the calculation of 'focality' scores. That is, for any given subject and any given test, Deviation score = Subject's percentile score on that test minus his mean percentile score on all 22 tests in Full testing (or all 11 tests in Short testing, as appropriate). This was to see whether decline could be predicted by poor score on certain tests relative to performance on other tests (rather than simply by poor score per se on certain tests).

(d) The four Full factor scores and the two Short factor scores.

(e) Factor difference scores, i.e. Factor 1 minus Factor 2, F1-F3, F1-F4, F2-F3, F2-F4, F3-F4, Sh.factor 1 minus Sh.factor 2, to see whether decline was related to poor performance on one factor compared to another rather than simply to level of performance on any factor in isolation.

Multiple regression analyses using as independent variables:

(f) The four Full factor scores.

(g) As (f) but including age as an additional variable.

(h) The two Short factor scores.

(i) As (h) but including age as an additional variable.

(j) Raw scores on all 11 test comprising Short testing, in the DAT group only as this is the only group with a sufficiently large number of retested subjects (43) to allow multiple regression with so many independent variables. (Even in this largest group multiple regression using the 22 Full test scores would be unjustifiable.)

(k) As (j) but including age as an additional variable, again of course only in the DAT group.

The numbers of subjects upon which analyses relating to Full test variables are based are 35 in the DAT group, 18 in the MID group, and 16 in the Other group. The numbers of subjects upon which analyses relating to Short test variables are based are 43 in the DAT group, 25 in the MID group, and 18 in the Other group. The results of the analyses will now be summarised. Significant correlations are defined as Pearson product-moment correlations with a significance level of .10 or better using a 2-tailed test.

Correlations with:

(a) Age, Duration of condition (where this was estimated), overall Full score, and overall Short score. None of these variables produced significant results in the DAT or Other subject groups. In MID subjects completing Short testing at Test 1 (but not in those completing Full at Test 1) there was a significant correlation of -.36 with age and of .56 with Short, suggesting greater decline in younger subjects who performed relatively well at initial testing.

This probably reflects the inclusion in this group (of subjects completing only Short testing at test 1) of a number of very old and very impaired subjects whose initial test performance was so low that they had no 'room' to show decline at test 2. This floor effect means that the findings are probably artefacts having no real significance.

(b) Raw score on each of the tests comprising Full and Short tests. In the DAT group, 5 out of the 22 tests in Full testing showed significant correlations: these were generally reflected in those Factor scores correlating with decline, and of course one would expect a couple of significant correlations by chance. One out of the 11 tests in Short testing showed a significant correlation. In the MID group, only 3 of the 22 Full tests showed significant correlations, but 9 of the 11 Short tests did: this is reflected in the overall Short score as discussed above. In the Other group only 1 of the 22 Full tests and 1 of the 11 Short tests showed significant correlations.

(c) Deviation scores on each of the tests comprising Full and Short tests based on the percentile distributions of test scores. In the DAT group, 4 out of the 22 tests in Full testing showed significant correlations (3 of these being in the direction of decline being associated with relatively good rather than poor performance on the test in question): these were generally reflected in those Factor scores and Factor difference scores correlating with decline; and again one would expect a couple of significant correlations by chance. Three out of the 11 tests in Short testing showed a significant correlation (1 of these associating decline with relatively good rather than poor performance). In the MID group, 5

of the 22 Full tests showed significant correlations (3 of these associating decline with relatively good performance) and only 1 of the 11 Short tests did (this associating decline with poor performance). In the Other group, 3 of the 22 Full tests showed significant correlations (1 of these associating decline with relatively good performance) and none of the 11 Short tests did.

(d) The four Full factor scores and the two Short factor scores. In the DAT group, there was a significant correlation of -.43 with Factor 1 and of .31 with Factor 3, suggesting greater decline in subjects performing badly on the 'memory' factor and in subjects performing well on the 'psychomotor' factor. This is reflected in the factor difference scores described below. In the MID group, there was a significant correlation of .53 with Sh.factor 1 and of .47 with Sh.factor 2, reflecting the findings with overall Short score as described above. There were no significant correlations in the Other group.

(e) Factor difference scores. In the DAT group, there was a significant correlation of -.47 with F1-F3, suggesting greater decline in subjects performing badly on the 'memory' factor relative to the 'psychomotor' one; and there was a significant correlation of -.29 with F2-F3, suggesting greater decline in subjects performing badly on the 'parietal' factor relative to the 'psychomotor' one. There was no significant relationship with what seemed from previous analyses to be the most interesting variable, i.e. F1-F2, which concerns the relative levels of performance on the 'memory' and 'parietal' factors. The significant results can only be taken to suggest some particularly nasty form of dementia if one decides after the event that there is something special

about 'psychomotor' performance, i.e. that it deserves to be treated as an important or absolute comparator. There is no real reason to do this. There is also a possibility of some artefact: subjects who initially perform well on the psychomotor tasks may have more 'room' to subsequently decline in view of the considerable contribution of the psychomotor tasks to the overall Short score, the basis of the decline measure. Inspection of scores suggested that this was not likely to be a major influence. There were no significant correlations in the MID group. In the Other group there happened to be a significant correlation of .47 with the Short factor difference score (Sh.fac.1-Sh.fac.2) indicating greater decline in subjects doing well on the 'performance' factor relative to the 'verbal' one: this is of no importance given the heterogeneous nature of the group.

Multiple regression analyses.

The standard regression method as provided by SPSS was used rather than the hierarchical method since no causal relationship between independent variables can be assumed. The tables below show values of R squared, the percentage of the variance accounted for by the multiple regression equation emerging from the analysis. (Results of multiple regression analyses on the DAT and MID groups combined are also shown: this was performed just in case there was any general predictability in an undifferentiated group of subjects with dementia.) All available variables were used in each equation except where indicated:

age: age not included in the multiple regression equation.
F1: Factor 1 not included in the multiple regression equation.
pm: Porteus Mazes not included in the multiple regression equation.

Multiple regression analyses using as independent variables:

(f) The four Full factor scores.

	n	R sq	
DAT	35	.28	
MID	18	.30	F1
Other	16	.26	
DAT+MID	53	.29	

(g) As (f) plus age.

	n	R sq	
DAT	35	.33	
MID	18	.30	age, F1
Other	16	.27	
DAT+MID	53	.33	

(h) The two Short factor scores.

	n	R sq	
DAT	43	.05	
MID	25	.33	
Other	18	.22	
DAT+MID	68	.13	

(i) As (h) plus age.

	n	R sq	
DAT	43	.07	
MID	25	.33	
Other	18	.22	age
DAT+MID	68	.13	

(j) Raw scores on all 11 tests comprising Short testing.

	n	R sq	
DAT	43	.45	
DAT+MID	68	.36	pm

(k) As (j) plus age.

	n	R sq	
DAT	43	.48	
DAT+MID	68	.38	pm

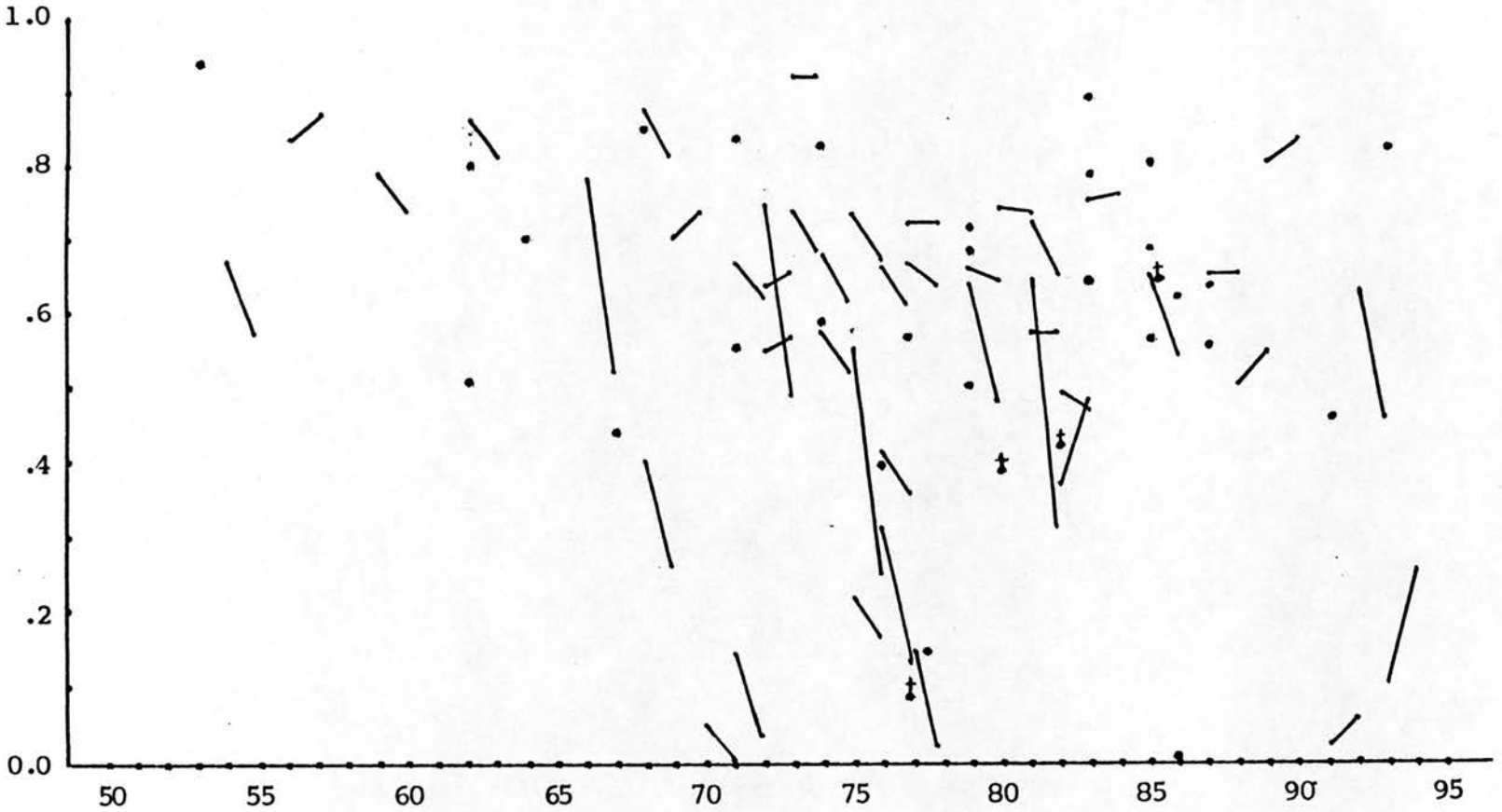
The best predictive equation for the DAT group involves the raw scores on all 11 tests comprising Short testing, plus age. This equation accounts for almost half the variance in this sample, though this is of course not to say that it would do the same in another sample. In the MID group, all equations give rather low and similar levels of prediction, but here subject numbers were too small to allow analysis using the individual tests even in Short testing. In both groups it is possible that a higher level of prediction might have been achieved had subject numbers been large enough to allow analysis on all 22 individual tests in Full testing (perhaps with the addition of age). In the Other group, all equations produced again give similar low levels of prediction:

here one would not expect any great increase in level had more detailed analysis using individual tests been possible. Levels of prediction in the combined DAT and MID group are generally in between those in each group alone: there may be some predictability in such an undifferentiated group, but there are probably predictors peculiar to each separate group.

It seems that decline over 10 months is to some extent predictable from aspects of performance at initial testing but that the level of prediction achieved in this study is not sufficiently high to be of clinical use in the individual case. A larger study including more subjects and a broader range of subject characteristics might produce results which were potentially more useful, but might not: a substantial proportion of the variance is always likely to remain unexplained.

Figures 15 and 16 show Short scores at Test 1 and Test 2 plotted for the patients who were seen twice. These are not intended to be considered in detail, but they give some idea of the considerable variability in degree of decline over 10 months. Figure 15 shows DAT patients, Figure 16 MID and Other. Most subjects have changed little, showing no more than slight deterioration or even improvement. A few, shown by the long downward-sweeping lines, have become much worse. A couple of MID subjects show remarkable improvement: this may reflect their having had a small stroke not long before Test 1 and subsequently recovering partially. One of the DAT subjects also shows a considerable improvement: this may reflect the presence of an unrecognised infection at Test 1 or perhaps mis-classification of the subject. Possible floor and ceiling effects are apparent in very few of the

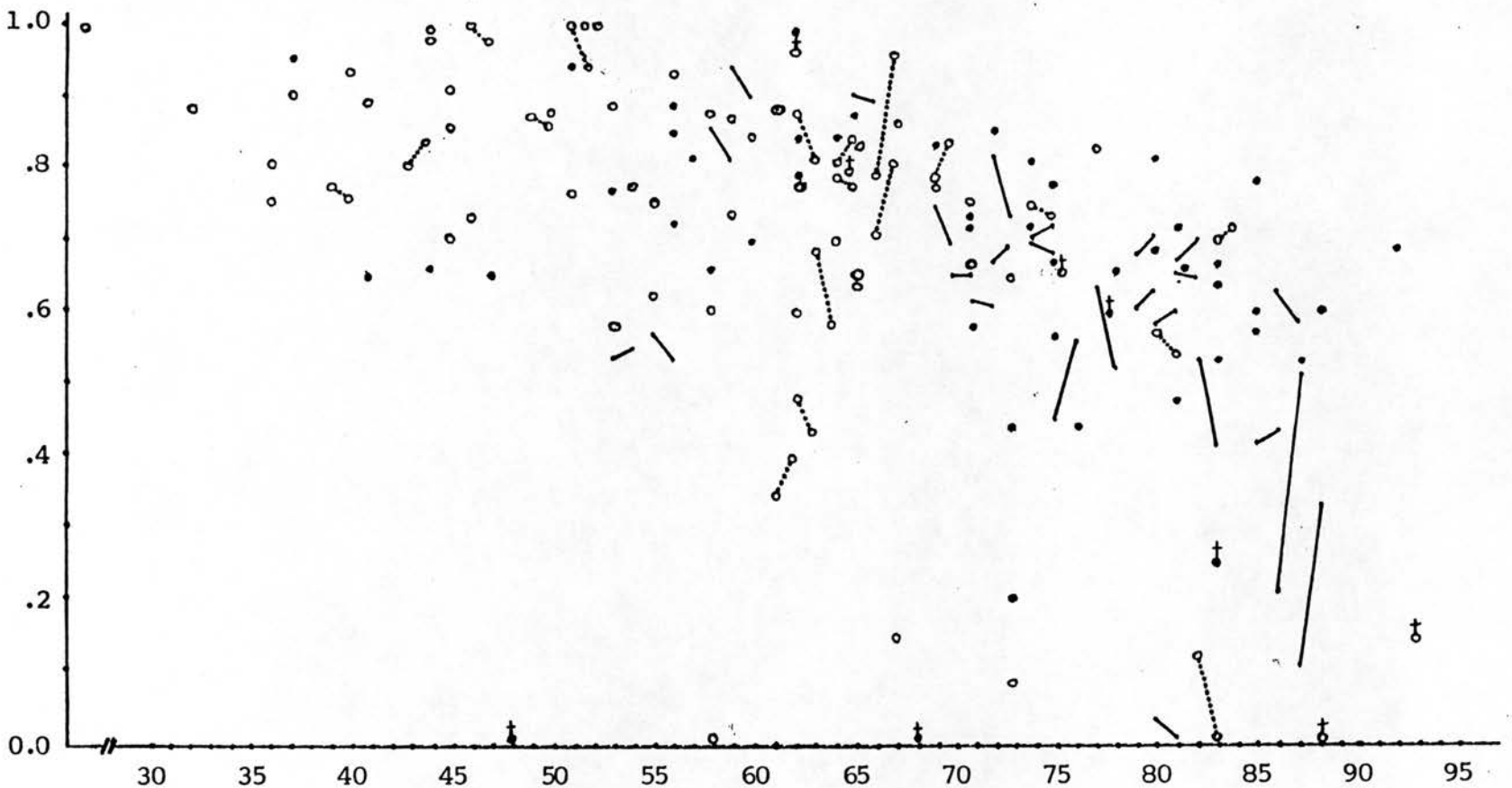
Figure 15 Propor. correct scores on short testing(s) of individuals with DAT, according to age at first testing.



Abscissa shows age at first testing, ordinate shows % score on short testing(s).

—— tested twice; • tested once only; † died before test 2.

Figure 16 Propor. correct scores on short testing(s) of individuals with MID or Other diagnoses, according to age at first testing.



Abscissa shows age at first testing, ordinate shows % score on short testing(s).

•—— MID; ◦····· Other; † died before test 2.

subjects retested. It may be, in view of ideas about terminal drop, that the subjects showing marked deterioration over time are getting close to death. At some stage it is hoped to collect data on subsequent mortality (and attempt to identify predictors of this in subjects' performances on testing, though this again is unlikely to produce results clear enough to make individual clinical predictions). Unfortunately reliable information about family history was not collected, and so the influence of genetic factors cannot be assessed.

Changes in focality.

Changes in focality scores, as defined in the previous chapter, will now be considered. Table 10 shows mean focality scores based on Full testing at Test 1 and Test 2 in those subjects completing Full testing twice, broken down by diagnostic category.

Table 10 Mean focality scores based on Full testing at Test 1 and Test 2 in subjects who completed Full testing twice, broken down by diagnostic category. (Standard deviations in parentheses.)

	Test 1	Test 2
DAT (n=27)	17.2 (5.6)	16.8 (6.6)
MID (n=14)	17.3 (5.1)	16.8 (4.4)
Other (n=13)	16.1 (6.2)	15.4 (5.6)

By correlated t-test, none of the three groups shows a significant change in focality. Table 11 shows the mean focality score based on Short testing at Test 1 and Test 2 of those subjects completing at least Short testing twice, broken down by diagnostic category.

Table 11 Mean focality scores based on Short testing at Test 1 and Test 2 in subjects who completed at least Short testing twice, broken down by diagnostic category. (Standard deviations in parentheses.)

	Test 1		Test 2	
DAT (n=43)	20.3	(9.2)	19.8	(9.4)
MID (n=25)	21.7	(8.1)	21.3	(7.3)
Other (n=18)	20.1	(9.4)	18.6	(8.7)

Again none of the three groups shows a significant change in focality by correlated t-test. Test-retest correlations on these focality measures are shown in Table 12. The measures seem to be reasonably reliable considering as usual the long test-retest interval (though again this might reflect a similar degree of artefact operating on both occasions).

Table 12 Test-retest correlations on the Full and Short focality measures in each diagnostic group and in the three combined. (Pearson product-moment correlations.)

	Full focality measure		Short focality measure	
	r	(n)	r	(n)
DAT	.69	(27)	.79	(43)
MID	.79	(14)	.68	(25)
Other	.75	(13)	.81	(18)
All	.72	(54)	.77	(86)

Another way of considering whether impairment becomes less focal or more global, generalised, or undifferentiated with the passage of time and presumed progress of the condition is to look at mean inter-test correlations at Test 1 and Test 2. Table 13 shows these means.

Table 13 Mean inter-test correlations (Pearson product-moment) for subjects completing Full testing at Test 1 and Test 2 and for those completing at least Short testing, broken down by diagnostic group. (Standard deviations in parentheses.)

	x Full				y Short			
	Test 1		Test 2		Test 1		Test 2	
DAT	.48	(.20)	.50	(.20)	.62	(.11)	.58	(.14)
MID	.46	(.23)	.45	(.25)	.60	(.15)	.56	(.16)
Other	.42	(.25)	.47	(.23)	.59	(.18)	.67	(.12)

x number of subjects are as given in previous tables, but number of correlations on which each mean is based = 231, i.e. all intercorrelations between 22 tests excluding redundant duplicate correlations.

y numbers of subjects are as given in previous tables, but number of correlations on which each mean is based = 55, i.e. all intercorrelations between 11 tests excluding redundant duplicate correlations.

By independent t-test on each pair of means shown in Table 13, no group shows a significant change in mean inter-test correlation from Test 1 to Test 2. Taken with the focality data, these results provide no evidence of de-differentiation of cognitive abilities over time in dementia.

Test-retest reliability.

Table 14 shows test-retest correlations, Test 1 vs Test 2 ten months later, for each individual test, Full and Short scores, and factor scores.

Table 14 Test-retest correlations, Test 1 vs Test 2 ten months later, for each individual test, Full and Short scores, and factor scores. (All diagnostic groups combined; Pearson product-moment correlations.)

	r	
FULL	.93	(n=54)
Factor 1	.82	
ORIENTATION	.91	
PARAGRAPH RECALL	.82	

MEMORY FOR DESIGNS	.84	
YES-NO PICTURE MEMORY	.77	
FACE-NAME LEARNING	.87	
PORTEUS MAZES	.86	
BLOCK DESIGN	.82	
WEIGL TEST	.76	
Factor 2	.82	
DIGIT SPAN	.66	
SENTENCE REPETITION	.82	
TOKEN TEST	.91	
SENTENCE PRODUCTION	.76	
WRITING AND READING	.70	
AUTOMATIC SPEECH	.67	
COPYING DESIGNS	.87	
MISC. VISUO. TASKS	.86	
ARITHMETIC	.92	
Factor 3	.82	
SPIRAL MAZE 1	.66	
SPIRAL MAZE 2	.79	
BOX-FILLING	.83	
Factor 4	.92	
OBJECT RECOGNITION	.89	
NOMINAL ABILITY	.91	
SHORT	.92	(n=86)
Short factor 2	.89	
ORIENTATION	.90	
DIGIT SPAN	.71	
SENTENCE REPETITION	.85	
AUTOMATIC SPEECH	.76	
Identifying SKP*	.81	
Writing Name/Reading Sentence	.83	
Short factor 1	.85	
SPIRAL MAZE 1	.75	
SPIRAL MAZE 2	.82	
BOX-FILLING	.88	
PORTEUS MAZES .	.86	
SQUARE WITH STICKS	.61	
(Supplementary Dysphasia	.72	n=31)

All these correlations are significant at better than the .001 level (2-tailed test), and most are fairly high in absolute terms. The lowest correlation in the tests comprising Full testing is .66 for the 'unspotted' Spiral Maze 1 and for Digit Span. In the tests comprising Short testing, Square with Sticks shows the poorest correlation at .61. Considering the length of time subjects had in which to actually change one way or another, this seems to indicate reasonably good test-retest reliability. Presumably the correlations would have been at least as good if a deliberate reliability study had been done with such numbers over a more usual period of time. Such reliability may not be wholly desirable as regards the study of change over time, but it does increase confidence that the patterns of performance found in the various groups at initial testing are reliable.

For practical purposes it is perhaps worth recording which individual tests showed significant decline over time in those subjects who were tested twice: in other words which tests appear to be most sensitive to deterioration. Table 15 summarises this information for each diagnostic group and for a combined DAT and MID group, i.e. a general 'dementia' group. The first part refers to subjects completing Full testing twice, the second to those completing at least Short testing twice.

As suggested by the previous factor graphs, memory tests (including several similar to traditionally used ones) do not seem to be particularly sensitive to decline. This may be partially explicable in terms of floor effects in some subjects on some of the memory tests: if a subject scores close to zero at initial testing then he cannot show much decline subsequently on that test.

Table 15 Significance levels of the drop in performance from Test 1 to Test 2 by correlated t-test in each diagnostic group and in the DAT and MID subjects combined. (+ significant at .10 level, 2-tailed test, * .05 level, ** .01 level, *** .001 level; imp = significant improvement at .10 level.)

Full Testing	DAT (n=27)	MID (n=14)	DAT+MID (n=41)	Other (n=13)
ORIENTATION		imp		
PARAGRAPH RECALL				
MEMORY FOR DESIGNS				
YES-NO PICTURE MEMORY			+	
FACE-NAME LEARNING				
PORTEUS MAZES		+	*	
BLOCK DESIGN				
WEIGL TEST				
DIGIT SPAN				
SENTENCE REPETITION	+		*	
TOKEN TEST	***	*	***	
SENTENCE PRODUCTION	+	+	*	
WRITING AND READING	*		**	
AUTOMATIC SPEECH				
COPYING DESIGNS	*		**	*
MISC. VISUO. TASKS	**		***	
ARITHMETIC	**		**	
SPIRAL MAZE 1	+			
SPIRAL MAZE 2	+	+	*	
BOX-FILLING	**		+	
OBJECT RECOGNITION	*		*	
NOMINAL ABILITY			+	

Short Testing	DAT (n=43)	MID (n=25)	MID (n=68)	Other (n=18)
ORIENTATION	+			
DIGIT SPAN		imp		
SENTENCE REPETITION	*			
AUTOMATIC SPEECH	*			
Identifying SKP		+	*	
Writing name/Reading sentence	*	*	**	
SPIRAL MAZE 1	*		*	
SPIRAL MAZE 2	*	*	**	
BOX-FILLING	**			
PORTEUS MAZES	*	*	**	
Square with Sticks	*	imp		

Group means, however, indicate that floor effects could only have occurred in some subjects: the means are well above the floor. The most change-sensitive tests according to the crude criteria used in the table would seem to be things like Arithmetic, Miscellaneous Visuo Tasks and especially the Token Test. The value of the Token Test will be returned to in a later chapter. These results could have implications for clinical practice as regards which tests might be most appropriate to use in judging whether progressive deterioration is occurring. It must be borne in mind, however, that this study concerned cases of established dementia: in dubious early cases where diagnosis is in question, retesting with memory tests over a period of time may still be the most sensitive way of detecting decline. Tests such as those in Miscellaneous Visuo Tasks are easy for unimpaired people, and it is likely that they will show no decline until a moderate degree of impairment is reached (i.e. until the subject drops below the low ceiling).

CHAPTER 4

Relationships Between Test Performance and Real Life.

A number of convincing arguments exist in favour of the judicious use of psychological testing in clinical settings with elderly cognitively impaired people (e.g. Miller, 1980; Woods & Britton, 1985); but the nature of relationships between performance on traditional psychological tests and measures of actual difficulties in carrying out the activities of everyday life are not well established in dementia. This is important as regards a number of clinical purposes. Work has generally focussed on relationships between overall intellectual level and everyday functioning in various groups of adult patients (as reviewed by Heaton & Pendleton, 1981) or on single aspects of function such as memory in tests and in life (e.g. Baddeley et al, 1982).

It was decided to investigate the relationship between performance on the tests used in this study and 'real- life' cognitive or behavioural competence. Self-report measures of everyday competence (as used by other workers with other populations) are most unlikely to be valid with all dementing subjects, and the most appropriate and practicable measure of everyday competence in a range of areas seemed to be some form of behaviour rating scale of the type commonly used with impaired elderly populations. Gilleard (1981) reviews the development of such scales, and accounts of the more popular ones can be found in texts such as Woods & Britton (1985) and Hanley & Hodge (1984). None of these is clearly superior to all the others in general terms, and any particular scale may be especially appropriate for certain purposes according to which aspects of functioning the scale covers in most detail. Most scales largely consist of items covering either or both of two main areas: behavioural disability

(or 'dependency', 'functional impairment', etc.) and behavioural disorder (or 'psychiatric disturbance', 'psychopathology', 'anti-social behaviour', 'social disturbance', etc.).

Measures used.

An appropriate rating scale for the purpses of the present study was assembled, in collaboration with the co-author of the CAPE, by combining items from two of the most commonly-used scales and adding a few items from other sources. It was hoped that the resulting combination of items would sample the main areas of behavioural functioning while remaining brief enough to be filled in willingly by appropriate raters. The two scales used as a basis were the Index of Independence in Activities of Daily Living (ADL, Katz et al, 1963) and the Behaviour Rating Scale from the Clifton Assessment Procedures for the Elderly (CAPE, Pattie & Gilleard, 1979).

The ADL was developed to assess impairment (and change in response to treatment) in elderly and chronically ill populations. It covers disability or dependency only, with three clearly defined levels of disability in six areas: bathing, dressing, toileting, transfer (i.e. mobility), continence, and feeding. According to the patterns of dependence in these six areas, subjects can be classified into one of seven levels of dependence. These dependency levels are defined according to their order of appearance in deteriorating subjects (or disappearance in recovering subjects); they have some of the characteristics of a Guttman scale, and also some theoretical relationship to the development of abilities in children. The scale has been used with very large numbers of elderly patients, and is known to have good reliability. The whole

of the ADL was used, as the first six items of the rating scale, in the present study.

The CAPE rating scale was developed from the longer Stockton Geriatric Rating Scale. It contains items concerning Physical Disability, Apathy, Communication Difficulties, and Social Disturbance, these four scales having been empirically derived from factor analysis of old people's scores on the scale. Again, this scale has been extensively used with elderly populations and is known to have adequate reliability. Twelve of the 18 CAPE items were included in the rating scale used here: the first four CAPE items were omitted as they are very similar to, and hence covered by, ADL items; CAPE item 6 (on 'appearance') was omitted as the rating would seem to depend greatly in some settings on the amount of attention paid by care staff to correcting the appearance of their charges; and CAPE item 8 (on 'helping out' in the home or ward) was omitted as being too dependent on how much opportunity someone is given to 'help out'. CAPE items 12 and 13 (on communication) were modified to cover communication by speech alone rather than communication by speech, writing, or gesturing. The reason for this modification now eludes the author. An item on night-time sedation was added as a supplement to the CAPE item on 'sleep pattern at night' since the use of sedatives might be a major determinant of sleep pattern. The layout of the CAPE items was slightly altered to match the ADL layout, with boxes to tick rather than numbers to circle.

Two items were added which were based on items shown to have predictive validity on a validation study of a Gresham Nurses' Psychogeriatric Inventory (Volans P J, 1982, personal

communication): these items respectively covered the ability to find one's own room (here extended to cover other rooms) and the ability or inclination to keep one's bed area tidy without assistance. (The author is grateful to Dr Volans for kindly providing a copy of her questionnaire and useful accompanying information.) Two more items were taken from a Dutch geriatric rating scale known as the B.O.P. (Van der Kam et al, 1971). These concern 'exaggerated repetitive movements' (pacing, rocking, etc.) and the production of sounds (muttering, moaning, etc.) which are not directed towards other people. Finally, three more items were made up to cover variability of functioning from day to day, variability over the course of single days, and emotional lability. The 26 main items in the resulting scale have three possible rating levels. The combined scale as actually presented to raters is shown in Appendix 2.

The scale was completed, on as many subjects as possible, by someone who knew the subject well and was in regular contact with him or her (most often a member of nursing or care staff, though occasionally a relative of the subject). Some subjects were rated both at the first testing and at the second testing 10 months later; others were rated only at the first testing, and a very few only at the second testing. Ratings were completed within a few days of testing. The reasons for subjects not being rated (ever, or on any given test occasion) included (1) the absence or unavailability of a sufficiently close informant and (2) authorial reluctance to ask busy nursing or care staff, particularly in locations where many subjects were being tested, to carry out time-consuming ratings purely for research purposes. For similar

reasons, data on inter-rater reliability were not collected. Almost all the items in the scale come from scales which have undergone proper reliability studies. Some data on test-retest reliability appear later in the chapter.

Subjects.

Initially, to maximise numbers, data from all subjects who were rated (at Test 1 or Test 2) will be lumped together (using only the data from Test 1 and the contemporaneous rating in subjects who were tested and rated twice). In this way, rating scale data and test data (on at least 'Short' testing) are available for 90 subjects (31 Male, 59 Female; mean age 76.0 years, range 39-93 years). This total comprised 37 DAT subjects (10 M, 27 F; 79.1, 56-93), 34 MID subjects (12 M, 22 F; 76.5, 44-92), and 19 subjects with Other diagnoses (9 M, 10 F; 69.0, 39-93). Of the total sample, 38 were on psychoactive medication. The locations of the subjects were: psychogeriatric day hospital 15, psychogeriatric in-patient 34, neurology in-patient 7, neurology out-patient 1, local authority part IV 28, other 5.

Results.

Firstly, individual rating scale items will be considered in relation to the overall score on Short testing (i.e. the mean proportion correct on all the tests comprising the Short collection as described in a previous chapter). Table 1 lays out the rating scale items (much as they appeared on the scale as presented to raters) with the numbers of subjects receiving a given rating level shown on the left and the mean Short proportion score for subjects receiving that rating shown on the right (together with details of a one-way ANOVA or t test for each item, indicating the

significance of the difference in mean Short scores when broken down by rated level on the rating scale item).

Table 1 Numbers of the 90 subjects receiving each of the three levels of rating for each of the 26 BRS items, and mean overall Short scores of these subject groups (plus standard deviations and results of one-way Anovas on the Short scores where possible. Where the n of cases in any group is less than 6, result of an independent t test on the other two groups is shown. * p less than .05, ** .01, *** .001).

n		Mean Short score	SD	F (or t)
1 BATHING (-either sponge bath, tub bath, or shower)				
21	Receives no assistance (gets in and out of tub by self if tub is usual means of bathing)	.73	.10	
25	Receives assistance in bathing only one part of the body (such as back or a leg)	.64	.14	22.1 ***
44	Receives assistance in bathing more than one part of the body (or not bathed)	.40	.26	
2 DRESSING				
41	Gets all clothes and gets completely dressed without assistance	.68	.14	
07	Gets clothes and gets dressed without assistance except for assistance in tying shoes	.62	.17	19.4 ***
42	Receives assistance in getting clothes or in getting dressed, or stays partly or completely undressed.	.40	.26	
3 TOILETING				
60	Goes to 'toilet room', cleans self, and arranges clothes without assistance (May use object for support such as stick, walker, or wheelchair; and may manage night bedpan or commode, emptying same in morning)	.63	.20	
23	Receives assistance in going to 'toilet room' or in cleansing self or in arranging clothes after elimination or in use of night bedpan or commode	.38	.26	12.9 ***
07	Doesn't go to room termed 'toilet' for elimination process	.38	.26	

4 TRANSFER				
71	Moves in and out of bed as well as in and out of chair without assistance (May be using object for support such as stick or walker)	.60	.20	
19	Moves in or out of bed or chair with assistance	.31	.27	t= 5.11 ***
00	Bedridden or chairbound	-	-	
5 CONTINENCE				
51	Controls urination and bowel movement completely by self	.61	.21	
30	Has occasional "accidents"	.53	.21	15.5
09	Supervision helps keep urine or bowel control; catheter is used; or is incontinent	.18	.23	***
6 FEEDING				
77	Feeds self without assistance	.59	.21	t=
12	Feeds self except for getting assistance in cutting meat or buttering bread	.24	.23	5.25 ***
01	Receives assistance in feeding; or is fed partly or completely using tubes or intravenous fluids	.20	-	
7 ORIENTATION				
50	Can find way to toilet room, own bed, and eating area without assistance	.67	.18	
16	Can find way to one or two of these without assistance	.53	.15	31.5 ***
24	Cannot find way to any of these without assistance	.29	.23	
8 CARE OF SLEEPING AREA				
21	Makes own bed and keeps area immediately around it tidy without assistance	.72	.10	
10	Makes own bed and tidies immediate area with assistance	.65	.13	12.0 ***
59	Doesn't make bed or tidy immediate area	.46	.26	
9 CONFUSION				
38	Not confused (Aware of situation & surroundings; recognises familiar faces; can find way around and rarely gets lost or loses possesions)	.67	.20	
34	Moderately confused	.55	.17	24.1
18	Severely confused; little or no grasp of situation	.27	.25	***

10 COPING OUTSIDE If allowed outside, would need supervision:			
18 Never	.65	.24	
32 Sometimes	.63	.16	10.3
40 Always	.42	.26	***
11 OCCUPATION Keeps self occupied in a constructive or useful way (works, reads, plays games, has hobbies, etc):			
14 Much of the time	.69	.13	
19 Sometimes	.63	.16	6.6
57 Almost never	.48	.27	**
12 COMPREHENSION OF SPEECH			
68 Understands almost everything you communicate	.63	.17	t=
20 Understands some of what you communicate	.30	.27	6.49
02 Understands almost nothing of what you communicate	.08	-	***
13 COMMUNICATION BY SPEECH			
70 Communicates well enough to be easily understood at all times	.59	.21	t=
19 Can be understood sometimes or with some difficulty	.38	.27	3.58 ***
01 Can rarely or never be understood for whatever reason	.03	-	
VARIABILITY			
14 From day to day, functioning in the kind of areas listed above tends to vary or fluctuate:			
66 Very little or not at all	.55	.25	t=
23 Moderately	.55	.25	0.00
01 Very noticeably	.20	-	n.s.
15 Over the course of a single day, functioning in these areas tends to vary or fluctuate:			
74 Very little or not at all	.56	.24	t=
15 Moderately	.45	.26	1.57
01 Very noticeably	.73	-	n.s.

SOCIALISATION

16

42 Establishes good relationships with others	.60	.20	
25 Has some difficulty establishing good relationships	.59	.21	6.6 **
23 Has a great deal of difficulty establishing good relationships	.39	.30	

17 Co-operation: Is willing to go along with requests:

65 Usually	.56	.23	t=
21 Sometimes	.56	.26	0.00
04 Almost never	.22	-	n.s.

18 Is objectionable to others (loud or constant talking, pilfering, soiling furniture, interfering with the affairs of others etc) during the day:

66 Rarely or never	.61	.21	
17 Sometimes	.35	.24	11.3
07 Frequently	.37	.32	***

19 And during the night:

70 Rarely or never	.60	.20	t=
16 Sometimes	.38	.30	3.53
04 Frequently	.18	-	***

20 Accuses others of doing him/her bodily harm or stealing his/her personal posessions (if accusations are all definitely true, tick "never"):

75 Never	.57	.23	t=
11 Sometimes	.38	.26	2.49
04 Frequently	.41	-	*

21 Hoards apparently meaningless items (wads of paper, bits of string, scraps of food, etc):

57 Never	.58	.22	
23 Sometimes	.46	.29	2.1
10 Frequently	.51	.24	n.s.

22 Engages in exaggerated repetitive movements (eg pacing about, rocking, wringing hands, fiddling with clothing):

53 Never	.62	.17	
26 Sometimes	.53	.26	20.0
11 Frequently	.19	.21	***

23 Makes sounds which are not particularly directed towards others (shouting, muttering, moaning):			
64 Never	.60	.21	
16 Sometimes	.47	.29	8.3
10 Frequently	.30	.23	***
24 Sleep pattern at night is:			
49 Almost never awake	.58	.20	
35 Sometimes awake	.51	.29	1.8
06 Often awake	.41	.29	n.s.
25 Night-time Sedation:			
64 Almost never has sedation	.55	.24	
12 Occasionally has sedation	.50	.32	0.2
14 Regularly has sedation	.56	.20	n.s.
26 Emotional lability.			
73 Mood is generally stable, with normal and understandable variation	.56	.25	t=
16 Moderate fluctuation in mood with some spells of laughing or crying for no apparent reason	.49	.24	1.00 n.s.
01 Marked emotional lability with frequent spells of innappropriate crying or laughing, the beginning and end of such spells often being sudden and abrupt	.14	-	

It can clearly be seen that for the great majority of items there is a highly significant relationship between level of rated impairment or abnormality and test score. Such a relationship is absent for the items concerning variability (both day to day and over the course of a single day); co-operation; hoarding of items; sleep pattern and night-time sedation (these two probably being related, and the finding being consistent with other results showing at least as good test scores in subjects taking psychoactive medication as in those taking none); and emotional lability. The picture was broadly similar in the DAT and MID groups alone.

It is possible to examine the relationship between particular

rating scale items and test performance in more detail by considering scores on specific tests as opposed to an overall mean test score. For example, in those 61 subjects who completed Full testing and were rated, Block Design performance seems to be more strongly associated with rated level on the Orientation rating scale item than with level on the Communication by Speech item; while Sentence Production score is more strongly associated with level on Communication by Speech than with level on rated Orientation (judging by the mean scores and the significance of the F ratio in one-way ANOVAs with test score broken down by level on the rating scale item); or, Block Design performance is more strongly associated with level on rated Orientation than is Sentence Production, and Sentence Production is more strongly associated with level on rated Communication by speech than is Block Design. Such findings parallel those of De Leon et al (1984) who found a relationship between wandering and parietal signs (but not performance on a general mental status test) in SDAT. These analyses will not be reported since the possible number of cross-comparisons of this sort is huge (and since some of the few comparisons which have been carried out do not come out nearly as neatly as the examples given). However some indication does seem to exist of specificity in the relationships between specific aspects of test and real-life performance.

To move on from individual items, a variety of compound scales could be calculated by combining scores on various rating scale items. Five such scales were calculated, after scoring each rating scale item as 0 for the most independent or 'normal' level, 1 for the intermediate level, and 2 for the most dependent or disturbed

level. The computed scales were:

(1) ADL: the total score on items 1 to 6 inclusive (bathing, dressing, toileting, transfer, continence, and feeding).

(2) Other Dependency: the total score on items 7 to 13 inclusive (orientation, care of sleeping area, confusion, coping outside, occupation, comprehension of speech, and communication by speech).

(3) Total Dependency: (1) plus (2).

(4) Social disturbance: the total score on items 16 to 24 inclusive and item 26 (establishing relationships, co-operation, objectionableness during day and during night, paranoid accusations, hoarding, repetitive movements, undirected noise-making, sleep pattern at night, and emotional lability).

(5) Overall total: (3) plus (4) i.e. the total score on all items on the questionnaire excluding item 25 (hypnotic usage).

These particular scales were constructed on a priori grounds, bearing in mind previous work, without recourse to factor analysis or other empirical derivation procedures. (In fact a later factorial procedure produced seven rather complex, difficult to interpret, and therefore unhelpful factors.) Intercorrelations between these five scales, and their correlations with age and Short score, are shown in Table 2.

All the correlations are significant and in the expected direction. The dependency measures (1), (2), and (3) intercorrelate very strongly. They correlate less highly with the social disturbance measure (4). The dependency measures correlate more strongly with Short score than does the social disturbance one. All measures correlate weakly with age in the direction of older subjects being more dependent, more disturbed, and (in this sample)

poorer on testing.

Table 2 Intercorrelations of the five compound scales and age and Short test score. Pearson product-moment correlations. n=90.

	(2) Other depend.	(3) Total depend.	(4) Social disturb.	(5) Overall total	Short score	Age
(1) ADL	.82	.95	.56	.89	-.67	.28
(2) Other depend.		.96	.57	.91	-.70	.35
(3) Total depend.			.59	.94	-.72	.33
(4) Social disturb.				.82	-.54	.26 $
(5) Overall total					-.71	.34
Short score						-.33

All correlations are significant at the .01 level of significance except $, significant at .02.

Consideration of correlations between the BRS compound measures and scores on each individual test is not very revealing: most tests correlate significantly in the expected direction with most BRS measures (usually excluding Social Disturbamce). Table 3 shows correlations between the BRS measures and test factor scores.

Table 3 Correlations between BRS scales and test factor scores and overall scores. Pearson product-moment correlation coefficients. ns: not significant (2-tailed test).

	ADL	Other Depend.	Total Depend.	Social Disturb.	Overall Total
Full (n=61)	-.53	-.53	-.56	-.11 ns	-.48
Factor 1	-.39	-.50	-.49	-.11 ns	-.43
Factor 2	-.25	-.13 ns	-.20 ns	.02 ns	-.15 ns
Factor 3	-.08 ns	-.13 ns	-.12 ns	.01 ns	-.07 ns
Factor 4	-.46	-.45	-.48	-.27	-.47
Short (n=90)	-.67	-.70	-.72	-.54	-.71
Short factor 2	-.62	-.65	-.67	-.55	-.68
Short factor 1	-.48	-.48	-.50	-.31	-.48

Correlations are in the expected directions. Full Factors 2 and 3 seem to be poorer predictors of real life performance than are Full Factors 1 and 4 and the Short factors. None of the correlations is extremely high.

It seems that the cognitive test measures do bear relationships to 'real-life' measures. However, these relationships are not sufficiently strong to allow reliable prediction of behavioural deficits from test impairments in the individual case. The fact that correlations between test score and rated dependency fall short of unity (around .7) may partly reflect the contribution of purely physical disability to dependence. The 'true' relationship between test performance and the cognitive components of behavioural incompetence may be stronger than that suggested by these results (although the 'other dependency' measure correlates with Short score no more strongly than does the ADL measure despite the 'other dependency' items being apparently less physical than the ADL ones). The correlation of only .5 between test score and social disturbance is unremarkable since intellectual deficit and psychiatric or social disturbance are not generally considered to increase apace in dementia.

Changes over time.

As previously stated, the above analyses were based on a combination of test 1 and test 2 data, to maximise numbers while never including a given subject twice. Subjects who were rated twice and received at least Short testing twice will now be considered. In all there were 37 such subjects (9 Male, 28 Female; mean age 75.8 years, range 39-93 years). This total comprised 15 DAT subjects (2 M, 13 F; 80.1, 68-93), 14 MID subjects (5 M, 9 F;

76.2, 55-87), and 8 subjects with Other diagnoses (2 M, 6 F; 66.9, 39-83). Of the total sample, 16 were on psychoactive medication. The locations of the subjects when seen initially were: psychogeriatric day hospital 5, psychogeriatric in-patient 5, neurology in-patient 1, local authority part IV 22, other 4. The location at the time of test two was usually the same.

A number of factors may contribute to findings (reviewed by Gilleard, 1978) that decline over fairly short periods in dementia may be apparent on rating scale measures but not on cognitive test measures. These include possibilities that (1) rating scales are more sensitive measures of impairment or change in level of impairment than are cognitive tests (or at least memory tests), and are less prone to floor effects; (2) subjects become physically frailer as they get older but suffer no measurable decline (regardless of method of measurement) in mental competence; (3) subject groups tend to be heterogeneous, and group means mask changes in some groups; and (4) practice effects occur with tests (thus masking an 'actual' decline) but not with rating procedures. The present study cannot tease out the relative contributions of such factors, but a brief look at changes over time in the BRS and test measures is of interest. Table 4 shows the changes in the various BRS scores and in Short test score from Test 1 to Test 2.

In the DAT group, the test score shows no change but some of the BRS measures do - essentially in fact the ADL measure. The ADL relates very much to physical dependence: the fact that the other (less physical) dependency measure shows no significant change suggests that possibility (2) above may be important, i.e. that subjects may be becoming physically frailer rather than measurably

Table 4 Mean deterioration on Short score and on each of the BRS measures from Test 1 to Test 2, in each group and in the three combined. (Standard deviations in parentheses; significance tested by correlated t-test.)

	DAT (n=15)	MID (n=14)	Other (n=8)	All (n=37)
Short	-.001 (.076)	.023 (.112)	-.048* (.048)	-.002 (.089)
(1) ADL	-1.53** (1.55)	-0.50 (2.62)	-1.00 (1.85)	-1.03** (2.08)
(2) Other dependency	-0.93 (3.13)	0.00 (2.00)	-1.75+ (2.61)	-0.76+ (2.65)
(3) Total dependency	-2.47* (4.17)	-0.50 (4.17)	-2.75+ (3.50)	-1.78* (4.06)
(4) Social disturbance	0.00 (4.00)	1.07+ (1.98)	-1.50 (2.67)	0.08 (3.16)
(5) Overall total	-3.27+ (6.86)	0.14 (6.21)	-4.38+ (5.32)	-2.22* (6.43)

Signs are adjusted so that negative sign always indicates deterioration (i.e. a drop in test score or a rise in BRS score).
+ p .10, * p .05, ** p .01 (2-tailed test).

more demented. In the MID group, neither test score nor ratings demonstrate deterioration, with an improvement in social functioning the only change approaching significance. In the very small Other group, the test score seems to be at least as sensitive to change as any of the rating measures. In the groups combined together, again the ADL emerges as the measure most sensitive to change.

Correlations (in each subject group and in the three combined) between change in Short score and change on the various BRS measures were all non-significant (at the .05 level, Pearson Product-Moment correlations, 2-tailed test) except for a significant correlation between change in Short score and change in Social disturbance score in the small Other group. Correlations

were calculated between change in each BRS measure and the initial score on that measure, in each group and in the three combined. A significant correlation was found between change and initial level only on the Social disturbance measure (in the DAT and MID groups and in the three groups combined, though not in the small Other group). None of the BRS change measures correlated with age.

In this small sample, then, it seems that ratings may register change more than tests do; that this may partly reflect the contribution of physical rather than psychological factors; and that the test and rating scale measures of change do not correlate highly with each other, perhaps again partly as a result of the contribution of physical changes to the rating measures.

A final note concerns reliability: Table 5 shows correlations between first and second ratings (10 months apart) for each of the 5 computed BRS scales (and for Short score in this sample).

Table 5 Test-retest correlations for the 5 computed BRS scales and Short score. Pearson product-moment correlations. n=37.

(1) ADL	.74
(2) Other dependency	.77
(3) Total dependency	.81
(4) Social disturbance	.50
(5) Overall total	.71
Short score	.94

The scales appear to have reasonably good test-retest reliability considering the length of time subjects had in which to actually change in level of functioning. The figure for the Social disturbance scale is rather disappointing. Looking at diagnostic groups individually, it appears that Social disturbance shows poor test-retest reliability in the DAT group but not in the other two groups. All other measures show good reliability, similar to the

above figures, when broken down into separate diagnostic groups. It may be that the DAT subjects did in fact change more one way or another as regards Social disturbance despite their mean change of zero.

Behaviour during testing.

This seems an appropriate point to consider briefly some other ratings used in this study. These concern aspects of subjects' behaviour during testing.

After every test session (or pair or triplet of sessions where testing was split up) ratings were made on 23 five-point scales and a further 11 three-point scales concerning aspects of the subject's behaviour during introductory conversation and testing. Many of these aspects were simply decided on from scratch as being apparently relevant; others came from the 'Philadelphia Geriatric Centre Minimal Social Behaviour Scale', from a paper by Ballinger et al (1982), and from a paper by Aggernaes et al (1975). (Another 14 items from the Philadelphia GCMSBS concerning very basic response to and interaction with the tester were also completed for each subject. 'Failure' on these items was so rare that these items will not be considered further.) Scores on seven scales were then calculated by adding together rated scores on various items. The items and the scales to which they contribute are shown in Table 5. The actual rating sheet used appears in the manual in Appendix 1.

The ratings were made on an impressionistic basis, and individual score levels on each item were not operationally defined. It was hoped that the author would use the ranges of scores in a consistent manner. The computed scales were constructed on a priori grounds, without the use of empirical methods of

Table 5 Aspects of behaviour rated after each testing, split according to the compound scales later computed. Plus and minus signs indicate whether the item contributes positively or negatively to the scale (i.e. whether the score is added or subtracted) according to whether a high rating score indicates normality or abnormality.

(Scored 0, 1, 2, 3, or 4:)	
+ Apparent confusion	
+ Impaired consciousness	
+ Distractibility	
+ Perplexity	'Confusion'
- Attention	
- Concentration	
- Insight re. own condition	
+ Engagement	
+ Cooperation	'Engagement'
+ Interest in tests	
+ Anxiety	'Anxiety'
+ Distress at failure	
+ Depression	
+ Elation	'Mood disturbance'
+ Emotional lability	
+ Disinhibition	
+ Confabulation	'Frontal behaviour'
+ Perseveration	
- Insight re. own condition	
+ Psychomotor slowing	'Slowing'
+ Apparent fatigue	
+ Paranoid features	
+ Hostility/aggression	
+ Suspiciousness	
(Scored 0, 1, or 2:)	
+ Delusions	
+ Hallucinations	
+ Thought disorder	'Other psychiatric
+ Inappropr. grimaces/mannerisms	symptomatology'
+ Bizarre sitting position	
+ In constant motion	
+ Motionless	
+ Disarranged clothes	
+ Drooling	
+ Nasal mucous/food conspicous (on clothes, face)	
+ S attempts to move away (without explanation)	

derivation such as factor analysis. The final computed scale described above, 'Other psychiatric symptomatology', will be mentioned little since very few subjects scored many points on it.

Differences between diagnostic groups will be considered only briefly: at least some of the scales may reflect features which have more to do with premorbid personality than with the nature of the dementing condition. Table 6 shows scores on the seven scales (at Test 1) of all subjects who completed at least Short testing, broken down by diagnostic category.

Table 6 Scores on the seven computed rating scales (at test 1) of all subjects who completed at least Short testing, broken down by diagnostic category. (Standard deviations in parentheses; F and significance of one way ANOVA at the right.)

	DAT (n=74)	MID (n=74)	Other (n=66)a	F	sig.
Confusion	-5.6 (4.2)	-6.6 (4.2)	-7.6 (4.2)	3.96	.02
Engagement	9.6 (2.2)	9.8 (2.3)	10.0 (2.4)	0.49	ns
Anxiety	1.6 (1.3)	1.4 (1.0)	1.2 (1.2)	2.54	.08
Mood disturbance	1.3 (1.1)	1.4 (1.1)	1.2 (1.1)	0.44	ns
Frontal behaviour	-0.3 (1.9)	-1.1 (1.7)	-0.9 (2.3)	3.39	.04
Slowing	1.1 (1.2)	1.7 (1.8)	0.8 (1.4)	7.10	.001
Other psychiatric symptomatology	0.5 (0.9)	0.5 (0.8)	0.8 (1.5)	2.69	.07

a One of the 67 subjects in this group was not rated by mistake.
A higher arithmetical score always indicates 'more' of the relevant feature: i.e. -5 indicates more confusion than -6.

The confusion scores are in the rank order one would expect from overall test scores. DAT subjects showed more Frontal behaviour

and less Slowing than did MID subjects. Otherwise the findings are unremarkable.

Pearson product-moment intercorrelations between these computed variables were calculated: most of the scales correlated significantly with one another in the expected direction (i.e. positively except where Engagement was one of the variables, in which case negatively). Predictably, Anxiety did not correlate significantly with Frontal behaviour or Slowing; rather less predictably, Frontal behaviour and Other psychiatric symptomatology did not correlate with Slowing, and Anxiety did not correlate with Confusion. Few of the correlations were very large, so there is at least a possibility that the different computed scales reflect different aspects of subjects' behaviour during testing (though of course the ratings were certainly not made independently of each other).

Anxiety correlated negatively with age in the DAT group and positively with age in the MID group. Frontal behaviour correlated positively with age in MID. Other correlations with age were not significant in the DAT or MID groups. There were some significant correlations, in predictable directions, between the computed scales and Short score. It would be possible to examine relationships between these seven scales and scores on the various individual tests. Since the ratings were completed after observing the test performances, the ratings are likely to be influenced or contaminated by the test performances: hence demonstration of specific relationships would prove little and will not be attempted.

Relationships between these test session ratings and the

independent ratings from the BRS were examined to determine whether behaviour ratings made on the basis of test sessions reflect aspects of everyday behaviour. Pearson product-moment correlations between the seven computed session rating variables and the five computed BRS scales previously described, in all 90 subjects, included several highly significant correlations in expected directions. The highest were between session ratings of Confusion and BRS dependency scores, though none exceeded .75. It seems that ratings made on the basis of a test session can be indicators of something 'real', as can the test scores themselves.

86 subjects were retested 10 months later on at least Short testing: 43 DAT, 25 MID, and 18 Other. (Subject details have appeared in a previous chapter.) The significant changes from test 1 to test 2 on the seven computed rating variables (at .05 level by correlated t-test, 2-tailed test) were as follows: DAT subjects showed less Anxiety, Mood disturbance, and Slowing, and more Frontal behaviour. Tendencies to show more Confusion and Engagement were significant only at the .10 level. MID subjects showed less Slowing and more Engagement. A tendency to show less Anxiety was significant only at the .10 level. Subjects with Other diagnoses showed only tendencies (significant at the .10 level) toward less Mood disturbance and Slowing. Test-retest correlations are shown in Table 7.

This suggests fair test-retest reliability for only a few of the measures: the length of time subjects had in which to change might have contributed to the other low correlations, but some of the measures may be genuinely unreliable. Findings must therefore be considered tentative.

Table 7 Test-retest reliability of the seven computed rating variables. (Pearson product-moment coefficients; n=86)

Confusion	.78
Engagement	.61
Anxiety	.43
Mood disturbance	.34
Frontal behaviour	.70
Slowing	.45
Other psychiatric sympt.	.33

CHAPTER 5

Variability Over Time in DAT and MID.

Accounts of the common clinical presentations of dementia assert that patients with MID show greater fluctuation in level of functioning over short periods of time (over the course of a day or from day to day) than do patients with DAT. This idea is enshrined in the Hachinski index, where a positive rating on the item 'Fluctuating course' earns two points. The reasons for such fluctuation are naturally thought to relate to the state of the cardiovascular system in MID, though opinions differ as to whether the most common sources of fluctuation are transient ischaemic attacks due to ephemeral emboli or more general fluctuations in blood pressure or flow or cerebral perfusion. If the day-to-day or hour-to-hour variation in level of functioning is marked in patients with MID, then interpretation of their test performances on any single occasion (either clinically or in theses such as this) must be guarded as regards judging both the overall level of impairment and the pattern of strengths and weaknesses.

Data on test reliability has already been described in Chapter 3. This next little study can be seen as providing additional information on the reliability of certain tests: primarily, however, it was conducted to compare DAT and MID subjects regarding the extent to which variation occurs from day to day by making three assessments within a two-week period. The tests chosen included those the author hoped would be <u>most</u> sensitive to short-term variation or fluctuation.

<u>Subjects</u>

The subject groups comprised 16 DAT Ss (3 M, 13 F; Mean age 78.4 yrs, range 71-91) and 13 MID Ss (7 M, 6 F; 77.5 yrs, 65-86), diagnosed in the usual way with the same inclusion-exclusion

criteria as described for the DAT and MID groups in the main studies. Almost all were day hospital patients. Selection was based largely on grounds of convenience i.e. that the patients were regular attenders at (or residents in) an easily accessible local facility. No special attention was paid to whether or not the MID patients scored positively on the 'Fluctuating course' item of the Hachinski index. Three members of the DAT group were on some psychoactive medication (2 on a small dose of a major tranquilliser, 1 on a hypnotic) and six of the MID group (5 on a small dose of a major tranquilliser, 1 on an anti-depressant). No subject had had any recent change in their medication regime, and none had any alteration in medication between first and third testing.

Method

All subjects initially received Full testing as described previously. Approximately one week after this, and then one week later again, each subject repeated those tests comprising Short testing, i.e. (in order of presentation):

1. ORIENTATION
2. AUTOMATIC SPEECH
3. DIGIT SPAN
4. BOX-FILLING
5.* Supplementary Dysphasia Test
6. Writing Name & Reading Sentence
7. SQUARE WITH STICKS
8. SENTENCE REPETITION
9. SPIRAL MAZE 1
10. SPIRAL MAZE 2
11. PORTEUS MAZES

* not all subjects received this dysphasia test at the first testing, though all received it at the second and third. Everyone was tested on the ability to identify a spoon, key, and pencil on all three occasions.

Session ratings as described previously were also completed on each test occasion. Testing was carried out as far as possible at

the same time of day on each occasion for any given patient. (Diurnal variation would be another interesting area of study, but to combine such an independent variable with that of day-to-day variation would have required more subjects than it was practicable for the author to assess.) There were slight variations in the times of day at which given patients were seen, just as there were slight individual differences in the exact number of days between testings. (Weekly intervals were chosen simply to fit in with day hospital programmes, where a patient's repeated availability was most likely at a given time on a given day of the week.) The possible influence of these variations is considered below.

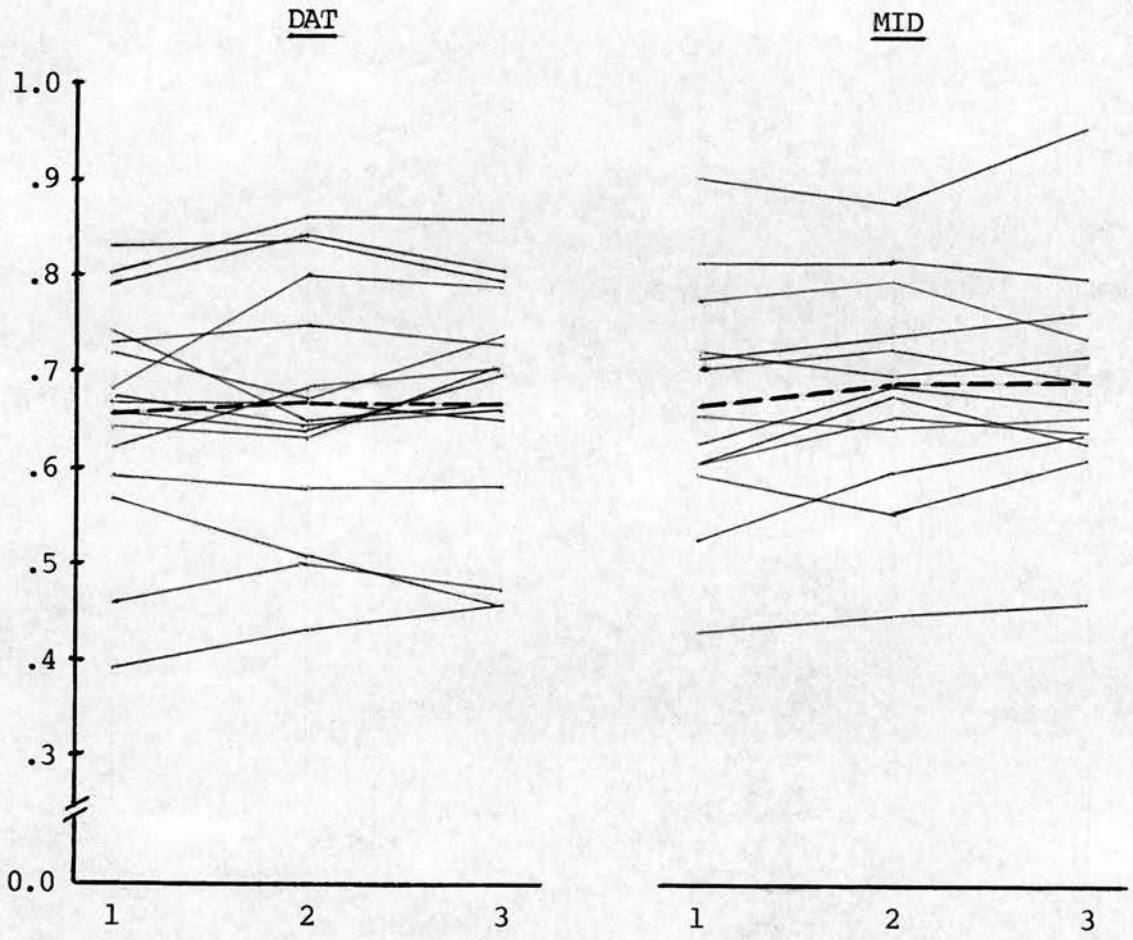
Results

The groups did not differ on overall level of performance at initial testing, as shown below. (The Full and Short overall scores and individual test scores are calculated just as described in Chapter 2, but are here expressed as proportions of one rather than as percentages; again, the Supplementary Dysphasia Test is not included in calculating the Short score as not every subject was given it on the first test occasion.) Standard deviations are shown in parentheses.

	DAT	MID
Full	.55 (.11)	.56 (.15)
Short	.66 (.12)	.67 (.13)

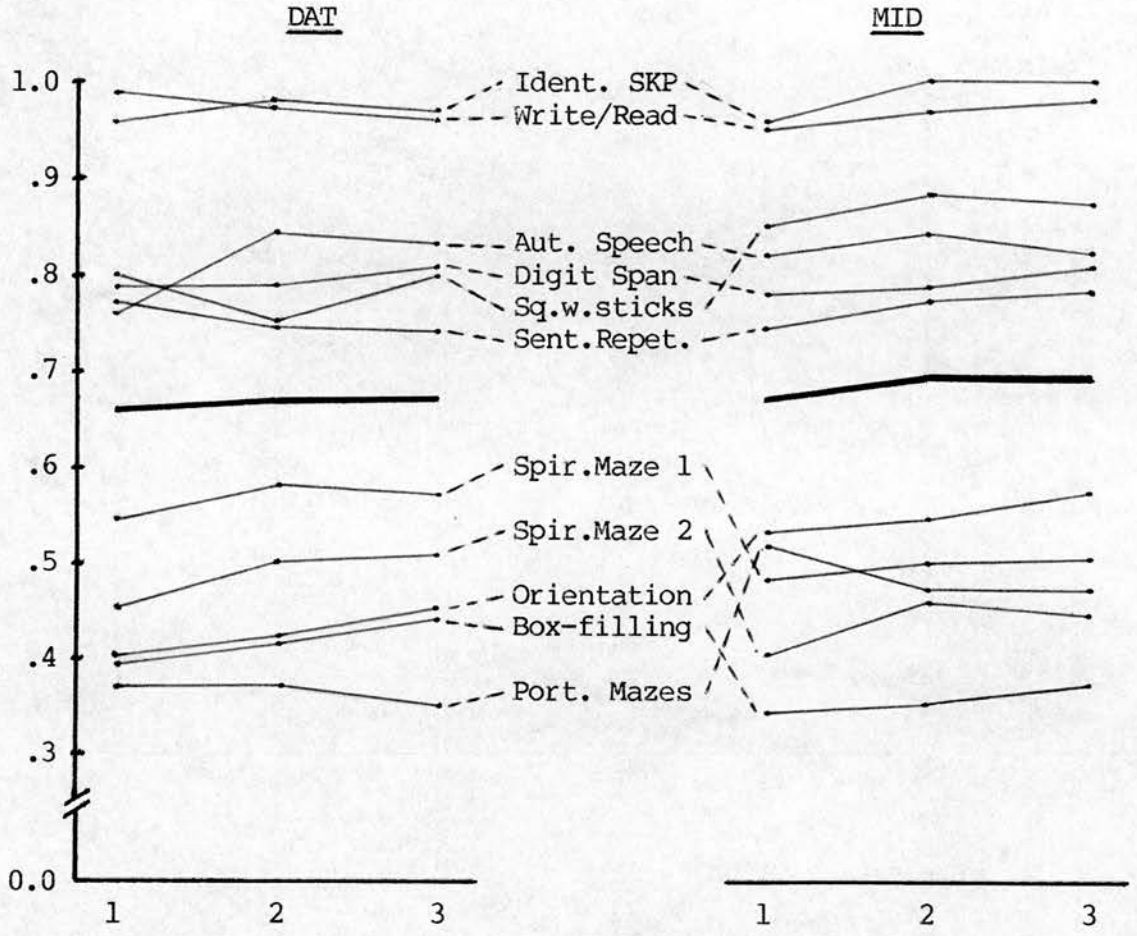
To give a general picture of the patterns of performance of the two groups, Figure 1 shows the overall Short score on each of the 3 test occasions for every subject. The heavy lines show the group means. Figure 2 shows the group means for each of the 12 individual tests on each test occasion, with the grand means again in heavy

Figure 1 Short score on each test occasion for every subject. (Group means shown by heavy line.)



Abscissa shows test number; ordinate shows overall Short score.

Figure 2 Group mean for each test on each test occasion. (Grand means shown by heavy line.)



Abscissa shows test number; ordinate shows mean proportion score.

line. It can be seen that the group means are very similar indeed on all test occasions. These means, however, may mask significant differences between the groups in terms of change in individuals' levels of functioning over the three testings. Therefore three difference scores were calculated for each subject for the overall Short measure and for each of the 11 constituent tests:

(1) test 2 score minus test 1 score

(2) test 3 score minus test 2 score

(3) test 3 score minus test 1 score (i.e. the sum of the above two)

Since the magnitude rather than the direction of change is of primary interest, four further difference scores were calculated for each subject, rather like the first three but using absolute rather than arithmetical values of the difference scores (i.e. ignoring the sign):

(4) Abs. value of (test 2 score minus test 1 score)

(5) Abs. value of (test 3 score minus test 2 score)

(6) Abs. value of (test 3 score minus test 1 score)

(7) (4) plus (5)

(Scores (6) and (7) are not necessarily equivalent, i.e. if a person's performance on a given test goes up then down again or down then up again over the course of the three tests.)

The two groups were compared on each of these 7 difference scores (for the overall Short measure and for each of the 11 individual tests) using independent t-tests. The results of these analyses are shown in Table 1. One would expect a few significant results by chance alone.

Table 1 Mean difference scores of the two groups on all test measures, with level of significance of the resulting t value in parantheses.

(a) Arithmetical values	(1) 2-1 DAT	(1) 2-1 MID	(2) 3-2 DAT	(2) 3-2 MID	(3) 3-1 DAT	(3) 3-1 MID
Short overall measure	.012	.017 (.75)	.005	.002 (.84)	.017	.019 (.89)
ORIENTATION	.014	.012 (.95)	.029	.030 (.99)	.043	.041 (.97)
AUTOMATIC SPEECH	.074	.024 (.44)	-.008	-.024 (.75)	.066	.000 (.37)
DIGIT SPAN	.002	.004 (.94)	.014	.021 (.85)	.015	.025 (.73)
BOX-FILLING	.012	.012 (.99)	.026	.012 (.56)	.038	.024 (.66)
(Supp. Dysphasia)	.025	-.007 a (.48) b	-.037	.010 (.11)	-.017	-.014 a (.94) b
Identifying SKP	.021	.039 (.55)	-.010	.000 (.62)	.010	.039 (.45)
Writing Name & Reading Sentence	-.010	.013 (.23)	-.007	.009 (.43)	-.017	.021 (.13)
SQUARE WITH STICKS	-.047	.039 (.25)	.047	-.019 (.24)	.000	.019 (.82)
SENTENCE REPETITION	-.025	.021 (.18)	-.010	.017 (.31)	-.035	.039 * (.03)
SPIRAL MAZE 1	.040	.022 (.67)	-.019	-.003 (.75)	.021	.019 (.97)
SPIRAL MAZE 2	.048	.057 (.82)	.011	-.017 (.44)	.058	.040 (.67)
PORTEUS MAZES	-.000	-.051 (.42)	-.017	-.004 (.79)	-.017	-.056 (.58)
£ (Total of 11 difference scores)	.128	.190 (.75)	.055	.021 (.84)	.183	.212 (.89)

a n=5 b n=6

£ i.e. total of the difference scores on the above individual test, excluding Supplementary Dysphasia since not all subjects received this at the first testing.

(b) Absolute values	(4) 2-1 DAT	MID	(5) 3-2 DAT	MID	(6) 3-1 DAT	MID	(7) (4)+(5) DAT	MID
Short Overall Measure	.042	.035 (.48)	.029	.034 (.51)	.046	.032 (.29)	.071	.069 (.87)
ORIENTATION	.053	.118 * (.03)	.077	.112 (.21)	.091	.089 (.93)	.130	.231 * (.04)
AUTOMATIC SPEECH	.129	.082 (.37)	.078	.101 (.53)	.152	.067 (.15)	.207	.183 (.73)
DIGIT SPAN	.049	.054 (.83)	.054	.083 (.32)	.049	.050 (.97)	.103	.137 (.45)
BOX-FILLING	.055	.058 (.81)	.061	.046 (.32)	.085	.048 (.09)	.116	.104 (.57)
(Supp. Dysphasia)	.042	.049 a (.84) b	.047	.042 (.83)	.033	.042 a (.77) b	.100	.083 a (.81) b
Identifying SKP	.021	.039 (.55)	.031	.000 (.11)	.031	.039 (.84)	.052	.039 (.74)
Writing Name & Reading Sentence	.031	.013 (.29)	.028	.017 (.54)	.038	.021 (.47)	.059	.030 (.30)
SQUARE WITH STICKS	.109	.077 (.62)	.047	.135 * (.05)	.156	.096 (.36)	.156	.212 (.47)
SENTENCE REPETITION	.058	.067 (.70)	.052	.046 (.75)	.057	.064 (.81)	.110	.114 (.92)
SPIRAL MAZE 1	.093	.094 (.97)	.125	.071 (.08)	.105	.135 (.33)	.218	.165 (.19)
SPIRAL MAZE 2	.107	.077 (.26)	.077	.069 (.68)	.106	.063 (.17)	.184	.146 (.23)
PORTEUS MAZES	.146	.111 (.38)	.094	.107 (.65)	.135	.133 (.95)	.240	.218 (.69)
£ (Total of 11 difference scores)	.850	.790 (.69)	.725	.787 (.50)	1.007	.804 (.21)	1.575	1.577 (.99)

a n=5 b n=6

£ i.e. total of the difference scores on the above individual tests, excluding Supplementary Dysphasia since not all subjects received this at the first testing.

In part (a) of Table 1 there is only one significant difference between the groups: on Sentence Repetition, the difference between the slight improvement in the MID group and the slight decline in the DAT group from test 1 to test 3 is significant. Otherwise there is no evidence of either group simply improving more than the other with repeated testing. Part (b) of the table is therefore interpretable without complication. (It is impossible to judge from the available data whether memory or learning might have been influential in another way, in the sense of helping to guard against poorer performance at the second or third testings. For example, a subject may perform at a particular level on Sentence Repetition at test 1 and then perform at a similar level at tests 2 or 3, even if his repetitive ability is in fact worse in some way, because some memories or savings of the sentences remain from the earlier tests. The use of different but parallel forms of testing might have clarified this point, but might also have complicated matters in the likely event of parallel forms not being of precisely equivalent difficulty.)

In part (b) of the table, the MID group show significantly more variability over time on two measures of change on Orientation - (4) and (7) - and on one measure of change on Square with Sticks - (5). This is very minor considering the number of comparisons made. It does not seem that there is an overall tendency for the MID group to be more variable but which is generally failing to reach significance because of the fairly small numbers of subjects: looking at the two overall measures in part (b) of table 1 (first and last rows), the MID figures are not consistently higher than the DAT ones. And considering the 11 individual tests (again

excluding Supplementary Dysphasia), the number of times that the MID figure is higher than the DAT one in the four columns of part (b), from left to right, are 6/11, 5/11, 4/11, and 4/11.

Variations in the session ratings (transformed into a series of scales as described in the previous chapter) were analysed in a similar manner. Out of a large number of comparisons, only two variables showed a significant difference between the groups (both in analyses using absolute values of difference scores): DAT subjects varied more than MID ones between test 2 and test 3 on the compound 'confusion-consciousness' scale, while MID subjects varied more than DAT ones between test 1 and test 2 on the compound 'anxiety' scale. These findings are probably attributable to chance.

There appear to be no artefactual reasons for the general lack of significant differences between the groups on testing. Figures 1 and 2 suggest no floor and few ceiling effects. The groups did not differ significantly (by t-test) regarding the mean number of days between tests 1 and 2, 2 and 3, or 1 and 3. The seven 'Short overall' and seven 'Total of 11 difference scores' measures shown in Table 1 were analysed broken down by sex: there were no significant differences by t-test between the sexes. These same measures were also broken down according to whether subjects were on any psychoactive medication or not: again, there were no significant differences by t-test between subjects on drugs and those on none. Table 2 shows some relevant correlations between the seven measures of change (using firstly the 'Short overall' and secondly the 'Total of 11 difference scores' measures) and overall Short score at initial testing, age, and the number of days

separating the tests.

Table 2 Correlations between change scores and initial Short score, age, and the number of days separating the relevant test sessions, in each group and in the two combined. Pearson product-moment correlations; * sig. at .05 level.

(a) Short overall measure

	Real			Absolute			
	(1) 2-1	(2) 3-2	(3) 3-1	(4) 2-1	(5) 3-2	(6) 3-1	(7) (4)+(5)
DAT (n=16)							
Short score	-.05	-.07	-.09	.04	-.02	-.14	.02
Age	.38	.17	.47	.08	.16	-.05	.19
n days between	.27	.61*	.48	.49	.50*	.57*	.59*
MID (n=13)							
Short score	-.44	.02	-.41	-.43	.48	-.17	.01
Age	.04	-.13	-.10	.04	-.31	.01	-.16
n days between	-.31	.07	-.42	-.45	.34	-.41	-.33
All (n=29)							
Short score	-.19	-.03	-.20	-.14	.20	-.16	.01
Age	.24	.03	.25	.08	-.05	-.01	.03
n days between	.07	.39*	.21	.14	.38*	.25	.23

(b) Sum of 11 difference scores measure (the correlations for the real values are as above)

	Absolute			
	(4) 2-1	(5) 3-2	(6) 3-1	(7) (4)+(5)
DAT (n=16)				
Short score	-.21	-.15	-.44	-.22
Age	-.03	-.01	.27	-.03
n days between	.42	.24	.24	.28
MID (n=13)				
Short score	-.52	.10	-.54	-.36
Age	.00	-.11	.13	-.06
n days between	-.43	.03	-.37	-.19
All (n=29)				
Short score	-.33	-.03	-.47*	-.27
Age	-.01	-.06	.22	-.04
n days between	.10	.12	.04	.14

The significant negative correlation between initial Short level and change measure (b) (6) in the two groups combined (and similar non-significant correlations) probably reflects ceiling effects in some subjects on some tests. This cannot account for all the

findings on all tests. Otherwise, the only significant correlations are between change scores and the number of days separating the relevant tests. As regards measures (1) to (3), the two significant correlations are positive: i.e. improvement from test 2 to test 3 correlates (really only in the DAT group) with a long gap between the two tests (or, deterioration correlates with a short gap). Why this might be so is not clear. Perhaps it is a chance finding. As regards the absolute measures, (4) to (7), the significant correlations are again positive (again particularly in the DAT group). This is more explicable in common-sense terms: subjects are more likely to show change (in either direction) the longer the interval in which they are able to change. This is interesting, but does not account for the lack of difference between groups: the groups do not differ in mean lengths of inter-test intervals, as previously stated.

It is not simply the case that no variability is present, as might be the case if the tests were mis-chosen and insensitive to change. Figures 1 and 2 give some suggestion of this variability and it was even clearer from inspection of individual subjects' performance on individual tests over the three occasions. It seems that the two subject groups do vary over time, but do so equally. The fact that the MID subjects were not selected according to whether they they scored positively on the 'fluctuating course' item of the Hachinski index undoubtedly has a bearing on these results. But, from the results from these particular samples, one may legitimately conclude that greater variability over time in MID than in DAT is not a universal characteristic of the conditions.

Test-retest reliability.

Finally, some comments will be made on test-retest reliability. Test-retest correlations (test 1:test 2, test 1:test 3) for the two groups combined are shown in Table 3. For most tests the correlations indicate fairly good test reliability. Again this is encouraging as regards the results of the main studies. The exceptions occur on tests where the range of scores in these particular samples was very narrow. The overall Short score is highly reliable. Correlations in the MID group alone were not consistently lower than in the DAT group alone, consistent with all the above evidence showing no greater variability in MID.

Table 3 Test-retest correlations in each group and in the two combined.

	All (n=29)	
	1:2	1:3
ORIENTATION	.88	.88
AUTOMATIC SPEECH	.64	.53
DIGIT SPAN	.80	.85
BOX-FILLING	.93	.89
(Supp. Dysphasia)	a .37ns	.87
Identifying SKP	.77	.45
Writing Name & Reading Sentence	.72	.43
SQUARE WITH STICKS	.31ns	.15ns
SENTENCE REPETITION	.91	.91
SPIRAL MAZE 1	.89	.81
SPIRAL MAZE 2	.85	.84
PORTEUS MAZES	.80	.76
Short score	.92	.92

Pearson product-moment correlations; ns not significant at .05 level; a n=11.

CHAPTER 6

Further Analysis of Memory Impairment.

Memory impairment ranks with general intellectual decline as a cardinal feature of established dementia. Its existence has been demonstrated in countless studies. A variety of hypotheses exist concerning the nature of the impairment in organic amnesia. Many of these have been based on relatively 'pure' types of amnesic syndrome as seen in Korsakov patients and cases of hippocampal or dorso-medial thalamic damage; but there is no reason why ideas derived from them should not have relevance to amnesia occurring in the context of more widespread cognitive impairments, as in the dementias. Such hypotheses will now be briefly reviewed.

The consolidation hypothesis.

This is among the older theories, deriving from the ideas of Hebb (1949), which relates short-term memory to the reverberation of neural circuits and longer-term memory to permanent change in them. Consolidation is the process of change from the first to the second: Baddeley (1975) uses the analogy of a concrete moulding which is soft and easily breakable when newly made but which sets with time to become hard and very durable. Milner (e.g. 1968) has been the main exponent of a consolidation hypothesis, stimulated largely by studies of patients such as the hippocampectomised H.M. It is an intuitively attractive view which is consistent with a large amount of empirical evidence. It necessarily adopts what is now seen as a simplistic or out-dated view of 'short-term' and 'long-term' memory stores; its most striking problem as a comprehensive theory of amnesia is that, to explain the finding of very long retrograde amnesia in some patients (Albert et al 1979), consolidation of a memory would have to continue for many years. Findings such as enhancement of performance with retrieval cues and

the occurrence of prior-item intrusions in list-learning can be accommodated only with difficulty. The hypothesis is perhaps more applicable to semantic than episodic memory (Kinsbourne & Wood, 1975).

The trace decay hypothesis.

This is an extension of the consolidation hypothesis. The memory trace is said to decay spontaneously over time; the structural neural changes 'melt away'. Wickelgren (e.g. 1970) is the main contemporary exponent of this idea. There is a large body of evidence suggesting that this is not the basis of normal human forgetting (e.g. Underwood, 1957) and it is consistent with only a limited amount of the amnesic evidence. It is perhaps best thought of as a shorthand expression for the effects of other influences such as interference.

The inhibition of competing responses hypothesis.

Otherwise broadly known as the retrieval-interference theory, this asserts that the amnesic subject's difficulty results from a retrieval problem rather than a registration or storage one. The problem arises from an inability to inhibit competing irrelevant material or responses; the problem is not that the amnesic subject remembers too little but that he fails to suppress other contents of the memory store.

The concept of inhibition in memory has a long history. Ribot in 1881 (quoted in Talland, 1965, p 310) said that the total obliteration as well as the momentary suppression of a vast store of information are necessary conditions of remembering. Luria's account of his 'memory man' vividly illustrates the problems of a man doomed to a life of inability to forget things. Underwood

(1957) has stressed the importance of various forms of interference in normal memory.

More recently, the competing responses hypothesis has been most strongly propounded by Warrington & Weiskrantz (e.g. 1970). Their evidence is of two main types. Firstly, prior-item intrusions in recall experiments which use two or more lists are particularly common in amnesic subjects (Weiskrantz & Warrington, 1970; Baddeley & Warrington, 1970). A similar intrusion effect has been found by Winocur & Weiskrantz during learning of paired associates. Secondly, amnesics seem particularly helped by the provision of partial information at the time of retrieval, a procedure assumed to reduce the potential number of competing responses. Warrington & Weiskrantz (1970) presented amnesic and control subjects with word lists and after one minute tested retention using one of four procedures. Amnesic subjects were severely impaired compared to the controls with free recall and yes- no recognition testing, but not when the partial information methods - providing fragmented words or the first three letters of words - were used. Other studies (Warrington & Weiskrantz, 1968; Milner, 1970) showed comparable effects using non-verbal material. Isaacson (1975) cites animal work compatible with the hypothesis, commenting that "one of the fundamental roles played by the limbic system is the suppression of established ways of responding".

Criticisms of the theory concern the possible existence of partial information benefits in normal subjects (Woods & Piercy, 1974), the possible non-existence of them in amnesic subjects (Meudell & Mayes, 1981), and the inability of the theory to account adequately for temporal gradients in retrograde amnesia.

The semantic encoding hypothesis.

This states that, although they may be capable of doing so, amnesic subjects do not usually encode material semantically to the same extent as normal people do, but rely on encoding in terms of physical (i.e. graphic or phonemic) features. Butters & Cermak (1980) provide the fullest exposition of this hypothesis, within the framework of an information-processing model. Elaborative encoding (Craik & Lockhart, 1972) strengthens the memory trace; accurate retention is directly proportional to the depth of the encoding process. Amnesic subjects' processing is shallow, so their memory is poor. The reason for the shallow processing might involve motivational factors or impaired rate of search of conceptual semantic memory (Cermak et al, 1978).

This approach has been criticised on the grounds of circularity (Baddeley, 1978). The difficulty Butters & Cermak (1980) have had in demonstrating the beneficial effects of encouraging semantic encoding, in comparison with Warrington & Weiskrantz's demonstration of the benefits of providing partial information, is notable. The picture is further clouded by the fact that several studies of the semantic encoding hypothesis have used manipulation of cues, distractors, or levels of processing at the time of retention testing rather than at the time of learning: as Kinsbourne & Wood (1975) point out, effects produced by such manipulations are at least as compatible with a retrieval hypothesis. Studies by Kinsbourne & Wood (1975) and Gardner et al (1973) involving deliberate comparison of cueing at time of learning and at time of retrieval supported a retrieval hypothesis. The theory has problems in dealing with some non-verbal memory

impairments, and cannot account for retrograde amnesia. Its proponents do not claim that it is the sole deficit in amnesia, and they acknowledge the importance of interference effects. They suggest that the lack of semantic encoding maximises the accumulation of proactive interference and prevents release from this interference: memories are poorly 'insulated' from each other (Cermak et al, 1974).

The imagery hypothesis.

Baddeley (1975) adopts the proposal made by Paivio (1971) and Bower (1972) that semantic memory has two components, one linguistic and one based on imagery, and asserts that it is the imagery component which is impaired in amnesia. The crucial finding (Baddeley & Warrington, 1973) was that while amnesics are able to take advantage of word clustering based on taxonomic category membership they derive no benefit from clustering based on visual imagery (a technique which is very helpful to normals); but subsequent studies have produced inconsistent results. This hypothesis has much in common with the semantic encoding hypothesis, though the respective bases of the inadequate semantic encoding are different. Again, the influence of encoding and retrieval factors is somewhat confounded. Subsequent studies have produced mixed results. The imagery hypothesis is more easily adapted to account for non-verbal memory deficits, and the two hypotheses can to some extent be considered complementary; it is unlikely that an imagery hypothesis alone could account for a large amount of the evidence.

The contextual hypothesis.

This proposes that amnesic subjects have nearly normal acquisition and storage capacities but are defective in discriminating the spatial and temporal 'contextual cues' associated with different experiences. They cannot retrieve specific information because the context in which it was learned cannot be reconstructed, and they easily confuse recently presented material and information acquired in the distant past. Kinsbourne & Wood (1975) and Winocur & Kinsbourne (1978) are among the advocates of this view; Talland (1965) also considered such difficulties important. The hypothesis is consistent with much evidence, but can be criticised for being rather vaguely formulated. It can be re-interpreted in terms of encoding rather than retrieval defects and can be reconstructed in terms of Tulving & Thomson's (1971) encoding specificity hypothesis.

The familiarity hypothesis.

This suggests that at the root of amnesia is the loss of the ability to judge the degree of familiarity of an item, whether the 'item' is something the subject is required to recognise or is a potential response which requres checking before emission. It has much in common with the retrieval-interference hypothesis. Gaffan (1974, 1976) proposed it on the basis of studies of fornix-lesioned animals. Such animals were unimpaired in learning associations but highly impaired in tasks requiring discrimination of familiarity (defined simply in terms of how often a stimulus had been presented). Though the effects of particular limbic lesions are notoriously different in animals and man (Horel, 1978), such a hypothesis is consistent with several features of human amnesic

syndromes. The theory's generality depends on the assumption that recognition is a sub-process of recall, a debatable point (Brown, 1976).

The affective-motivational hypothesis.

This derives from the well-known fact that the limbic system, in which the typical lesions of 'pure' amnesic patients lie, plays a crucial role in the regulation of human emotion and affect (Papez, 1937); the hypothesis overlaps with Gaffan's. It was proposed by Mackay (quoted in Sweet et al, 1959, p80) and later by Simpson (1969) and asserts that the memory failure relates to a failure to attach affective importance or 'significance' to events or memories. More crudely, amnesic subjects are not motivated to remember. This is an appealing hypothesis but is rather imprecisely stated; it has no specific experimental support in humans and does not explain such phenomena as the benefit of partial information at retrieval test.

The premature closure of activation hypothesis.

This is based on the concept of memory search, and was originated by Talland (1965); memory search progresses by stages, each phase being terminated by some implicit act of recognition. Amnesic subjects terminate search prematurely, with an incorrect match; this early termination is related to a lack of affect or motivation. Talland presents data consistent with the hypothesis, but there is little direct experimental support for it (largely because it is very difficult to test). There is an interesting parallel with the competing responses hypothesis: Talland (1965, p318) states during a discussion of the role of the hippocampus that "the orderly operation of matching and information filing

would be disrupted by the invasion of messages that are normally suppressed". Mattis et al (1981) have proposed a related idea based on Hull's (1941) concept of reactive inhibition. They suggest that a disorder of this internally generated inhibitory process underlies Korsakov amnesia; the inhibitory process is too strong (in contrast to the Warrington & Weiskrantz idea of inadequate inhibition). Again, little direct evidence exists.

The cognitive map hypothesis.

O'Keefe & Nadel (1974) proposed this hypothesis on the basis of the huge body of evidence from studies of animals with hippocampal lesions. They suggest that the role of the hippocampus lies in constructing a cognitive map of the environment; if it is damaged the animal can still learn to associate stimuli and responses but is unable to form this spatial picture of its world. Despite some attempts to couch this in terms applicable to human amnesia, the link seems tenuous.

These hypotheses generally relate to episodic rather than semantic memory (Tulving, 1972). A recent perspective on the neuropsychological distinction between the two is provided by Warrington (1986). All the theories may have value in the sense of providing frameworks for research analysis. There is considerable overlap between them. None can account by itself for all the existing evidence, partly because even the relatively 'pure' amnesias may differ from each other (Lhermitte & Signoret, 1972; Iversen, 1977; Rose & Symonds, 1960). It has become clear that the stages of processing are so highly interdependent that the impairment of one must always be reflected in the impairment of others (Butters & Cermak, 1980); increasingly, syntheses of

encoding, storage, and retrieval hypotheses are being attempted (Wickelgren, 1979; Eysenck, 1979). Squire (1980) discusses the logical and practical difficulties involved in such formulations.

Memory changes in normal ageing and in dementia are reviewed by Miller (1977). Much of the evidence is consistent with more than one of the hypotheses described above. It seems likely that a number of different components of the memory process are disturbed in dementia, that the relative contributions of these disturbances may vary with the severity or stage of dementia, and that no single available hypothesis can account for all the findings. A comprehensive review of memory research in dementia is out of the question, and only studies directly relevant to the studies to be described will be mentioned. Studies on short-term or primary or working memory will be ignored except where relevant to impairments of long-term or secondary memory, which the following studies address. Such studies provide equivocal results concerning whether STM impairment exists in the early stages of the condition: at severe levels of dementia STM, like all other cognitive functions, is undoubtedly impaired or unassessable. Evidence is reviewed by Miller (1977). Similarly, the possibility of impairments in sensory or iconic memory will not be considered.

Studies in the literature, reasonably enough, generally compare amnesic subjects and normal controls. Apart from comparisons of dementia and depression, few studies compare one memory-disordered group with another. Corkin (1982) compared a DAT group with a group of 'pure' amnesic subjects (of mixed aetiologies) using a number of test of memory and learning. Some suggestions of different patterns of memory impairment emerged, but the main difference simply

concerned severity of impairment (with this particular DAT group showing very severe impairment). Suggestions that the memory impairments seen in depression are not identical to those seen with organic damage are reviewed by W Miller (1975) and Woods and Britton (1985, Ch4). Three studies of memory impairment in different organic amnesias (including DAT and MID) and depression will now be described, followed by brief and practical analyses of memory performance on some of the tests used in the main study.

Response Competition in Recognition Memory DAT, MID, and Korsakov's Syndrome.

Miller (1975) has confirmed the ability of cueing to improve performance in presenile dementia to normal levels, and Morris et al (1983) have replicated this in senile dementia. Davis & Mumford (1984) on the other hand failed to replicate this in SDAT, finding that semantic cueing at recall produced no significant benefit and that the benefit from initial-letter cueing was not significantly greater than that seen in control subjects. They concluded that their results were more compatible with a deficit at the acquisition stage. Miller (1977) also points out that the cueing effects might have something to do with subjects' encoding strategies and notes that subjects do not produce an excess of wrong words in free recall tasks (as the disinhibition of competing responses hypothesis might predict). Data concerning the occurrence of irrelevant or prior-list intrusions in list-learning experiments have also been equivocal as regards the hypothesis. Miller (1978, Expt 2) developed another elegant paradigm to test the competing responses hypothesis: a recognition memory task using learning of

word lists where at retrieval each target word appeared in a choice array of two, four, or eight words. He reasoned that his DAT subjects should be particularly disadvantaged (compared with normals) by larger numbers of response alternatives, and this was what he found.

In view of the diversity of amnesic syndromes (as reviewed by Hirst, 1982), the present study was designed to compare three amnesic conditions - alcoholic Korsakov's (K), DAT, and MID - using a variant of Miller's design. Differences exist in the locations and types of anatomical damage in the three conditions (relatively characteristic in K and DAT and variable in MID) as well as in their clinical characteristics. One might therefore expect the nature of the memory impairment to differ between the groups. This study stemmed in part from a previous small study (Taylor, 1981) comparing Korsakov patients with a mixed group of elderly memory-impaired subjects. The data on the K patients is retained, but here comparison is made with equal numbers of patients with DAT and with MID. Comparison of the three groups with two other subject groups - normal elderly and depressed elderly - was also attempted: unfortunately this had to be abandoned, for reasons described later.

Subjects.

The subject groups consisted of 12 chronic hospitalised K subjects (10M, 2F; Mean age 59.1 yrs, range 47-69), 12 DAT subjects (9M, 3F; 75.5 yrs, 64-82), and 12 MID subjects (9M, 3F; 75.1 yrs, 65-83). All the K patients were hospital in-patients, and had been so for at least a year. All had histories of alcoholism, but had not (as far as is known) been using alcohol since admission. In all

cases a Consultant Psychiatrist's diagnosis was of Korsakov's attributed to the alcoholism. None had any known features of a more global dementia in their histories or at interview: the author did not formally test this himself, but another psychologist had done so some months before as part of another study on Korsakov's syndrome. The DAT and MID subjects were diagnosed in the usual way. Most were day hospital patients, and none had known histories of alcohol abuse. No case was included in any group where there was doubt about the diagnosis. As ever, the possibility of misclassification cannot be completely ruled out in the absence of post mortem information. No subjects had histories of previous psychiatric disturbance, and none had had ECT. No subjects had uncorrected impairments of sight or hearing: straightforward clinical tests of vision and hearing, including reading and comprehension, were carried out before testing. Patients with prominent dysphasic difficulties or any dyslexic problems were excluded as valid testing was not feasible in such cases. Two members of each group were on some psycho-active medication (generally small doses of major tranquillisers).

Method.

After introductory conversation and an explanation of the nature of the tests, each subject was tested on twelve brief recognition memory tasks: three each of four types of test as follows.
(1) A prose paragraph was read out to the subject. After 30 sec. of general conversation (to prevent rehearsal), he or she was asked a series of six questions about it, in each case a card being shown with the correct answer amongst one, three, or seven plausible but incorrect filler answers (all in lettering 1 cm high). The subject

was required to choose the correct answer, and to guess if he did not know.

(2) A list of six common words (Thorndike & Lorge category AA) was read to the subject at a rate of one every 2 sec. After 30 sec of conversation he was shown a series of six cards, each with one target word and one, three, or seven fillers drawn from the same pool (again in 1 cm lettering). The subject was asked to choose which word had been read out, and to guess if he did not know.

(3) A series of six meaningful photographs (5 by 6 cm, taken from colour magazines) was shown at a rate of one every 2 sec. After 30 sec of conversation the subject was required to choose which picture he had seen from each of six arrays (the one, three, or seven filler pictures being taken from the same source) and to guess if he did not know.

(4) This was as (3), but using meaningless designs.

Precise details of the test materials appear in Appendix 2. Three 'sets' (a,b, & c) of each of these four types of test were used (to collect enough data without making any single test too long or difficult): i.e. three paragraphs, three lists of 6 words, three sets of 6 photos, and three sets of 6 designs. They were presented in a cyclical fashion (1a, 2a, 3a, 4a, 1b, 2b, etc). Any given target item appeared in a two-choice array for a third of each subject group, a four-choice array for a third, and an eight-choice array for a third. The order (from 1 to 6) in which two-, four-, and eight-choice arrays appeared was also appropriately counterbalanced. The order of appearance of the target items in the choice arrays (from 1 to 6) was the same as their initial order of presentation. The physical position of correct choices in arrays

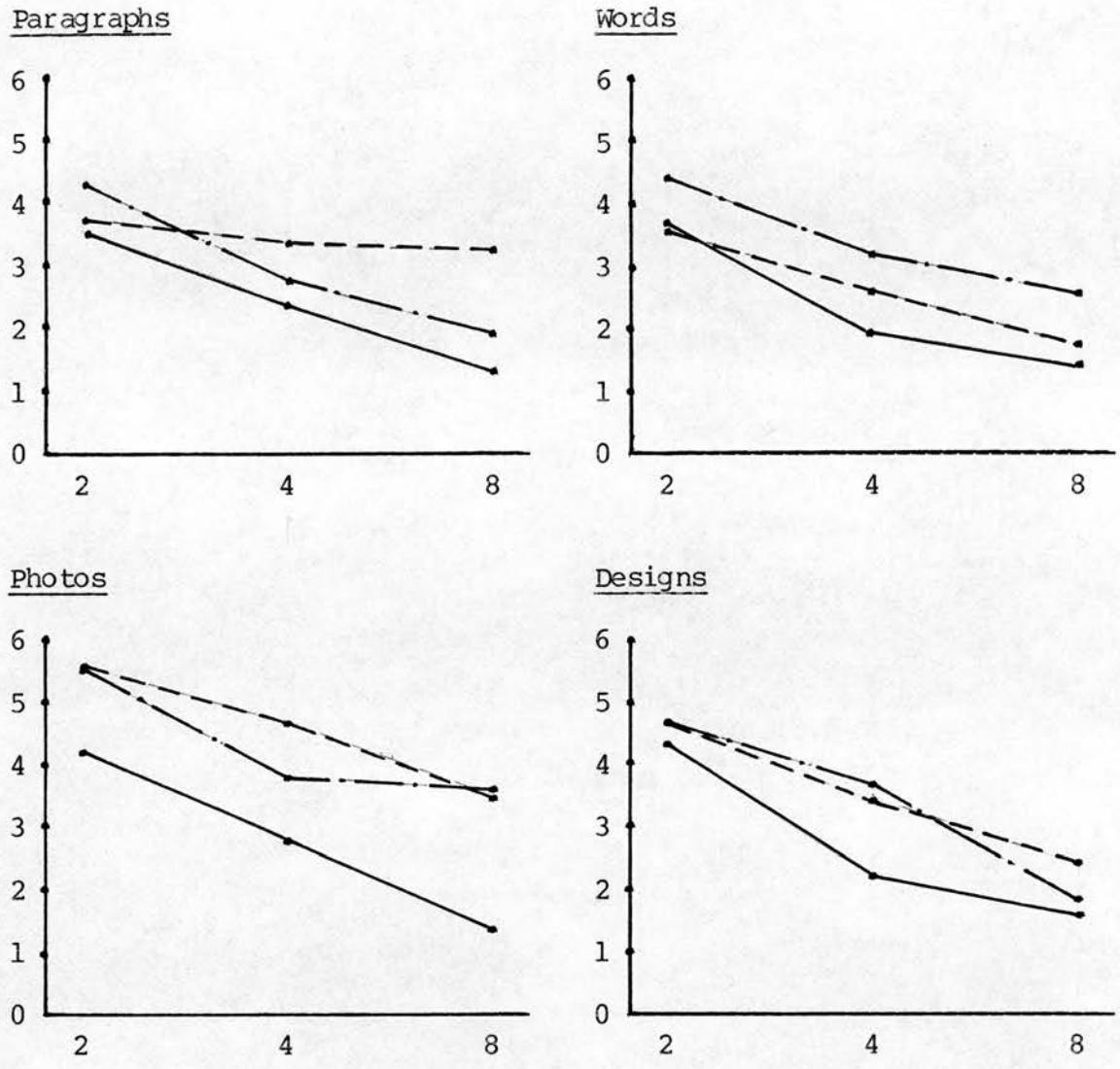
was randomised. No subject ever saw the same filler item twice.

Results and discussion.

The mean numbers of correct choices (possible score 0 - 6) in each condition for the three groups are shown in Figure 1. All the data were analysed using a repeated-measures analysis of variance with diagnostic group as the between subjects factor and type of material and number of response alternatives as the within subjects, repeated-measures factors (Winer, 1971, p539 onwards). A smaller repeated-measures analysis of variance (Winer, 1971, p518 onwards) was also carried out for each of the four types of material separately. To summarise: The DAT group happened to perform rather more poorly than the other groups, and Photos was an easier test than the others. The effect of number of response alternatives (NR) was, not surprisingly, always highly significant. The interaction effects of interest - those involving subject group by NR - were not significant (though there was some trend in Paragraphs for the MID group to be less affected than the other groups by increasing NR).

Obviously, a fall in scores from the 2 to the 4 to the 8-choice condition would be expected on the basis of chance alone. To give an idea of the patterns of performance when chance is taken into account, chance-corrected scores were calculated according to the formula: Chance-corrected score = $(Po - Pc)/(1 - Pc)$ where Po is the subject's obtained proportion correct and Pc is the proportion expected correct by chance. Hence chance performance gives a corrected score of 0, perfect performance 1. Mean chance-corrected scores in each condition are graphed in Figure 2. An alternative transformation uses the d' measure from signal

Figure 1 Mean scores of the three groups under the three recognition conditions for each type of material (n=12 for each point plotted).



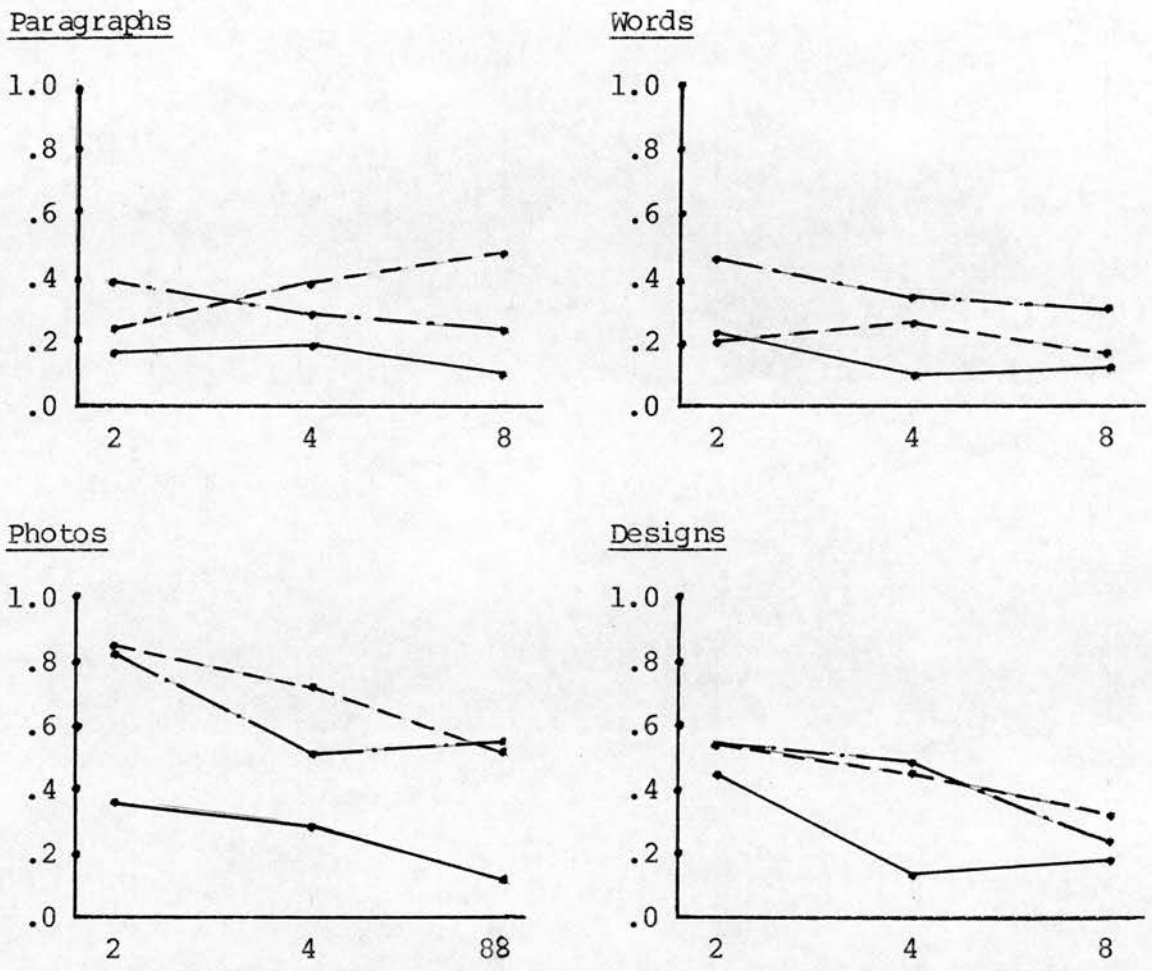
Abscissa shows recognition condition; ordinate shows score.

K

DAT

MID

Figure 2 Mean chance-corrected scores of the three groups under the three recognition conditions for each type of material (n=12 for each point plotted).



Abscissa shows recognition condition; ordinate shows chance-corrected score.

K —·—

DAT ——

MID - - - -

(Scores corrected for chance according to the formula CCS = (Po - Pc)/(1 - Pc) where Po is subjects' obtained proportion correct and Pc is proportion expected correct by chance. Hence chance performance gives a CCS of 0, perfect performance 1.)

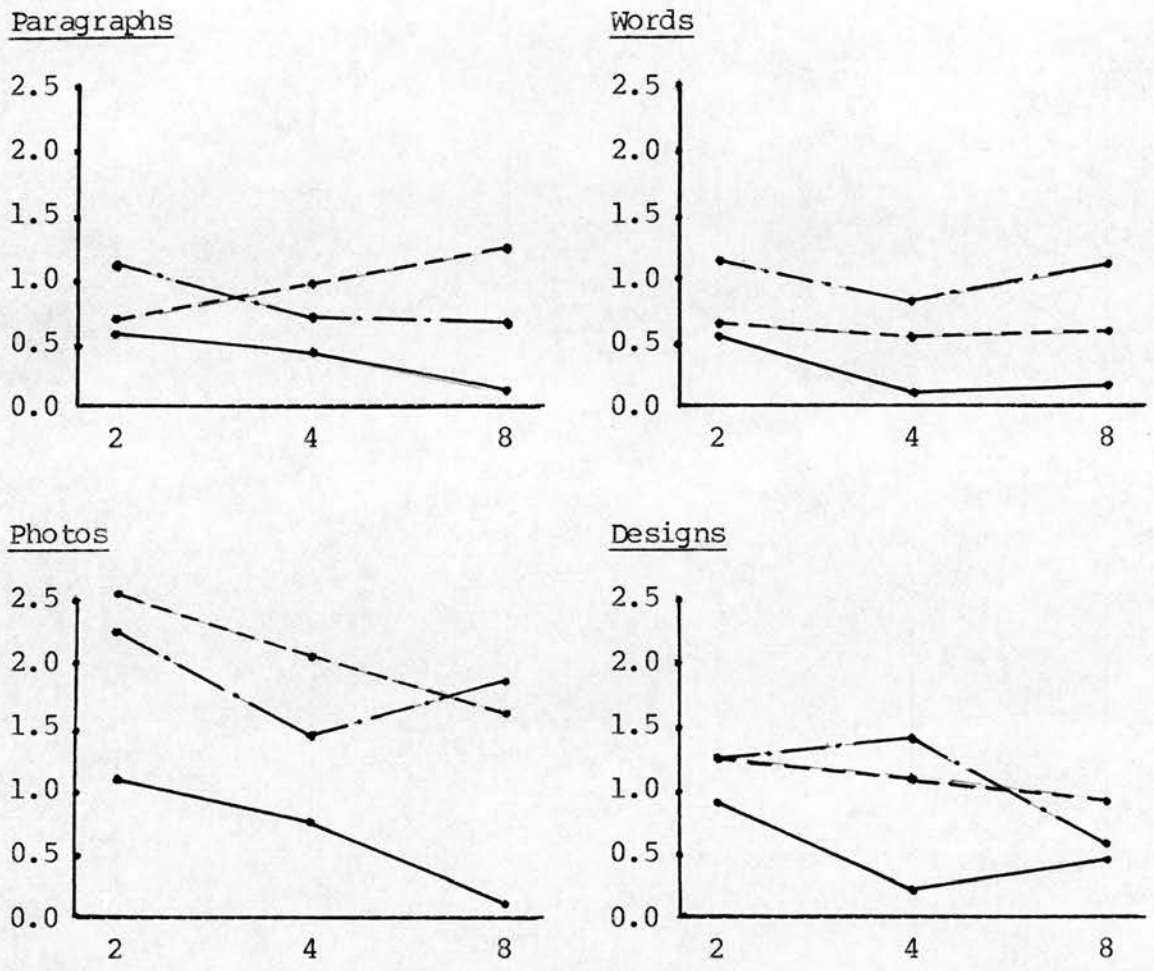
detection theory (as discussed in Meudell & Mayes, 1981). Mean scores using this transformation are graphed in Figure 3.

The analyses were repeated using the chance-corrected scores and then using the d' scores. The outcome was largely unchanged, the point of note being that the main effect of NR was obliterated in the two verbal tests but remained in the two non-verbal ones. The significance of the various F ratios using the three types of score (raw, chance-corrected, d') are shown in Table 1.

To investigate other relationships in the data, difference scores for each type of material for each subject were calculated: score in 2-choice condition minus score in 4-choice; 4-choice minus 8- choice; and 2-choice minus 8-choice (= the sum of the first two). Hence 12 difference scores were calculated for each subject. Correlational analyses (in each group separately, and in all three combined) indicated no consistent relationship between overall levels of performance and the extent of the decline in performance as NR increased. Similarly neither the overall levels nor the extent of decline correlated with age in any group or in the three groups combined. The only apparent ceiling effect was in the MID group in Photos two-choice, and this does not distort the results of the analyses in an important way.

(The main effects in the study can also be analysed using these difference scores. The difference scores in the three groups were compared using a series of 12 ANOVAs. Only one of the 12 produced a significant result: the drop in performance between the 2-choice and the 8-choice conditions in Paragraphs is smaller in the MID group than in the other two groups, confirming the trend noted above.)

Figure 3 Mean d' scores of the three groups under the three recognition conditions for each type of material (n=12 for each point plotted).



Abcissa shows recognition condition; ordinate shows d' score.

K —·—·—

DAT ————

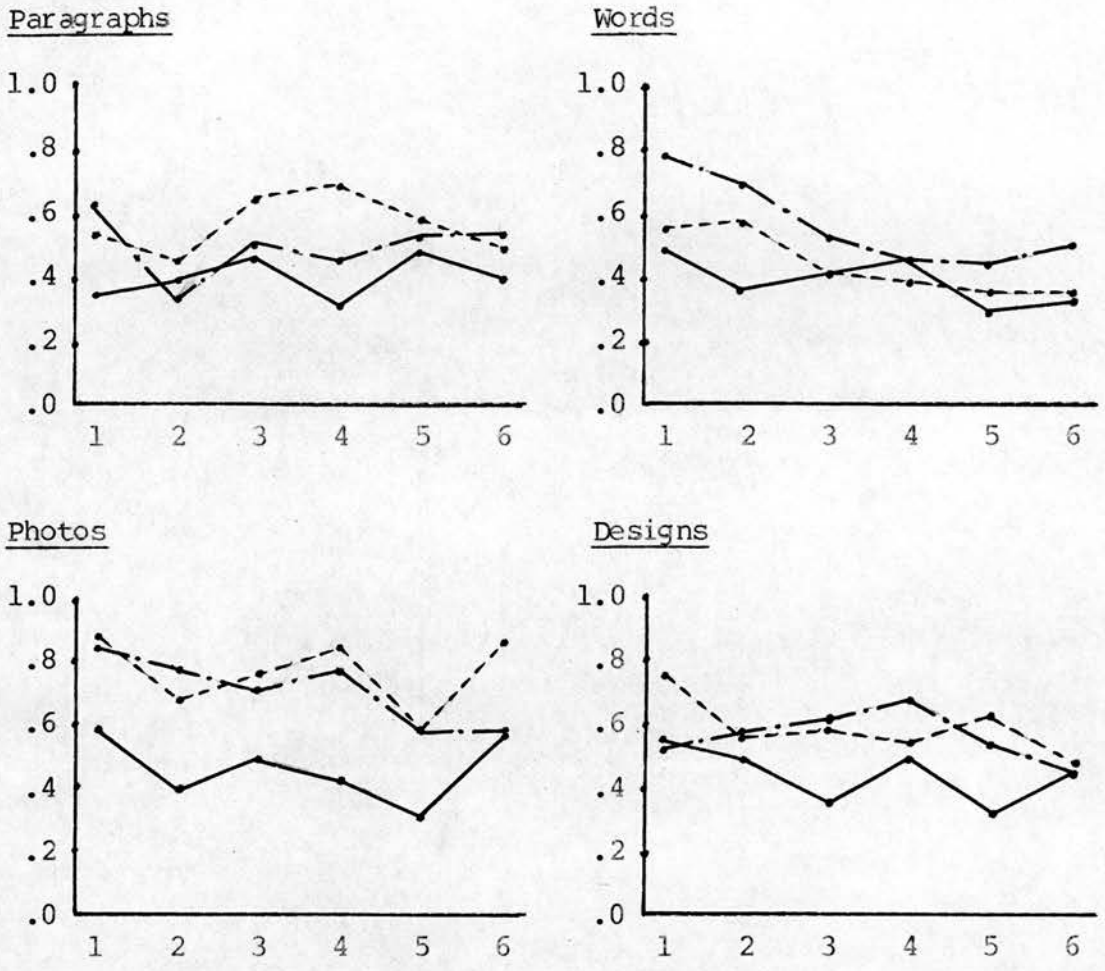
MID - - - -

Table 1 Significance of the F ratio in the various analyses.

	Significance of F ratio		
All Data	Raw	Chance-corrected	d'
S group (A)	.0027	.0052	.0058
Type of material (B)	.0000	.0000	.0000
A X B	.0216	.0388	.0298
No. of choices (C)	.0000	.0004	.0065
A X C	.0661	.4014	.7303
B X C	.4816	.0827	.4385
A X B X C	.3902	.6669	.5556
Paragraphs			
S Group (A)	.1297	.2328	.2840
No. of choices (B)	.0000	.9413	.9167
A X B	.1177	.3014	.4103
Words			
S Group (A)	.0695	.1059	.0772
No. of choices (B)	.0000	.6166	.5419
A X B	.8008	.8762	.9506
Photos			
S Group (A)	.0000	.0000	.0001
No. of choices (B)	.0000	.0002	.0036
A X B	.2750	.5974	.4510
Designs			
S Group (A)	.0437	.0629	.1113
No. of choices (B)	.0000	.0007	.0752
A X B	.1368	.3378	.3149

Primacy and recency curves for the three groups for each type of material are shown in Figure 4. (2,4, and 8-choice data are of course combined, but this does not distort the picture since the position of 2,4, and 8-choice tests was counterbalanced.) The

Figure 4 Serial position curves of the three groups for each type of material (n=12 for each pont plotted).



Abscissa shows serial position; ordinate shows proportion correct.

K —·—

DAT ——

MID ·-----

curves are far from smooth, but at least there is little sign of recency effects (and in some areas there are primacy effects), which is consistent with the interval between presentation and test having been adequate to test 'long-term' memory.

Inspection of individuals' performances suggested that the subjects who were on some psycho-active medication showed patterns typical of their groups; the numbers are so small that statistical confirmation of this is not feasible. Similarly, inspection revealed no obvious sex differences in patterns of performance, but again the numbers of women are too small to allow statistical analysis.

Comparison of the three groups with normal elderly and depressed elderly subjects was abandoned because the tests were too easy for these intact subjects: those tested scored five or six in all conditions. This ceiling effect would have led to a spurious difference between the patterns of performance of the impaired and the intact groups. Consideration of Miller's (1978, p146) graph suggests that a ceiling effect in the two-alternative condition might have occurred in his study also, though Meudell & Mayes (1981) argue against this interpretation on the basis of their signal detection analyses.

Despite the differing neuropathological bases and clinical features of the three conditions, little difference between the groups emerged in this study regarding the importance of failure to inhibit competing responses. The extent to which <u>all</u> the groups were suffering from such a failure cannot be judged here in the absence of normal control data. The avoidance of ceiling effects in unimpaired Ss as well as floor effects in amnesic subjects is

problematic, but controls must have 'room for improvement' just as amnesics do under the easiest or most helpful experimental manipulation if conclusions are to be drawn about essential defects in amnesia. Squire (1980) discusses whether amnesic subjects are uniquely benefited by certain kinds of prompts given at the time of the retention test, and concludes that they are not. Similarly, Meudell & Mayes (1981) found with two- vs. eight-choice word recognition that normals after long retention intervals showed similar patterns of performance to those seen in amnesic subjects after short intervals. Using a signal detection analysis they also found that d' scores were no worse with eight than with two choices: they suggest that a failure of response inhibition may not be present and that the concept of 'weak memory' is just as compatible with the findings. In the present study, a d' analysis removed the effect of increasing NR only in the two verbal tests. The study suggests that a particular recognition memory paradigm which is presumed to involve the inhibition of competing responses shows little difference between three amnesic groups, but that failure of response inhibition may nevertheless be present at least with certain types of material.

Note: The 12 Korsakov patients and 6 of the patients in the groups with dementia were tested towards a thesis submitted for a previous degree (Taylor, 1981). Parts of the review of theories of amnesia at the beginning of this chapter are adapted from that thesis. A paper based on this study has been published as Taylor R & McGuire R J (1985) Response competition in recognition memory in three amnesic groups. British Journal of Clinical Psychology 24: 145-147. This appears in Appendix 6.

Encoding Preferences in DAT, MID, and Depression.

No attempt was made in the previous experiment to manipulate the degree of acoustic or semantic similarity of target and filler items. This study considers aspects of encoding.

A basic distinction is allowed within the encoding hypothesis between the ability to use deep encoding strategies and the tendency to use them spontaneously. Experimental manipulations which force subjects to process material deeply or which provide appropriate cues at retrieval can improve performance of amnesic subjects (though not always to normal levels), suggesting that the subject's tendency or encoding preference plays a part as well as any actual incapacity to use deep levels of processing. Moscovitch (1982) and Perlmutter & Mitchell (1982) suggest that the tendency is much more affected than the capacity in normal old people. This might well also be the case in dementia or depression.

Weingartner et al (1982) found that subjects with DAT were not helped by the presence of semantic relationships between stimulus words in a free recall task: unlike controls, they appeared to fail to make use of the potential for semantic organisation or clustering. These authors also found that the patients with dementia did poorly on verbal fluency tests: they interpreted this as further evidence of a deficit in accessing semantic memory. The verbal fluency measures correlated strongly with the memory measures, though of course this does not in itself imply that both types of test impairment should be attributed to the same underlying deficit.

Corkin (1982) found that DAT subjects, unlike normal subjects, were not helped in recalling words by having previously made

semantic judgements (as opposed to sensory or phonological judgements) about the words. Again this was interpreted as indicating a semantic encoding impairment.

Miller (1977, p58) describes an experiment comparing DAT and normal subjects using recognition memory for words where recognition alternatives at testing included a semantic, an acoustic, and an unrelated distractor as well as the target word. No differences in the pattern of errors made by the two groups were found, though the normal subjects presumably made fewer errors. The assumption in this type of experiment is of course that occurrence of a disproportionate number of acoustic as opposed to semantic errors reflects shallower processing.

Warren & Groome (1983) used a recognition memory for words test where at testing subjects were required to select the words they had been given from a continuous list containing semantic, acoustic, and unrelated distractors as well as the target words. Subjects included three brain-damaged groups and a depressed group as well as normal controls. Overall performance levels differentiated between some of the subject groups, but differences were not found in error patterns. However the numbers of error choices made was very small in some groups, presumably reflecting a generally low false positive rate.

Larner (1977) tested memory for words using a continuous recognition format in demented, depressed, and physically ill control subjects. He interpreted some of his results in terms of an encoding breakdown present in both dementia and depression (but more severe in dementia).

Some feature of encoding ability or tendency may therefore

contribute to memory diificulties in some cognitively impaired groups. This study assesses the possibility that this occurs in some elderly memory-impaired groups more than others. It uses a similar task to that of Warren & Groome (1983) except that forced-choice rather than yes-no recognition testing is used in an attempt to overcome the problem of low false positive rate.

Subjects.

The subject groups comprised 22 DAT subjects (8 M, 14 F; Mean age 74.9 yrs, range 60-84), 22 MID subjects (10 M, 12 F; 75.7 yrs, 57-87), and 19 elderly depressed subjects (6 M, 13 F; 74.6 yrs, 60-86). The DAT and MID subjects were diagnosed in the usual way. Most were day hospital patients. The depressed subjects were selected on the basis of a consultant psychiatrists's current diagnosis of depressive illness (alone, with no secondary diagnoses); any suspected of having any organic dementia were excluded. Most of the depressed subjects were day hospital patients, with a few in-patients included; from this it may be assumed that most were probably not suffering from severe depression. Some of the depressed subjects had had ECT in the past, but not within the month prior to testing. Otherwise, inclusion and exclusion criteria were exactly as described for the previous study. 11 members of the DAT group were on some psychoactive medication (hypnotic 2, Major tranquilliser 6, Combination of two 3), 14 of the MID group (Minor tranquilliser 1, Major tranquilliser 10, Anti-depressant 2, Combination of two 1), and 16 of the depressed group (Hypnotic 1, Minor tranquilliser 1, Anti- depressant 8, Combination of two - usually anti-depressant plus tranquilliser - 6).

Method.

After an introductory conversation, each subject received two brief tests of recognition memory for word lists. First it was explained to the subject that he or she would be given a list of ten words and that he was to concentrate on the words and try to remember them, because afterwards a choice of words would be shown and he would have to pick out the words which had been given. Nothing else was said or done which might encourage encoding of the words in any particular way. Once this was understood, ten words were presented, each written on a flash card in 1 cm lettering, at a rate of one every 2 sec. Each word was spoken aloud by the tester as the card was shown. After the tenth card, the subject was engaged in general conversation for about 30 sec in order to prevent rehearsal. He was then presented with a card showing the the first word in the target list amongst three filler words: one with a similar meaning to the target (semantic distractor), one with a similar sound to the target (acoustic distractor), and one not related to the target in any way (unrelated distractor). For example, the first target word, Chair, appeared among Seat, Share, and Hand. The subject was asked to choose which one of the words had been given, and to guess if he did not know. The other nine, similarly-constituted, cards were presented in the same way. A second list of 10 words and 10 choice cards was then presented in an identical manner. (Two lists of 10 words were used rather than a single list of 20 to collect enough data without making any single test too long or difficult.) The order of appearance of the target words in arrays (from 1 to 10) was the same as their initial order of presentation. The physical position of target words and the

three types of distractors in the 2 X 2 arrays was appropriately counterbalanced, so that tendencies to choose a particular position on the card could not bias the results. No subject ever saw the same distractor twice. The actual words used are listed in Appendix 2. They are the same as those used by Warren & Groome (1983) in their continous- recognition task, excluding half of the words in their 'unrelated distractors' category; this half in fact consisted of 'list' distractors, i.e. words acoustically similar to a previous semantic distractor, and these did not seem particularly relevant to the present study. (The author is grateful to Dr. D H Groome for kindly providing the word lists and additional information about his study: The target words and semantic distractors were selected from the Palermo & Jenkins (1964) word association tables. Acoustic distractors were words which rhymed with target words. Care was taken to ensure that no acoustic or semantic similarities occurred between any of the words other than those specified by the distractor conditions. All words were monosyllabic and had a frequency of at least 20 per million according to the Thorndike & Lorge (1944) tables. Since word frequency is known to have marked effects on recognition, the different categories of words were approximately matched in this respect.)

In the present study, no attempt was made to introduce a category of 'physical' distractor, i.e. words with similar shape and size to the target words when written down: this would have been feasible given the inclusion of visual presentation here, but might have served only to confuse the results, given that many of the other types of distractors are physically similar to the target

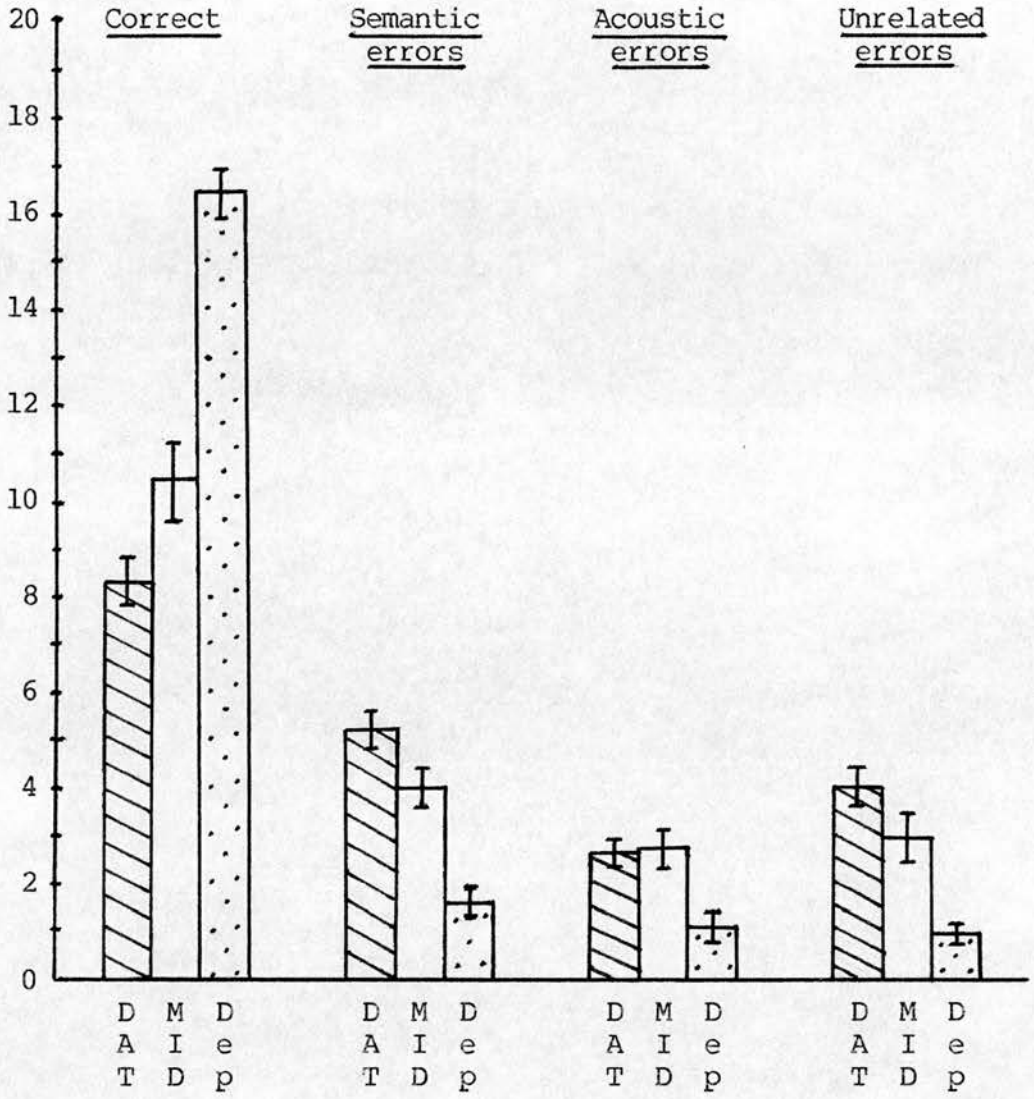
words in any case.

Results and discussion.

Mean numbers of correct choices and of the three possible types of error are shown in Figure 5. (Data from the two 10-word lists are combined.) Results of one-way ANOVAs for each type of choice appear at the foot of the figure. The depressed subjects clearly made more correct choices and fewer of all types of error than did either group of dementing subjects. This indicates that, in these particular samples of depressed and dementing subjects, a straightforward quantitative measure of level of memory performance discriminates well between depression and dementia, which in itself is not surprising. The depressed subjects were perhaps not sufficiently depressed to show major memory impairment. Patients with cognitive impairment sufficiently severe to raise the question of organic dementia were excluded from the depressed group; and the dementing groups contained only definite cases of dementia, some of whom obviously had considerable memory impairment. As regards diagnostic discrimination, of course, this means that the most interesting subjects - those with possible depression or dementia (or both) - are excluded. Woods and Britton (1985, Ch4) conclude that while impairments of memory and learning can occur in depression, they are not invariably present. Hence an alternative possibility here is that, rather than being insufficiently depressed, the depressed subjects simply did not all have memory impairment as a prominent symptom.

The more interesting aspects of the data concern the relative numbers of of the different types of error in the three groups. Simple inspection of Figure 5 does not suggest major differences

Figure 5 Mean numbers of correct choices and of semantic, acoustic, and unrelated errors made by the three groups.



ANOVA result (dfs 2,60)

F =	43.5	26.9	8.5	14.0
sig.	.0001	.0001	.001	.0001

Abscissa shows subject group and type of choice; ordinate shows mean number of choices; bars show standard errors.

between groups in pattern of error performance. This was investigated statistically by first calculating various error ratios for each subject:

1. (No. of semantic errors + 1) / (No. of acoustic errors + 1)
2. (No. of semantic errors + 1) / (No. of unrelated errors + 1)
3. (No. of acoustic errors + 1) / (No. of unrelated errors + 1)
4. No. of semantic errors / Total no. of errors
5. No. of acoustic errors / Total no. of errors
6. No. of unrelated errors / Total no. of errors

(The addition of 1 to each of the first 3 denominators was to avoid zero denominators; to avoid distortion as far as possible, 1 was also added to the numerators.)

One-way ANOVAs were carried out separately for each of the 6 ratios, broken down by diagnostic group. The only significant result (at the .05 level) was with ratio 6: this ratio was higher in the DAT group than in the MID group and lowest of all in the Depressed group. In other words the DAT group were most likely to make errors completely unrelated to the target words, and the Depressed group least so. These unrelated errors may be considered almost 'chance' errors, and the result may be interpreted in terms of the DAT subjects being more likely not to process words in any significant way at all (not even acoustically). The test design does not allow specific interpretation in theoretical terms of whether this defect is more one of encoding or attention, retention, retrieval, or some complex combination of these.

Miller (1972) and Morris (1984) suggest that the impairment of primary memory in DAT may be related to inefficient acoustic encoding, on the basis of findings that control subjects are more

disadvantaged than are DAT subjects by phonemic similarity in verbal material to be remembered. Wilson et al (1983), using a variant of the usual free recall of word lists paradigm, found not only an impairment of primary memory compared to the control group but a positive correlation between primary and secondary memory in the DAT group which was not present in the control group of normal elderly subjects. Wilson et al interpreted these results in terms of inattentiveness in DAT patients, and suggested that at least some of the well-documented impairment of secondary or long term memory could be attributed to this inattentiveness and primary memory impairment. Miller (1971), however, tried to reinstate a missing primacy effect in DAT patients in free recall of word lists by slowing the presentation rate of items (thereby presumably avoiding overloading of primary memory): this was not successful, suggesting that here the observed impairment of secondary memory could not be explained in terms of primary memory deficit. Of course, the rather vaguer concept of inattentiveness could still have been applicable. Attention failures can obviously be assumed to limit the possibility of any deep or any other kind of processing, and this may account for the excess of unrelated errors in the DAT group in this study.

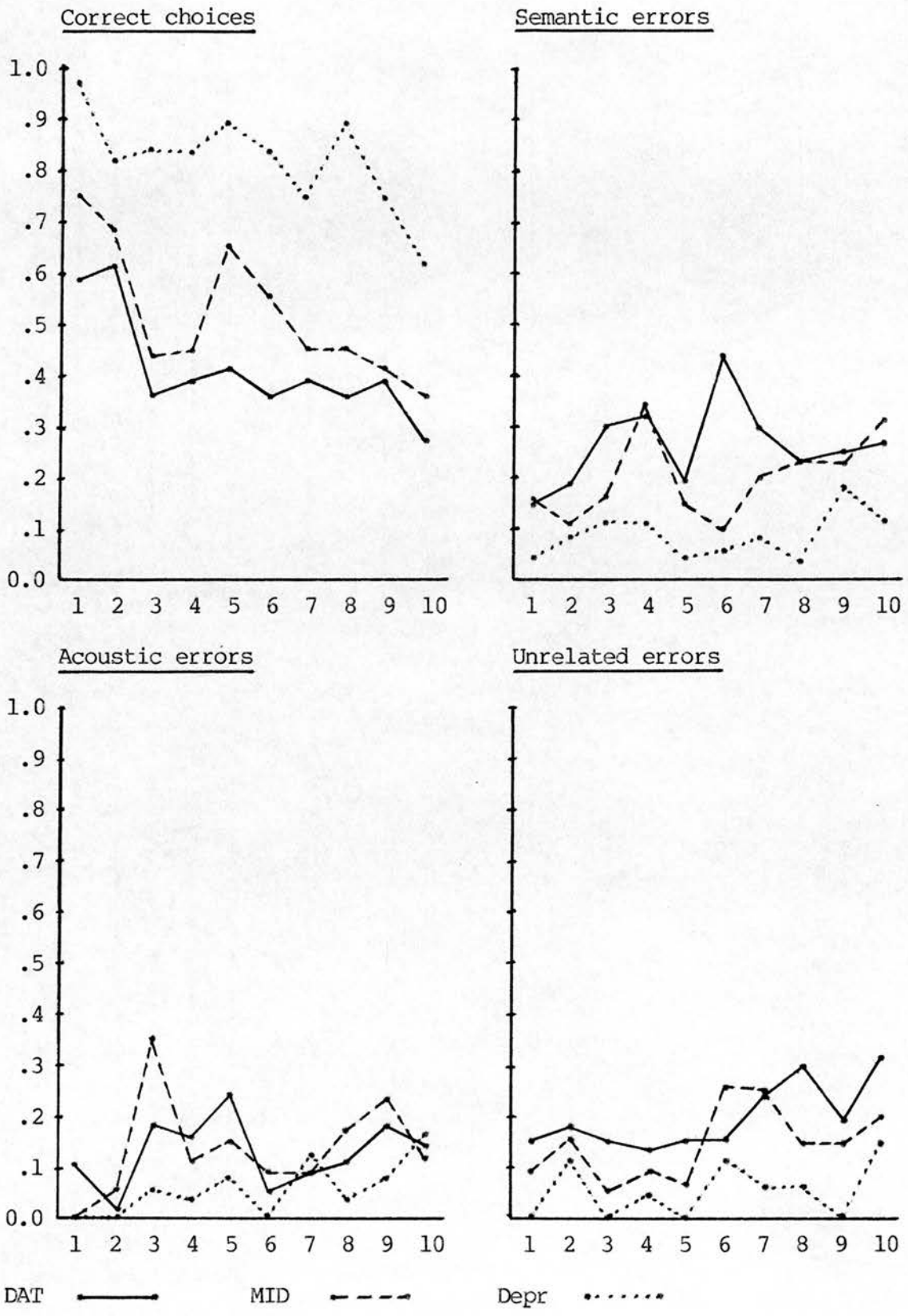
Intercorrelations (Pearson product-moment) between error ratios (as described above), total correct score, and age, were calculated in each subject group and in the three groups combined. Total correct scores correlated negatively with age except in the depressed group. A significant correlation between ratio 6 and total number correct appeared in the DAT group alone (i.e. in the group with the highest levels of ratio 6): hence not only did the

DAT group as a whole tend to make the most unrelated errors, but the more impaired DAT subjects showed this tendency most strongly. The other correlations were unremarkable.

One further possible area of difference between the patterns of performance in the three groups lies in primacy and recency effects. Figure 6 shows the distibutions of correct choices and types of error according to serial position (i.e. order of presentation and hence of testing) from 1 to 10 in the three groups. Data from the two 10-word lists are combined. The curves for correct choices indicate some primacy effect but no recency effect in all three groups, suggesting that the test procedure was adequate to test long-term rather than short-term memory; there are no signs of any consistent differences between the groups. Similarly, there appear to be no consistent differences between the groups as regards the curves for semantic, acoustic, or unrelated errors. The occasional peaks and troughs in some of the curves probably relate to certain types of errors being inherently easier to make over certain test words (or pairs of words, since the data from two lists are combined).

Finally, brief analyses were carried out concerning sex differences and drug use differences on the various test variables by means of t-tests (Males versus Females; on psychoactive medication versus not) in each subject group separately and in the three groups combined. Almost no significant differences emerged. In the MID group, those subjects on psychoactive medication made more semantic errors than those on none: there is no obvious explanation for this isolated finding, and it may be attributable to chance in view of the number of comparisons made by t-test. In

Figure 6 Primacy and recency curves for correct choices and for semantic, acoustic, and unrelated errors in the three groups.



Abscissa shows serial position; ordinate shows proportion of choices.

the Depressed group, the males happened to have lower total correct scores than the females.

Hence there appear to be no artefactual reasons for the relative lack of significant differences between groups in this small study.

A Signal Detection Analysis of Recognition Memory in DAT, MID, and Depression

Miller & Lewis (1977) compared demented, depressed, and normal control subjects using a continuous-recognition memory-for-designs test. The depressed and normal groups did not differ from each other in sensitivity: both were significantly more sensitive than the dementia group. The depressed group had significantly higher response criteria than either of the other groups. Criteria in the demented group were not significantly higher than those in the normal group: subjects with dementia made many false positive as well as false negative errors. These results suggested that the test impairments sometimes observed in depression might be partly attributable to response biases rather than to fundamental memory deficits.

Larner (1977) used a continuous-recognition word-memory test to compare demented, depressed, and physically ill control subjects. Depressed subjects showed conservative performance with relatively few errors, while the subjects with dementia made more errors partly because of their more liberal response criterion and partly because of their poorer encoding abilities or strategies.

Interestingly, Ferris et al (1980) found no difference between DAT subjects and normal elderly controls in any aspects of a signal

detection analysis of performance on a face-recognition test. These older groups differed from young controls in sensitivity but not response criterion, which is consistent with the findings of Harkins et al (1979) concerning memory for words in young and old normal women. Wilson et al (1982) did find a significant impairment in facial recognition memory in DAT subjects compared to controls, though the authors did suggest that this facial memory impairment was perhaps less severe than the verbal memory impairment found.

Whitehead (1975), using recognition memory testing with both verbal and pictorial material, found that DAT subjects were better on forced-choice recognition than on yes-no recognition. She suggested on the basis of an analysis of errors that this was not simply because the forced-choice format reduced the possibility of false positive responses, but rather because it overcame subjects' hesitancy or uncertainty as regards responding positively when unsure. However there were no control data for comparison, and the tests were not matched for difficulty: the forced-choice form may have been inherently easier. Miller & Lewis (1977) found a slightly higher response criterion on average in DAT compared with controls, but this difference was not significant partly because of the variability within the DAT group.

A lowering of response criterion might be partially explained in terms of disinhibition of competing responses. One might also speculate that response criteria might vary with severity of dementia (mildly impaired subjects with insight into their problems being more cautious), anxiety, frontal lobe dysfunction, and so on.

Strategy rather than ability per se may therefore be important in the memory performance of some cognitively impaired groups. This

study compares the extent of this phenomenon in two types of dementia and in depression.

Subjects.

The subject groups comprised 21 DAT subjects (8 M, 13 F; Mean age 77.7 yrs, range 60-86), 19 MID subjects (9 M, 10 F; 77.8 yrs, 65-92), and 19 elderly depressed subjects (6 M, 13 F; 74.7 yrs, 60-86), diagnosed in the usual way, with inclusion and exclusion criteria exactly as described for the previous study. Most were day hospital patients. 5 members of the DAT group were on some psychoactive medication (Hypnotic 2, Major tranquilliser 3), 11 of the MID group (Major tranquilliser 6, Anti-depressant 3, Combination of two 2), and 14 of the depressed group (Hypnotic 1, Major tranquilliser 1, Anti-depressant 6, Combination of two - usually anti-depressant plus tranquilliser - 6).

Method.

After an introductory conversation, each subject received two continuous-recognition memory tests. The first used meaningful photographs taken from multiple copies of colour magazines; the second used meaningless designs. (Details of the materials appear in Appendix 2.) The construction of the two tests was similar. Each consisted of a set of 100 cards, comprising 5 identical sets of 8 target items, and another 60 completely different filler items (taken from the same pool of items). Each fifth of the pile of 100 contained a set of 8 target items amongst 12 fillers. In each of these divisions of 20, the 8 target items appeared in the same order (from 1 to 8) but not in the same positions relative to the fillers: i.e. the order of appearance of target items compared with fillers was randomised within each division of 20. (These

'divisions' of 20 refer simply to the construction of the tests: they would not be apparent to a subject.)

The Photographs test was always given first. The subject was told that the pictures would be shown one at a time, and that he or she was to concentrate on them and try to remember them; the first 20 would all be different, but after that some 'repeats' would come up and he would be required to identify them. Once this was understood, the cards were presented at a rate of 1 every 3 sec. After the first 20 the tester began asking whether the subject had seen each picture earlier in the pile or not, requiring him to guess if he was unsure. Feedback on whether the subject was right or wrong each time was avoided lest the tester might be unconsciously shaping the pattern of 'yes' and 'no' answers, though appropriately encouraging remarks were made at the end of each test. A duplicate of the 100th card was presented as a 101st card as a check on whether the subject had retained the idea and purpose of the test. Subjects who said 'no' to this card were excluded. Occasionally someone would erroneously say 'yes' to the 100th card (and then again to the 101st), in which case this check is invalid. In such cases the subject was questions about the nature of the tests to clarify whether he had retained the desired idea. Almost no subjects had in fact to be excluded. After a suitable rest period, the Designs test was given in the same way after an explanation of the different type of test material.

Results and discussion.

Mean values of d' and B for each group are shown in Table 2. A one-way ANOVA was carried out on the d' and B scores for each type of material, broken down by diagnostic group: the F ratios and

levels of significance appear at the foot of the table.

Table 2 Mean values of d' and B for the three groups in each test (standard deviations in parentheses).

	Photographs		Designs	
	d'	B	d'	B
DAT (n=21)	1.61 (1.03)	4.37 (4.78)	0.77 (0.62)	1.56 (0.96)
MID (n=19)	2.44 (1.40)	4.66 (3.99)	1.30 (0.79)	1.99 (2.08)
Depr (n=19)	3.42 (0.75)	5.63 (4.24)	1.57 (0.59)	2.02 (2.38)
F	13.78***	0.45	7.48**	0.37
sig.	.0001	.64	.0014	.70

It can clearly be seen that in both tests the d' measure shows significant differences between the groups while the B measure does not. As it happens the MID group are the better of the two groups of dementing subjects in terms of sensitivity (i.e. d'), but both are poorer than the depressed group. As with the study on encoding preferences, the traditional idea that memory performance is quantitatively worse in dementia than in depression is supported. Possible alternative explanations are as discussed in that study. There is a slight but interesting trend for the decision criterion to be higher in the MID group than in the DAT group, and highest of all in the Depressed group (as might have been predicted); but the spread of scores is so great that this trend is far from significant.

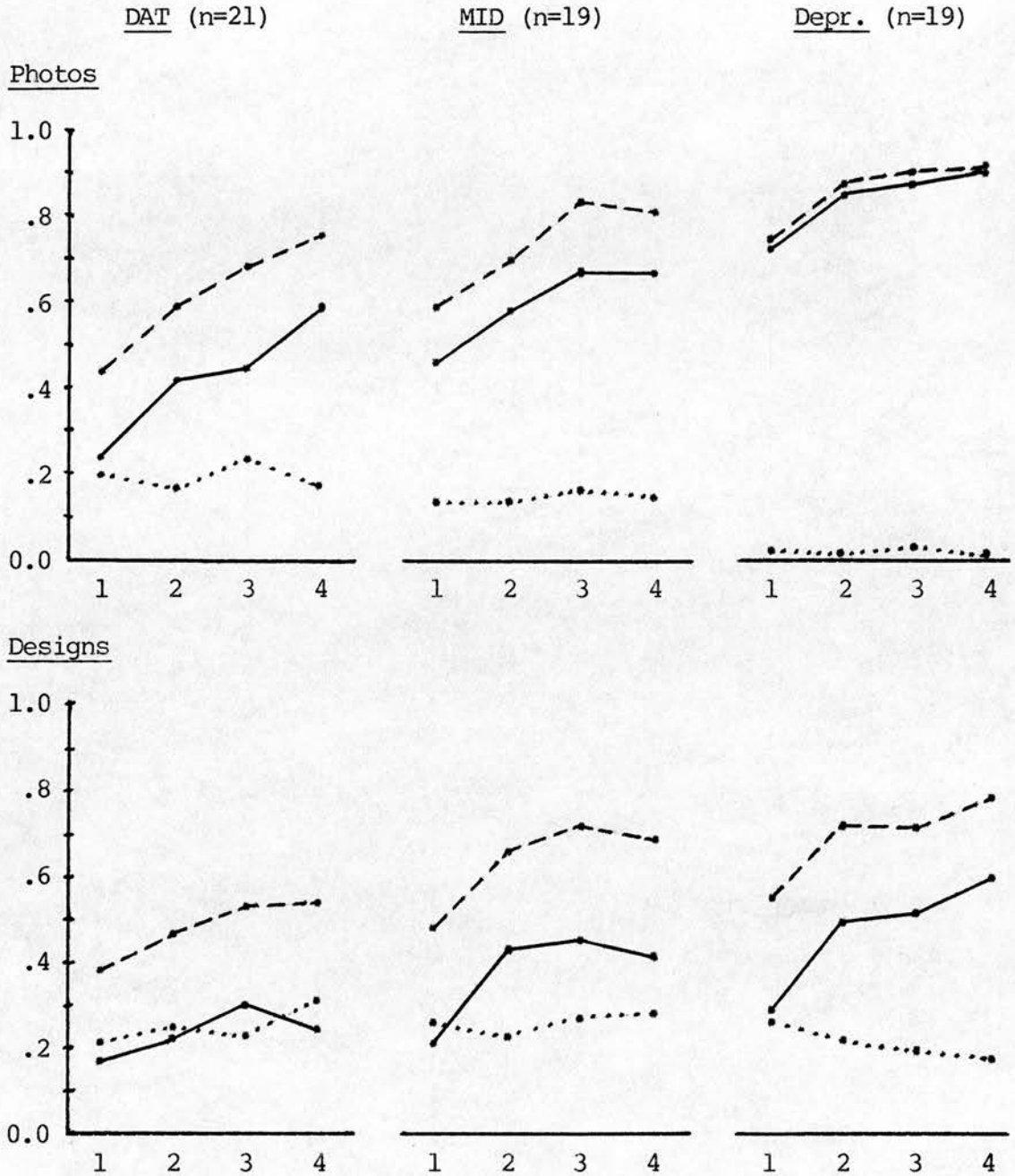
Pearson product-moment correlations between the various test measures and age were calculated for each group and for all three combined. The intercorrelations were largely as one might expect: d' scores did not correlate with B scores. d' scores correlated weakly with age in a negative direction (i.e. older subjects were less sensitive). B scores showed no relationship to age except,

interestingly, within the depressed group on the Photographs test: here older subjects tend to have higher response criteria than younger ones. Correlations between Photographs B and Designs B were non-significant in the DAT and Depressed groups, suggesting that the response criterion or the factors influencing it may vary according to the nature of the material involved. Photographs d' and Designs d', however, always correlated significantly together. Sensitivity or memory 'strength' may be a more stable parameter, less affected (in populations such as these) by variations in the type of test material. Designs is certainly a harder test than Photograp hs judging by mean d' scores. To a subject it also looks at face value a much harder test, and the designs seem more 'confusable' than the photographs: it would be difficult to say a priori whether these factors should raise or lower B.

It is possible to make use of the test data in a different way. Since the tests involve repeated presentation of the same material (i.e. the same patterns or designs), they may be considered tests of learning.

Performance in each 'block' of 20 cards (excluding the first 20 cards of each test, where no response was required) was therefore considered in terms of simple proportion of hits and false alarms. These performances are presented graphically in Figure 7. The measure of overall sensitivity (shown in solid line) is simply the proportion of hits minus the proportion of false alarms. There are no striking differences between the patterns of performance of the groups in the Photographs test. The Depressed group is clearly better than the other groups overall. To gauge learning over trials in some way, difference scores between the the first and last

Figure 7 Mean proportions of hits and false alarms (and the difference between the two) in the three subject groups.



Abscissa shows 'block' number; ordinate shows proportion.

blocks of 20 were calculated for the proportion of hits, of false alarms, and the sensitivity measure. One-way ANOVAs showed no significant differences between the groups on these difference scores. In fact the improvement from first block to last is largest in the DAT group and smallest in the Depressed group: this probably reflects a ceiling effect in at least some of the Depressed subjects.

The patterns of performance in the Designs test are rather more interesting. Here there are no suggestions of ceiling or floor effects. Over the 4 'blocks', the Depressed group shows a fairly steady rise in hit rate, a fall in false alarm rate, and hence a rise in overall sensitivity. The two dementing groups show some rise in hit rate (though this tends to have 'tailed off' by block 4) but they show either no reduction in false alarms or, in the DAT group, an actual rise in false alarms. Conequently the overall sensitivity measure in these groups peaks and then tails off, almost as if the dementing subjects eventually became overloaded by the ever-increasing number of new designs presented. Difference scores between blocks 1 and 4 were calculated as before: the ANOVAs which indicated significant differences (at the .05 level) were those on the sensitivity measure and the the false alarm rate. Change in hit rates did not differ significantly between the groups. It seems that these depressed subjects show a more consistent improvement in performance with repeated presentation of this type of material, particularly as regards reducing their false alarm rate.

Brief analyses were carried out concerning sex differences on the various test variables by means of t-tests (male versus

female) in each group separately and in all three combined: a tendency for males to perform significantly better than females on the d' measure was explicable wherever it occurred in terms of the males being significantly younger (in view of the correlations between d' and age). Similarly, test performances were analysed according to drug use (on psychoactive medication versus not): where any differences occurred they were in the direction of higher d' scores in subjects on drugs. This again is confounded by subjects on drugs being significantly younger than those on none. There is no good evidence that sex or drug use had a major independent influence on d' scores. There were no significant differences concerning B scores.

Overall, the experiment suggested that, in these particular samples of depressed and dementing subjects, depressed patients show (1) quantitatively better performance on testing and (2) more learning over time (at least with one type of material), with a steady decrease in false alarms despite the fact that the groups do not differ on overall levels of the response criterion B. Perhaps one could say that the depressed subjects were better able to inhibit competing responses or judge familiarity as the task progressed, but the result is explicable in other terms as well. The test design does not permit specific formulations of whether poor performances result from defects in registration, retention, retrieval, or some combination of these.

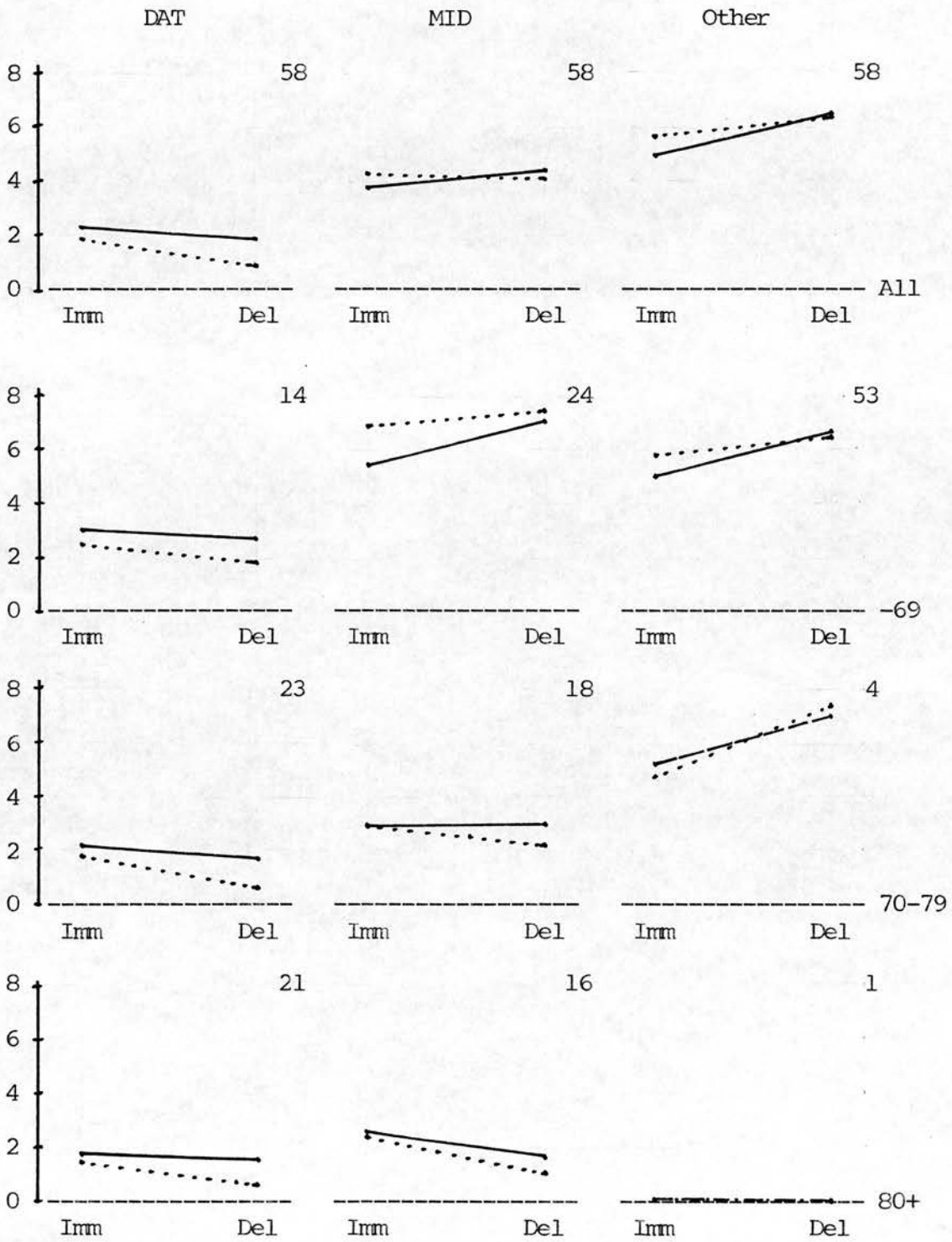
Memory performance in the main study: Paragraph Recall.

The importance of testing for delayed as well as immediate retention of new material in memory tests is stressed by a number of authorities (e.g. Russell, 1975), with the assumption either

that delayed testing measures something different from what is measured by immediate testing or that it measures the same thing but with greater sensitivity. It is clear from clinical experience that some amnesic subjects may perform reasonably well at immediate testing but then go on to score zero at delayed testing some time later. Delayed testing on Paragraph Recall (as well as on Face-Name Learning) was therefore incorporated into the present study. Since the test was part of Full testing, the subjects referred to in this section are exactly those described in previous chapters concerning subjects who completed Full testing. The administration and scoring of the Paragraph Recall test was summarised in Chapter 2, and appears in detail in Appendix 1. It may be remembered that for each subject, at immediate and at delayed testing, a score based on specific questioning was recorded as well as one consisting of the number of ideas spontaneously recalled. So far the analysis of Paragraph Recall has used only the composite score to questions of the immediate and delayed recall tests. Now immediate and delayed recall, and both the questioning and spontaneous recall measures of each, will be considered separately. Figure 8 shows these four scores in each diagnostic group and then in each group broken down by age as in previous analyses.

It can be seen that the recall measure based on questions gives very similar patterns to those based on spontaneous recall. In all the DAT groups, delayed recall performance is almost identical to, or marginally poorer than, immediate performance. This is also true of the two older MID groups, but in the youngest MID group and in the Other groups, delayed scores tend to be a little higher than immediate ones. It must be remembered that the questioning

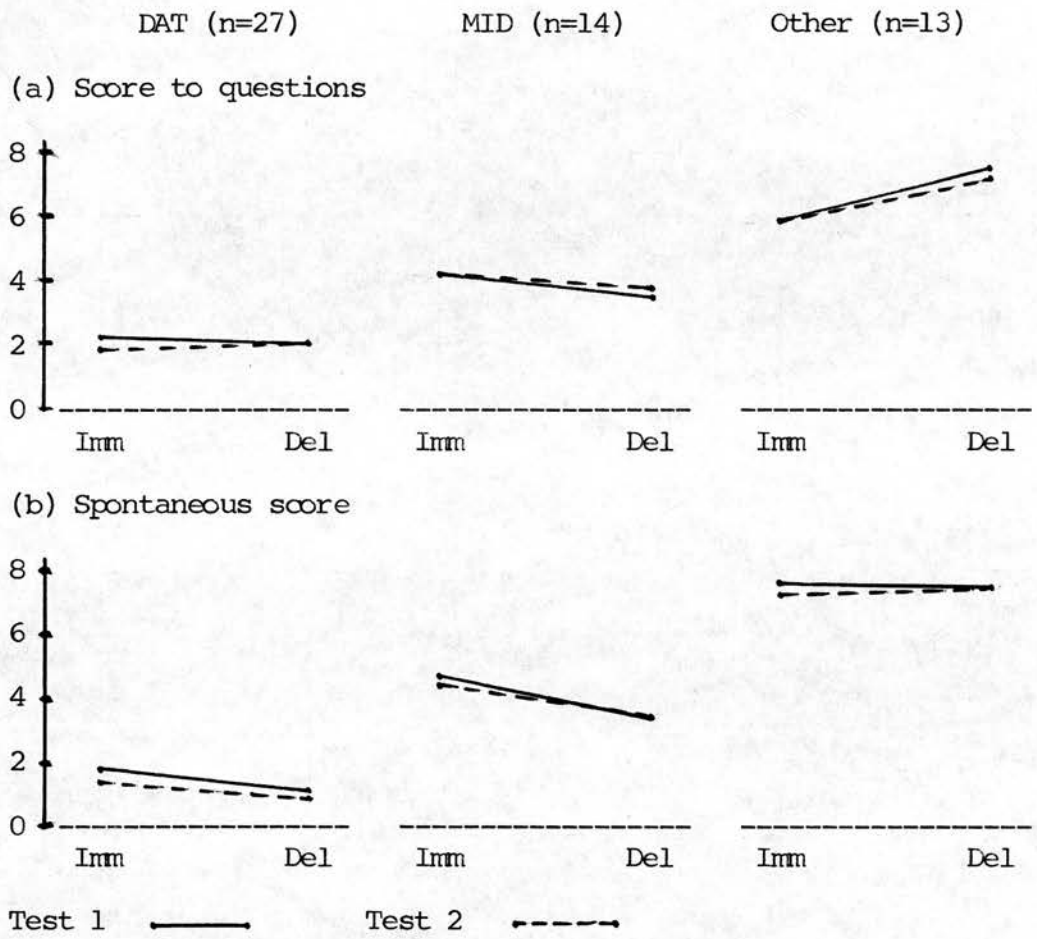
Figure 8 Mean scores on immediate and delayed Paragraph Recall (spontaneous recall scores and scores on questioning) in each diagnostic group and then in each group broken down by age. Number of cases shown at top right of each graph.



procedure used provided virtually a second presentation of the important points of the paragraph after the first presentation and test (by giving the correct answers to the six questions) as well as providing prompts or reminders at any given testing simply through asking the questions. Perhaps the slight differences between groups simply reflects the greater ability of some subjects to make use of this 'second presentation' and the prompts. The groups which show higher scores at delayed than at immediate testing are also the groups which show higher overall levels of performance on the test. The time period between immediate and delayed testing varied somewhat between subjects because it depended on how long subjects took to perform a certain number of intervening tests. The influence of using a test-determined rather than a time-determined interval cannot be assessed here. Intercorrelations (Pearson product-moment) between the four measures (immediate recall to questions and spontaneously and delayed recall to questions and spontaneously) in each diagnostic group and in the three combined were all highly significant and usually around .7 or .8 (though slightly lower in the DAT group). The method of scoring the answers to questions (with bonuses for 'spontaneous answering' of questions) contributes to the correlations between spontaneous recall and questioning measures.

Figure 9 shows Paragraph Recall scores (to questions and spontaneously) at Test 1 and at Test 2 in subjects tested twice, broken down by diagnostic group. The patterns at Test 1 are very similar to those seen in the previous figure (i.e. are representative), and there is extrememly little difference between the patterns at Test 2 compared with those at Test 1.

Figure 9 Paragraph recall scores (a) with questioning and (b) spontaneously, at test 1 and test 2 in subjects tested twice, broken down by diagnostic group.



Abscissa shows recall condition; ordinate shows test score.

These group analyses show few interesting differences between diagnostic categories concerning immediate versus delayed recall of paragraphs, nor between a measure of recall based on questioning compared to a spontaneous recall measure. This certainly does not imply that delayed testing or the use of questioning is pointless in all cases: the group means may hide considerable individual variation, and these procedures may yield helpful information in individual cases.

Face-Name Learning.

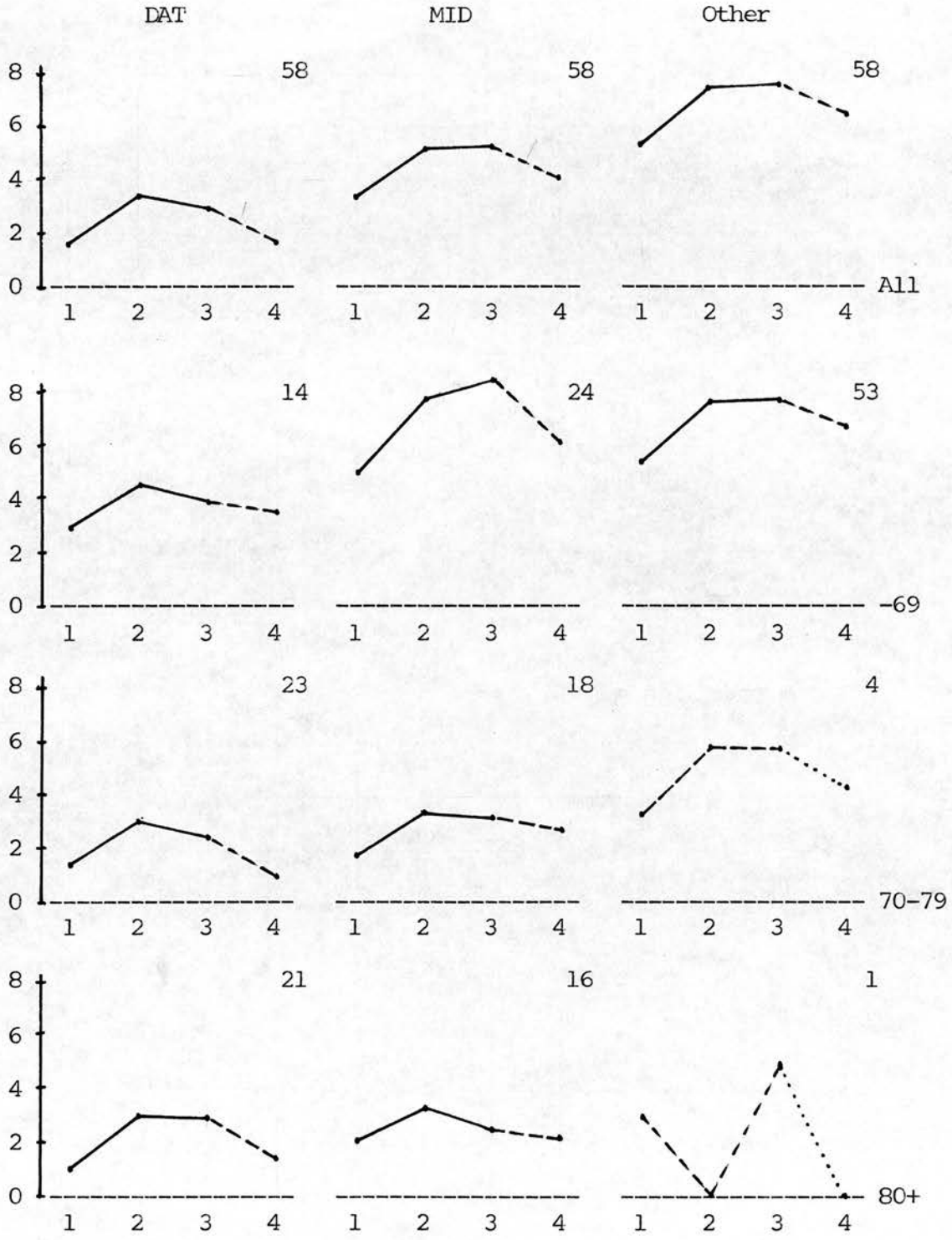
Paired-associate word learning tests are known to be sensitive and useful measures of impairment in organic brain syndromes (Erickson & Scott, 1977). The Face-Name Learning test used in this study is clearly a test of learning in that it involves repeated presentations of material and repeated tests of retention. It involves association learning to some degree, though this is perhaps more complex than is the case with word association learning. As previously mentioned, the test involves association between the first and second names of the pictured people as well as between the names and the faces. As with Paragraph Recall, the procedure includes a delayed retention test, and so also provides some measure of the loss of the acquired information over time.

The administration and scoring of the Face-Name Learning test was summarised in Chapter 2, and appears in detail in Appendix 1. Since the test was part of Full testing, the subjects referred to in this section are exactly those described in previous chapters concerning subjects who completed Full testing. So far only the composite score of the total number of points achieved over the four retention tests (three immediate and one delayed) has been

considered. Patterns of learning and retention can be examined by plotting separately scores at each retention test. Figure 10 shows mean score on each of the four trials of Face-Name Learning in each diagnostic group and then in each group broken down by age as in previous analyses. The curves are similar in all the DAT groups and in the two older MID groups, with no more than slight improvement over the first three trials (mostly accounted for by improvement between trials 1 and 2) followed by a slight drop again at delayed testing. The Other groups show a similar pattern, but at a higher overall level of performance. Only in the youngest MID group is there evidence of improvement between trials 2 and 3 as well as between trials 1 and 2: this is then followed by a considerable drop at delayed testing, perhaps because of the artefact that it is only possible for individual subjects to show a large drop between trials 3 and 4 if they have achieved a fairly high level on trial 3. As with Paragraph Recall, the time period between immediate and delayed testing varied between subjects: again no data are available concerning the importance of this. Scores on the four test trials intercorrelated highly in each of the three groups.

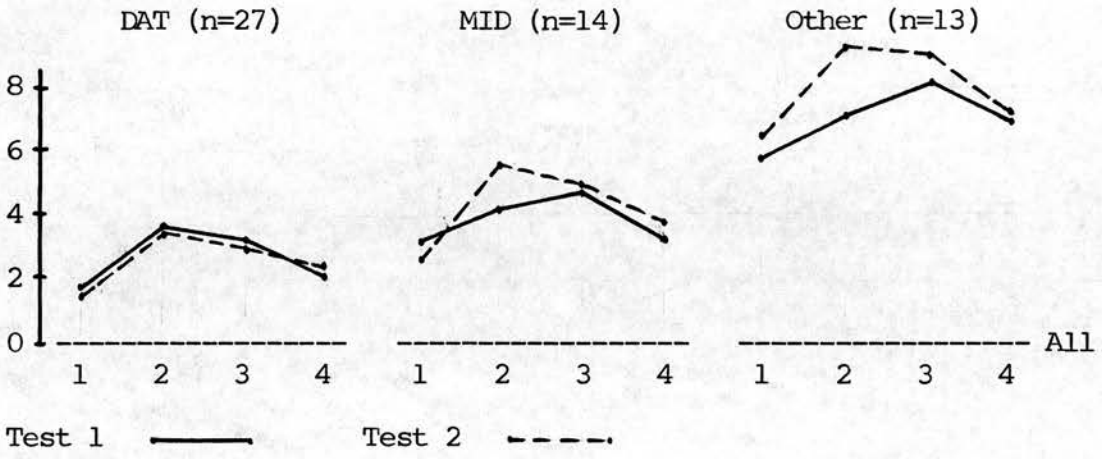
Figure 11 shows mean score on each of the four trials of Face-Name Learning at Test 1 and Test 2 in subjects tested twice, broken down by diagnostic group. The patterns at Test 1 are very similar to those seen in the previous figure (i.e. are representative). There are few notable differences between the patterns at Test 2 compared with those at Test 1 except for the fact that at Test 2 the youngest MID group have reverted to a pattern more typical of all the rest of the groups, i.e. with improvement tailing off between trials 2 and 3 and then only a moderate drop by trial 4.

Figure 10 Mean score on each of the four trials of Face-Name Learning in each diagnostic group and then in each group broken down by age. Numbers of cases shown at top right of each graph.



Abscissa shows recall test number; ordinate shows test score.

Figure 11 Mean score on each of the four trials of Face-Name Learning at test 1 and test 2 in subjects tested twice, broken down by diagnostic group.



Abscissa shows recall test number; ordinate shows test score.

These analyses show few striking differences between diagnostic groups concerning patterns of learning and subsequent retention on the Face-Name Learning test. As with the Paragraph Recall data, the group means may hide considerable individual variation and analysis of acquisition and retention curves may be of value in individual cases.

Yes-No Picture Memory.

Clearly, the brief Yes-No Picture Memory test included in the main battery of tests might be susceptible to a signal detection analysis. On such a brief test, however, it was fairly common for individual subjects to score no false alarms or 10 out of 10 hits. Calculation of d' and B is questionable under such circumstances, but an attempt was made by considering a proportion of zero (false alarms or hits) as equivalent to a very small proportion (i.e. not zero but much less than 0.1, which is the proportion representing one hit or one false alarm), and similarly considering a proportion of 1.0 as equivalent to a proportion of just less than 1. A table of values of d' and B (shown in Appendix 1) was drawn up on this basis, using the usual formulae provided by Hochhaus (1972). The administration and scoring of the test was summarised in Chapter 2, and appears in detail in Appendix 1. Since the test was part of Full testing, the subjects referred to in this section are exactly those described in previous chapters concerning subjects who completed Full testing. As described in Chapter 2, a 'raw' sensitivity score (proportion of hits minus proportion of false alarms) had already been calculated. Table 3 shows mean values of d' and B broken down by diagnosis.

Table 3 Mean values of d' and B broken down by diagnosis. Standard deviations in parentheses. Numbers of cases as before.

	DAT		MID		Other	
	d'	B	d'	B	d'	B
All	1.72 (1.10)	1.97 (1.37)	2.54 (0.98)	2.00 (1.30)	2.56 (0.98)	2.03 (1.31)

Mean d' scores are worse in the DAT group as a whole than in the MID and Other groups, which is of little interest, and mean B scores are almost identical in all three diagnostic groups. Pearson product-moment correlations between raw score, d', and B were calculated in each diagnostic group and in the three combined. The high correlations found between raw score and d' are not surprising, and the low correlations found between d' and B are desirable since the two are supposed to be theoretically independent. B did not correlate significantly with age or Full score in any group. Raw score and d' correlated significantly with Full in all groups and (negatively) with age in the MID, Other, and combined groups but not in the DAT group alone. In every group d' correlated significantly with estimated premorbid I.Q. (though reservations have been expressed about the validity of these estimates) but B did not.

As a check on the validity of the method of calculating d' and B for this brief test, scores derived from it were correlated with the corresponding scores on the longer 100-item picture-memory test previously described. Thirty-eight subjects had completed the 100-item test within a few weeks of completing the brief test as part of Full testing. The Pearson product-moment correlation coefficients were .72 for the Raw sensitivity measure, .70 for d', and .32 for B. The coefficients for raw sensitivity and d' are

acceptably high, but that for B is disappointing. It is possible that the B measure derived from the brief test is not invalid, but simply different from the B measure derived from the long test because of the differences in the tests or because of the time separating their administrations; we have already seen in a previous section that differences in test materials may affect the B measure, and it may simply be a less robust measure across situations. There is no way of testing this further here. In the 54 subjects tested twice on Full testing, test-retest correlations (Pearson product-moment) were .77 for the Raw measure, .67 for d', and only .22 for B. This makes the B measure highly suspect.

Clearly any given moderate Raw score can be achieved either with a relatively large number of hits and of false positives or with a smaller number of both. This of course relates theoretically to the B measure itself but, in view of the dubious validity of the B measure derived from this test, mean numbers of hits and false positives in different groups will be considered directly. Table 4 shows mean numbers of hits and false positives, and mean raw scores (consisting of hits minus false positives), in the three main diagnostic groups.

Table 4 Mean numbers of hits and false positives, and mean raw scores (consisting of hits minus false positives), in the three main diagnostic groups. Standard deviations in parentheses.

	Hits	False Pos	Raw score
DAT (n=58)	6.9 (2.9)	1.8 (2.9)	5.1 (3.3)
MID (n=58)	8.2 (2.2)	0.6 (1.7)	7.5 (2.8)
Other (n=58)	8.1 (2.3)	0.6 (1.6)	7.5 (2.9)

All the table seems to show is that DAT subjects have on average a lower overall raw score than MID or Other subjects and that their

poorer scores result from both a lower hit rate and a higher false positive rate. This trend is fairly consistent across age groups within the main diagnostic groups.

Pearson product-moment correlations between hits, false alarms, age, overall Full score, and estimated premorbid IQ produced no surprises. A primacy-recency analysis concerning the likelihood of hits and of false positives at different serial positions (from 1 to 10 in each case) revealed no interesting differences between the three main diagnostic groups. All three groups showed a drop in hits at the final end of the serial position curve on hits, but otherwise the curves concerning both hits and false positives were more or less flat in all groups.

Analyses of changes in performance in subjects retested 10 months later indicated that DAT subjects became less conservative as well as less accurate with time. In MID subjects there was no change in either parameter. In the Other group there was a slight and uninteresting trend for subjects to become more accurate but less conservative.

Allowing for differences in overall levels of test performance, the analyses show few differences between diagnostic groups and subgroups concerning patterns of performance on this test. Again, the group data may hide considerable individual variation.

Orientation.

Accounts, sometimes unsubstantiated and contradictory, appear in the literature to the effect that some aspects of orientation are more affected than others in different conditions. Individuals in this study did seem to differ in this respect. Benton (1968) associates poor orientation in time with bilateral frontal lobe

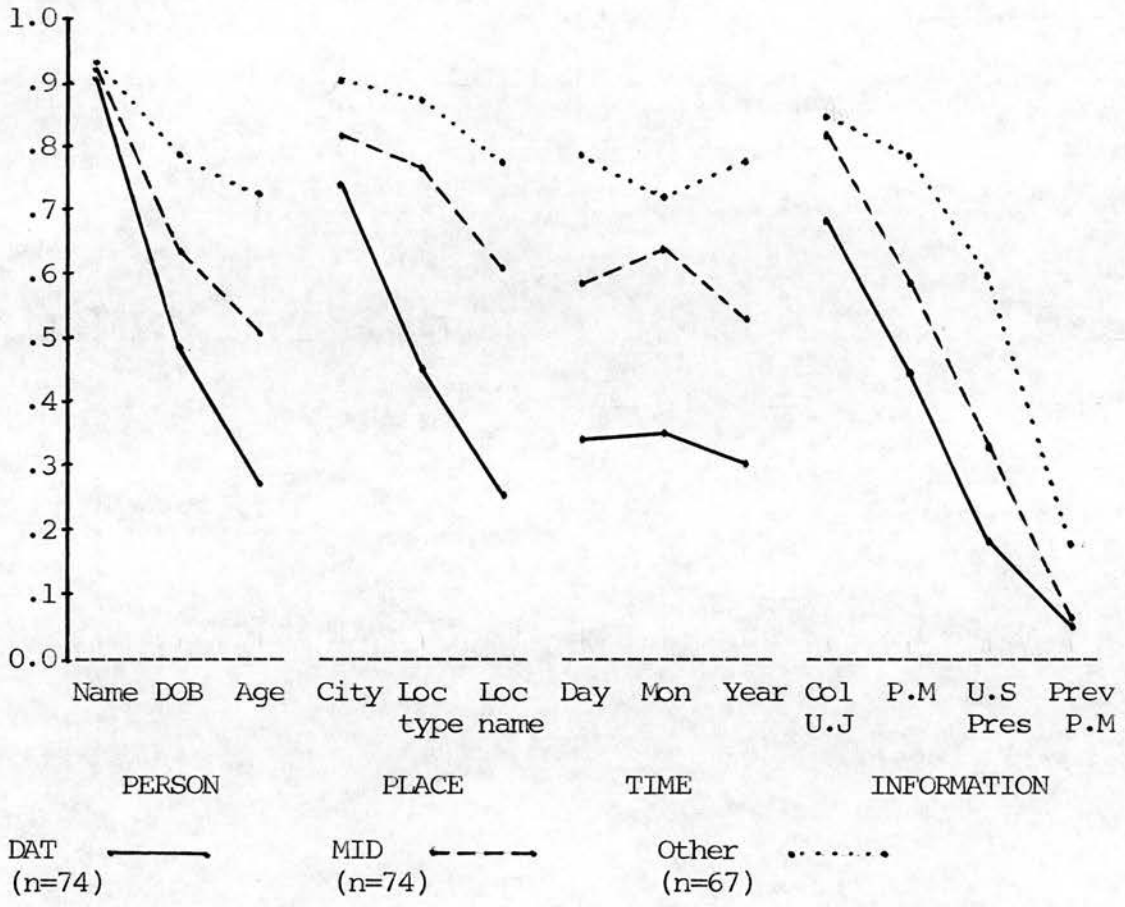
dysfunction. Different aspects of Orientation performance in this study were therefore examined. Since the Orientation test was part of Short testing, subject details relevant to this analysis are identical to those given previously for subjects completing Short testing. Details of the test appear in Chapter 2 and Appendix 1.

Clearly, some Orientation questions are harder than others. Figure 12 shows the mean proportion of subjects answering each Orientation question correctly, broken down by diagnosis. There appear to be no striking pattern differences between the groups, allowing for the fact that the Other group generally perform best (mean overall score 9.7, SD 3.5), followed by the MID group (mean 7.9, SD 3.8), and finally by the DAT group (mean 5.5, SD 3.6). Four scores were computed for each subject concerning different aspects of Orientation: Total number correct on

(1) Personal information (name, age, and date of birth)
(2) Place information (nature of location, its name/address, and the name of the town or city)
(3) Time information (weekday, month, and year)
(4) Other information (name of the Prime Minister, name of the U.S. President, colours of the Union Jack, and name of the previous Prime Minister).

Mean patterns of performance in different subject groups and subgroups (as in previous analyses) were then examined. Data will not be presented: few interesting differences emerged apart from a slight trend for DAT subjects to be better oriented in person than in place but for MID and Other subjects to be better oriented in place than in person. (As stated above, MID subjects were better oriented overall than DAT ones, and Other subjects best of all.)

Figure 12 Mean proportion correct on each Orientation question, broken down by diagnosis.



Abscissa shows Orientation question; ordinate shows proportion correct.

This trend was consistent across age groups except that the youngest DAT group showed a pattern more like the MID and Other groups. Orientation in time was at the same level as score on other information (using proportion correct out of 3 and 4 respectively) in all DAT groups, while in the MID and Other groups orientation in time was better than score on other information. In all three diagnostic groups, overall performance levels were worse in older age groups. Breakdown according to overall Orientation score was unremarkable. Correlational analyses involving the computed variables and age and Short score produced no surprises in any group.

However, cross-tabulation of each of the four computed Orientation variables with all of the others (1 with 2, 1 with 3, 1 with 4, 2 with 3, 2 with 4, and 3 with 4) showed the extent of individual differences: in each diagnostic group, with every combination of variables, there were almost always some subjects who scored at least two points more on the first variable than on the second and others who scored at least two points more on the second than on the first. The different parts of the Orientation questionnaire may therefore be assessing different aspects of cognitive functioning; but any differences would seem to be largely unrelated to diagnosis, age, or level of impairment. Perhaps differences have more to do with the locations in which subjects were tested or some other aspect of their daily lives and activities. The cross-tabulation data also indicated that the Orientation scale would not have the characteristics of a Guttman scale and that any further abbreviation of (already short) orientation questionnaires would not be justified.

The only memory test used in the main testing which has not been considered in detail is Memory for Designs. Data on this are limited to simply a single score for each subject, and so there is no scope for detailed analysis of component parts of performance on the test. The data were simply analysed in terms of mean scores broken down by diagnostic group and by age groups within diagnostic groups, followed by appropriate correlational analyses. No remarkable findings or trends emerged which would not have been predictable from more general aspects of test performance in the various groups (such as patterns of factor scores or of overall Full scores), and so details of these analyses will not be presented. Investigation of relationships between performance on this test and on Copying Designs similarly produced no findings worth reporting.

CHAPTER 7

Further Analysis of Language Impairment.

Some disturbance of language function is known to be common in dementia but, particularly in DAT, the disturbance is not generally considered to mimic any of the classic or conventional dysphasias seen with focal lesions such as tumour, single stroke, or localised head injury. Bayles (1982) assessed patients with DAT using a large battery of language tests, and found that linguistic abnormalities, where present, were mainly semantic in nature. The patterns of difficulties did not conform closely to traditional subcategories of aphasia. Appell et al (1982) noted differences between DAT subjects and normal and stroke subjects respectively. In DAT they found that disturbances akin to Wernicke's dysphasia and transcortical sensory dysphasia were common, while disturbances akin to other categories were rare. Albert (1980) suggests that the overall picture of language impairment in dementia may often appear similar to that of Wernicke's, but that there are underlying differences: in particular that the 'general intention' of utterances is retained in Wernicke's but lost in dementia. A philosophical discussion of the relationship between thought and language is beyond the scope of this chapter and author: Vygotsky (1978) discusses this in depth.

The impression gained from assessing the DAT and MID subjects in this study was that dysphasic problems, if present, tended to be of a basically receptive or receptive/conductive type, i.e. with poor comprehension and essentially fluent speech sprinkled with literal paraphasias (as opposed to the semantic paraphasias considered typical of 'classic' receptive dysphasia). Severe impairment of speech fluency as seen in Broca's aphasia was very rare (non-existent in the DAT subjects) in accord with the notion that a

typical Broca's aphasia tends to require a relatively sudden large lesion rather than a gradually progressive one or an accumulation of small serial ones. Certainly the quantity and content of some subjects' speech was very limited and impoverished, but this usually seemed to reflect at least in part a general lack of spontaneity or volition rather than a true dysphasia: the speech output, though limited, was often gramatically correct with near normal prosody, unless dementia was fairly severe. Reduction in fluency as measured by tests requiring the production of as many words as possible beginning with a given letter or belonging to a particular category (within time limits) and in spontaneous output is known to be associated with frontal lobe damage outside specific language areas. Hence such reduction might be expected in dementia, and has often been found (e.g. Rosen, 1980; Weingartner et al, 1981). General verbal intellectual level and speed (because of the use of time limits) may be important contributors to this impairment in dementia (Miller, 1984c; Miller & Hague, 1975).

A formal analysis of the categorisation of language disturbance in dementia in terms of traditional systems of subdivision cannot be attempted on the basis of the data available here. However, some aspects of performance on some of the language tests used can usefully be considered in detail. These aspects are: (1) Nominal Ability (in conjunction with Object-Recognition ability); (2) expressive speech as measured by the Sentence Production test; and (3) the relationship between comprehension deficit (as measured by the Token Test) and other test performances.

Object Recognition and Nominal Ability.

Kirshner et al (1984) remind us that there must be at least three steps or processes involved in the naming of an actual or pictured object: the visual recognition of the object, the finding of its name, and the transformation of the name into an act of speech. Each step may be divisible into subprocesses, but this is not crucial here. The third step can be assumed to be largely intact in dementia if the ability to either converse, read aloud, or repeat verbal material is adequate. The impairment in confrontation naming in dementia has often been assumed to lie in the second, word-finding, stage, but this traditional view has been challenged by a number of authors. Rochford & Williams (1964) suggested that in dementia there is evidence of misperception of objects rather than difficulty in naming correctly perceived objects, that an object can usually be named if it can be easily recognised, and that naming failures are not related to the word frequency of the name being sought. (Failures are related to word frequency in dysphasia, and theoretically this is thought to implicate the second, semantic search, stage.) Rochford (1971) concluded that the naming problem in dementia (as contrasted with dysphasia) lay in the first, perceptual, stage on the basis of a qualitative analysis of the types of errors made in naming pictures of objects plus the fact that dementing subjects did very well when asked to name body parts (which are unlikely to be misrecognised). This improvement with body parts was not considered attributable to an effect of word frequency since some body part names had a low frequency in the Thorndike-Lorge (T-L) count; it was noted that dysphasics did not show a comparable improvement over object naming

in their naming of body parts. Barker & Lawson (1968) and Lawson & Barker (1968) did find that failures and long response times in object naming in dementia were associated with low word frequency (as judged by the T-L magazine count), suggesting difficulties in the second stage. However, they also found that demonstration of the use of objects improved subjects' performance, probably indicating impairment of the first stage. It is of particular note that Barker & Lawson excluded patients thought to have visual agnosia. Kirshner et al (1984) investigated naming performance in DAT using materials of varying perceptual difficulty: masked drawings, plain drawings, and photographs of objects, and actual objects, with the names of the objects used being classifiable as high or low frequency. They concluded that both perceptual difficulty and word frequency were important determinants of naming performance in dementia. They found that naming problems could occur even in mild dementia and in dementing subjects whose overall language function remained 'normal' on an aphasia battery; the problem worsened in proportion to the degree of overall language impairment on the aphasia battery and to the degree of overall cognitive impairment on other neuropsychological tests. Miller (1981a) notes that if objects whose names are infrequent in language are also more difficult to recognise then the finding of a word frequency effect in naming objects but not in conversational speech (Miller & Hague, 1975) is explained: recognition is not required in conversation.

Clearly there has been disagreement over the principal processes impaired in the naming disorder in dementia. Some of the differences of opinion may arise from differences in study

methodology and subject selection: but in a condition such as DAT, which by definition involves deterioration in a wide range of cognitive functions, it seems highly likely that (as Kirshner et al conclude) both the first and second stages may be involved (though perhaps in different proportions in different subjects). Or it may be that the two stages are simply not as separable or sequential as is commonly assumed: controversy continues over the status of the concept of visual object agnosia, but the phenomenon is clearly no simple perceptual deficit. Finally it is perhaps worth noting that in fairly severe dementia the third, articulatory encoding, process may also be disturbed, but this will not be considered here.

The Nominal Ability/Object Recognition test incorporated into Full testing in this study had been designed with the earlier of the above findings in mind. The 23 objects were chosen so that their names covered a range of word frequency according to the T-L overall count. The drawings of the objects were arranged in order of decreasing word frequency, i.e. most common first. Details of the administration and scoring appear in Chapter 2 and Appendix 1. Appendix 2 includes the actual pictures. To recap briefly, each subject was asked to name each object shown in turn and, if he was unable to name it in 15 seconds, to explain or demonstrate how the object was used or what it was used for, or otherwise show that he had correctly recognised it. Naming of seven body parts (indicated by the tester on his own body) was also requested, again with a maximum allowance of 15 seconds. The number of seconds taken to name each object was recorded for every subject. If the subject could not name the object in 15 seconds but showed correct recognition of it (or named it after 15 seconds during the attempt

to describe its use) a response latency of 16 seconds was recorded. If he failed to recognise the object, a category of recognition failure was recorded. Latencies were recorded in the same way with body parts, but here no category of recognition failure was included: as previously stated, subjects never seemed unable to indicate on their own bodies which part had been indicated even if they could not name it, which was itself rare. (No distinction was being made between left and right body parts: if the tester indicated his elbow, the subject simply had to touch or indicate either of his own elbows for recognition to be considered correct.)

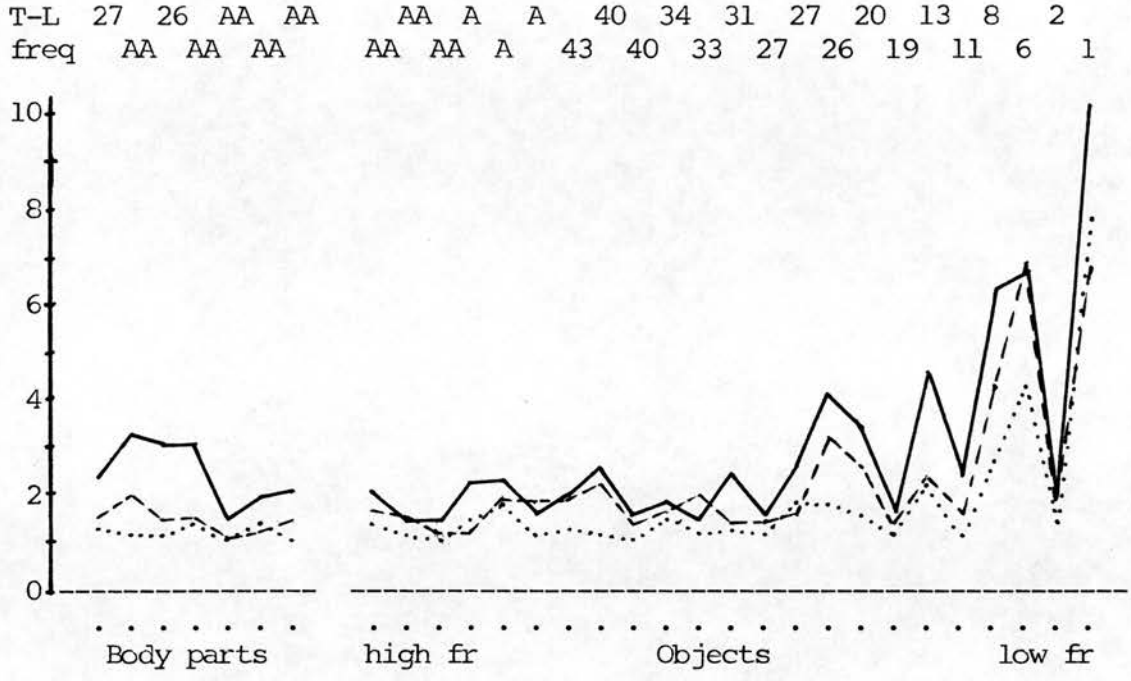
As this test was part of Full testing, subject details are exactly as described for Full testing in Chapter 2.

Figure 1 shows mean latencies in naming each body part, mean latencies in naming each object in those who could recognise the object, and numbers of subjects failing to recognise each object, broken down by diagnostic group. The latency graphs are very similar in the three diagnostic groups, and there is clearly some association between word frequency and naming latency. There are clear 'dips' in latency with some of the objects with lower frequency names (comb, bicycle, and teapot): this will be discussed below. Body parts are generally named about as quickly as objects with high frequency names: unfortunately none of the body part names required had very low word frequencies.

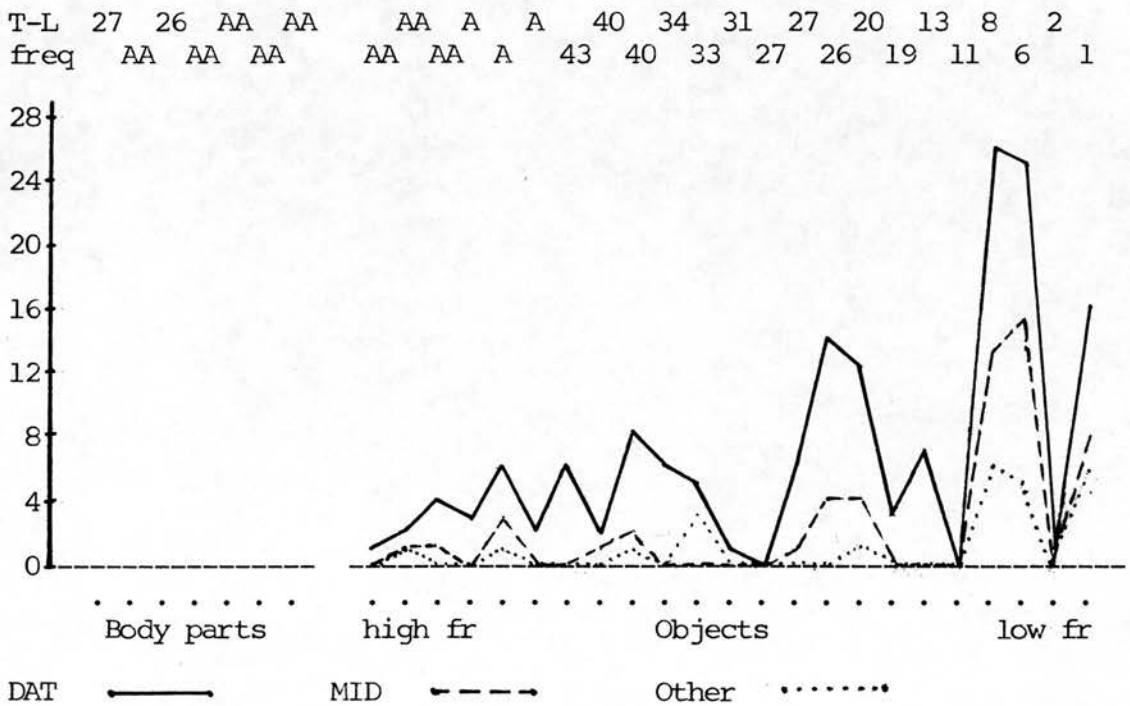
It is difficult to tease out the contributions of recognition and word-finding problems in failure to name objects, but with fairly large numbers of subjects (as in this study) a measure of recognition difficulty with any given object is the proportion of subjects failing to recognise it at all. The second part of Figure

Figure 1 (a) Mean latencies (in seconds) in naming each object and body part in those who could recognise any given object and (b) numbers of subjects failing to recognise each object, broken down by diagnostic group. Thorndike-Lorge word frequency shown above.

(a) Naming latency



(b) Recognition failures



Abscissa shows stimuli; ordinate shows (a) naming latency (in seconds) (b) number of subjects failing to name stimuli.

1 suggests some relationship between recognition difficulty and word frequency. Spearman rank-order correlations between word frequency, mean naming latency, and numbers of subjects showing recognition failure were calculated for each diagnostic group (excluding data on naming of body parts, where no recognition failure category existed): the correlations appear in Table 1.

Table 1 Correlations between word frequency, mean naming latency, and numbers of subjects showing recognition failure in each diagnostic group. (Spearman rank-order correlations incorporating correction for ties; each correlation based on 23 pairs of scores; ** significant at .01 level, * .05 level, 2-tailed test.)

	Frequency: Latency	Frequency: n fail rec.	Latency: n fail rec.
DAT	.63 **	.33	.54 **
MID	.57 **	.25	.52 *
Other	.52 *	.26	.34

Many tied scores occurred in the Other group because of the relatively low rate of recognition failure. The correlations between frequency and latency are as one would expect from previous work. The correlations between frequency and recognition failure are not significant, but they are positive and of a similar size in each group. These are difficult to explain away in terms of some artefact of testing or scoring procedure. Subjects who received this test were generally not severely demented, and it did not seem that subjects were simply unable to show in any way that they had recognised an object if they could not name it. Had this been the case, there should have been no relationship between recognition failure and word frequency i.e. recognition failures should have been randomly distributed across the various objects, unless some objects are inherently more difficult to show recognition of (by

demonstrating or describing their use) than are others: no data are available on this point, but it would be surprisingly coincidental if such difficulty was related to the word frequency of the objects' names. It is conceivable that, if wrong names are produced as a result of disconnection of visual and verbal areas, when asked to demonstrate the use of an object the subject might respond on the basis of the erroneous name he has generated. This possibility cannot be judged in the present study. It should be pointed out that the author was certainly not expecting any relationship between recognition failure and word frequency or naming latency, so the findings cannot be accounted for in terms of tester bias.

There is no objective or independent way of judging whether the drawings used in this study vary in terms of 'inherent' perceptual difficulty. All were simple line drawings immediately recognisable to normal people. The degree of perceptual difficulty of particular drawings might conceivably depend on the subject or on semantic associations of the pictured object rather than on physical characteristics of the picture itself. Hence perceptual difficulty would have to considered as a dependent rather than an independent variable.

Perhaps, then, there is a real relationship between access to the name of an object and knowledge of its nature. There are perhaps four aspects of an object's frequency which are of relevance to a discussion of object recognition and naming: the frequency of its name in published literature (as in the various T-L counts); the frequency of its name in oral use (i.e. in everyday conversation, either spoken or heard by an individual); the frequency of its name in internal thought (which is

unmeasurable but potentially important); and finally the frequency with which the actual object (not its name) is seen, used, or encountered in any way in everyday life. This concept of 'encounter frequency' becomes important if the very seeing, using, or in any way sensing of an object causes activation of some sort of semantic network, one constituent of which will be the object's name even though the name itself need not be spoken, heard, or consciously thought of: this might in some way make the subsequent finding of the object's name more easy. Unfortunately this seems virtually impossible to test without a confounding of encounter frequency and other variables such as word frequency. The influence of the four aspects of an object's frequency described above cannot be judged here, but it seems possible that the low naming latencies and rates of recognition failure for objects such as the comb, bicycle, and teapot may have something to do with the fact that the objects would have high encounter frequencies and their names high conversational frequencies despite their low T-L frequencies. Objects such as the screw, rake, dice, whisk, and corkscrew are plausibly less often encountered and talked about. Similarly the names of some of the body parts used by Rochford (1971) may be low in T-L frequency but they may have high conversational or internal thought frequencies, and the parts themselves clearly would have about the highest encounter frequency possible.

Assuming that the method of differentiating nominal from recognition failure is valid, the relationship between ease of recognition of an object and the T-L frequency of its name is interesting. This may relate to recent single case findings of an association between recognition and semantic category, where a

subject may show agnosia for objects in one category but not for those in another (Warrington, 1986).

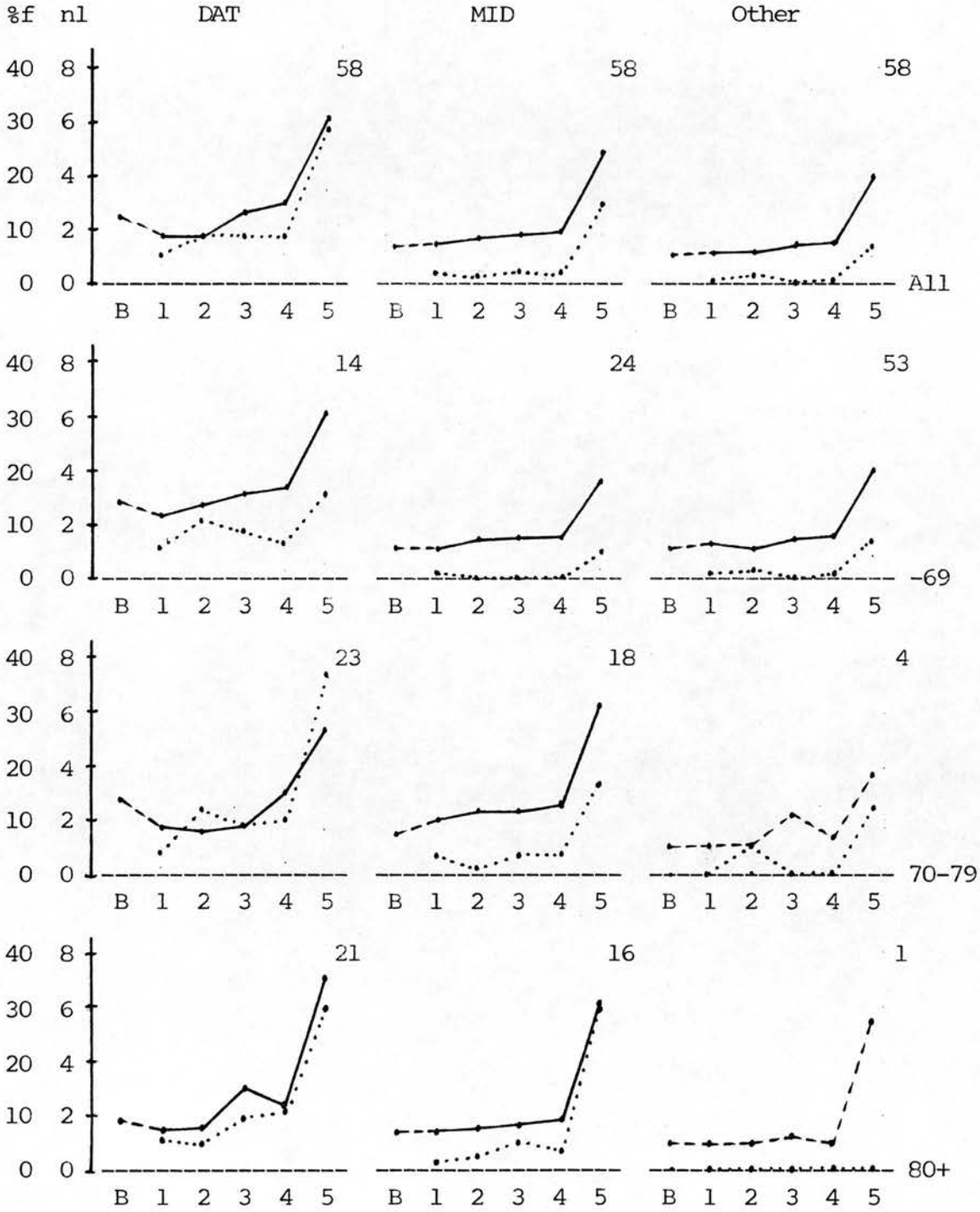
To facilitate interpretation of patterns in different subgroups, the objects and body parts were grouped together according to the frequency of their names as follows:

1. 6 items with T-L frequencies of at least 50 per million: Chair (AA), Table (AA), Book (AA), Key (A), Pen (A), Clock (A).
2. 5 items with frequencies between 33 and 43 per million: Candle (43), Pencil (40), Drum (40), Hammer (34), Spoon (33).
3. 4 items with frequencies between 26 and 31 per million: Fork (31), Kettle (27), Leaf (27), Anchor (26).
4. 4 items with frequencies between 11 and 20 per million: Screw (20), Comb (19), Rake (13), Bicycle (11).
5. 4 items with frequencies between 1 and 8 per million: Dice (8), Whisk (6), Teapot (2), Corkscrew (1).

'B'. 7 body parts with frequencies of at least 26 per million: Thumb (27), Finger (AA), Elbow (26), Knee (AA), Nose (AA), Eye (AA), Ear (AA).

Figure 2 shows mean naming latencies and mean percentages of subjects failing to recognise objects, broken down by diagnostic group and then by age as well, with body parts grouped together and objects of similar T-L word frequency grouped together. The Figure clearly shows that a dramatic increase in naming latency and in proportions of subjects failing to recognise objects only occurs with objects with T-L frequencies less than 10 (i.e. grouping 5). This is true for all three diagnostic groups and for all age-defined subgroups, and is consistent with previous findings. There are no striking pattern differences between groups.

Figure 2 Mean naming latencies (nl) in seconds and mean percentages of subjects failing to recognise objects (%f), broken down by diagnostic group and then by age as well, with body parts grouped together and objects of similar T-L word frequency grouped together. Number of cases shown at top right of each graph.



Naming latency ---——— % Recognition failures .....
Abscissa shows stimulus group: B: body parts, 1: 50 per million or over, 2: 33-43 pm, 3: 26-31 pm, 4: 11-20 pm, 5: 1-8 pm. Ordinate shows mean % of subjects failing to name stimuli and mean naming latency.

Figure 3 shows mean naming latencies and mean percentages of subjects failing to recognise objects at Test 1 and at Test 2 in subjects tested twice, broken down by diagnostic group, with body parts grouped together and objects of similar T-L word frequency grouped together. The patterns at Test 1 are very similar to those seen in the previous figure (i.e. are representative), and there is very little difference between the patterns at Test 2 compared with those at Test 1. The patterns for both latency and recognition would seem to be reliable.

In view of known sex differences in normal people concerning some aspects of language ability, a breakdown by sex of the naming latency and recognition failure data was carried out. The proportions of the sexes differed in different diagnostic groups and so any sex differences might have influenced patterns. Analysis was by inspection of graphs only. In the DAT group, females were generally rather slower to name all classes of object. In the MID and Other groups, females were slower only with the lowest-frequency object names. In the DAT and MID groups females were more likely to fail to recognise objects, though no such trend appeared in the Other group (where recognition failure was generally rare). Overall the magnitude of the sex difference was small. Slight differences on individual items (such as whisk or anchor) may reflect life experience in some way.

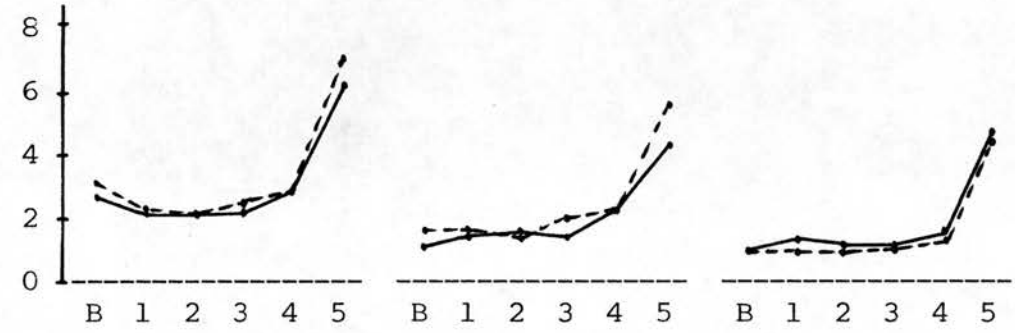
One other linguistic characteristic of words studied by Kirshner et al (1984) was word length: in view of their negative findings regarding this, the present results have not been analysed with respect to it.

A final general comment about the use of naming tests in

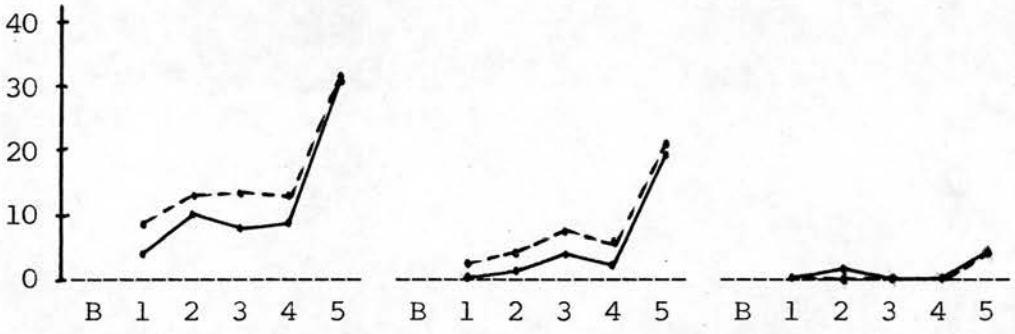
Figure 3 (a) Mean naming latencies and (b) mean percentages of subjects failing to recognise objects, at test 1 and test 2 in subjects tested twice, broken down by diagnostic group, with body parts grouped together and objects of similar T-L word frequency grouped together.

DAT (n=27) MID (n=14) Other (n=13)

(a) Naming latency



(b) Recognition failures



Test 1 ——— Test 2 - - - -

Abscissa shows stimulus group: B: body parts 1: 50 per million or over 2: 33-43 p.m. 3: 26-31 p.m. 4: 11-20 p.m. 5: 1-8 p.m.
Ordinate shows (a) mean naming latency (in seconds) and (b) mean % of subjects failing to name stimuli.

dementia may be in order. Skelton-Robinson and Jones (1984) suggest using nominal ability as an index of the severity of dementia. The use of any single test for such a purpose can be criticised in terms of the possibility of producing seriously inaccurate impressions in atypical cases (and, as we have seen, it seems that there may be no such thing as a typical case), but in any case this particular type of test does not seem to be the best choice. In the present study, nominal ability correlated only about .5 with the overall Full score; most other tests correlated more highly with Full score, including an orientation questionnaire (correlation about .8) of the simple and brief type already widely used. Nominal aphasia tests may be sensitive to the presence of early dementia but still be relatively poor indices of severity in established cases. Orientation tests are at least as good, and better still when combined with other simple spatial tasks such as those in Miscellaneous Visuo Tasks here (which itself correlates with Full very highly, perhaps because it contains a number of rather different tasks and so may sample a range of abilities or impairments).

Sentence Production.

Miller & Hague (1975) found no evidence of restriction of the range of words used nor of abnormally infrequent use of rare words in samples of conversational speech of patients in the early stages of PDAT when compared with controls. They also found that abnormally infrequent production of rare words on Thurstone's word fluency test in such patients probably resulted from an artefact caused by the time-limited nature of the test. Hutchison & Jensen (1980) did find abnormalities in the conversational speech of five

elderly dementing patients when compared with controls: the dementing patients tended to elaborate less on each conversational topic, to introduce irrelevant or tangential topics or material, and to break conventional conversational 'rules' concerning the continuation of topics. Possible explanations for such general abnormalities are wide-ranging and include general intellectual impairment, impaired memory, and various aspects of personality change including that of increased 'egocentrism' (Woods & Britton, 1985). Similarly Albert (1980) stresses the importance of general intellectual and thinking difficulties in the production of apparently disordered language in dementia.

The test chosen here to provide some objective measure of some aspects of language production was the Sentence Production test from Schuell's (1965) Minnesota Test for the Differential Diagnosis of Aphasia, as described in Chapter 2 and Appendix 1: after introductory examples, the subject is given six words in turn and each time asked to make up a sentence containing the given word. There are no time limits.

This type of method has a considerable history in studies of aphasic and schizophrenic speech. It has a number of advantages and disadvantages compared with studies of spontaneous or conversational speech. The Sentence Production test is structured and the requirement standard for every subject, which cannot be the case with samples of spontaneous speech, but it is clearly less naturalistic and 'ecologically valid'. Some subjects with dementia will have extremely limited spontaneous verbal output (which may result from a variety of intellectual or personality changes rather than any specific or fundamental changes in language capacity). The

Sentence Production structure may partially overcome such problems. The choice of the test was also influenced, as explained in Chapter 2, by the finding of Kaszniak et al (1978) that the test had predictive value concerning longevity in DAT.

As this test was part of Full testing, subject details are exactly as described for Full testing in Chapter 2.

Five measures were derived from the test for each subject:

(1) Number of correct sentences (out of 6) produced as defined in the Schuell manual.

(2) Mean length (in letters) of correct sentences.

(3) Mean length (in words) of correct sentences.

(4) Mean length (in letters) of words in correct sentences.

(5) The Type-Token ratio.

The first is the basic measure used by Schuell herself and by Kaszniak et al (1978). The next three concern sentence and word length in an effort to examine some rather more qualitative aspects of performance. This type of analysis is common enough in aphasia research: Ludlow (1977) for example used sentence production and mean sentence length (among other things) in studying recovery from aphasia. The Type-Token ratio is a measure of repetitiveness or restricted vocabulary use which has a long history in studies of schizophrenic language (Maher, 1972). It has usually been used with samples of spontaneous speech or writing: it can be applied to any other variety of language output, though it may not then have the same meaning as it does with spontaneous output (as will be discussed below). The ratio is simply the number of different words used in a sample of output divided by the total number of words in the sample. Thus the two sentences

'The cat sat on the mat. The fat man sat on the cat.'

have 13 words in all but only 7 different words, i.e. 13 minus the 6 repeated words (underlined). This gives a Type-Token ratio of 7/13, i.e. about 0.54. Any sample which contained no repeated words would have a ratio of one; normal language use would of course have a ratio of considerably less than this. The sample used here for each subject consists of all correct sentences. Table 2 shows mean scores on the five measures broken down by diagnostic group.

Table 2 Mean number of correct sentences, length (in letters) of correct sentences, length (in words) of correct sentences, length (in letters) of words in correct sentences, and Type-Token ratio, broken down by diagnostic group. Number of cases shown at left beneath diagnostic category: the first number refers to measure (1), the second to the other four measures where subjects producing no correct sentences are excluded. Standard deviations in parentheses; significance of one-way Anova shown below.

	(1) n of sents.	(2) Sent l (letters)	(3) Sent l (words)	(4) Word l (letters)	(5) Type-Token
DAT (58,52)	4.14 (1.86)	26.5 (10.1)	7.62 (2.69)	3.46 (0.30)	0.77 (0.10)
MID (58,52)	4.53 (1.83)	24.8 (7.5)	7.14 (1.71)	3.44 (0.35)	0.77 (0.09)
Other (58,55)	4.93 (1.41)	26.3 (9.5)	7.39 (2.26)	3.53 (0.41)	0.76 (0.10)
Sig.	.05	ns	ns	ns	ns

Measure (1), the number of correct sentences, is the only measure to show a significant difference between the groups. The nature of this difference is very much consistent with overall levels of impairment as described in earlier chapters: i.e. best performance in the Other group, worst in the DAT group, with the MID group lying in between. The five measures were analysed broken down by age group as usual: there were no significant or interesting differences between age groups. The absolute levels of

the Type-Token ratios in the table cannot be compared with other published figures (such as Maher's) because of the unusual nature of the collection of the speech samples in the present study.

One factor contributing to the lack of inter-group differences on the various length and Type-Token measures may be some sort of principle of economy of effort in some subjects: more able subjects may realise that they can perform the task perfectly adequately by producing short, simple sentences which may be quite similar to each other in structure and wording. The test instructions do not suggest that long or elaborate sentences are necessary or desirable; the introductory examples are themselves short and simple. Less able subjects, on the other hand, may produce simple repetitive sentences because it is the best they can do. Consideration of certain individual cases confirms that at least occasionally subjects with undoubtedly excellent verbal skills will receive fairly low Type-Token ratios on this test.

Pearson product-moment correlations between the 5 measures, age, and overall Full score were calculated for each diagnostic group. The number of sentences produced correlated with overall Full score but not with age (except weakly negatively in the Other group). Relationships between the 5 measures of sentence and word length were unremarkable. In the MID group sentence length correlated with the number of sentences produced. In the DAT group, higher Type-Token ratios were associated with the production of fewer sentences. There may be some artefact in calculating Type-Token ratios when speech samples are extremely limited (e.g. if only two correct sentences are produced). In the MID and Other groups, lower Type-Token ratios were associated with the use of shorter words.

There were few interesting correlations between the five sentence production measures and other individual tests in the Full battery. There were some associations between estimated premorbid IQ and both the number of correct sentences produced and their length, but previous reservations about the validity of the premorbid IQ estimates must be borne in mind.

Change over time in scores on the five measures was studied in those subjects completing Full testing twice. By correlated t test, both the DAT and MID groups showed significant drops (at .10 level, 2-tailed test) in the number of correct sentences produced. None of the other measures showed significant changes in any group, though there was a tendency in the DAT group for subjects at second testing to use slightly longer words and to have, surprisingly, slightly higher Type-Token ratios.

Correlations calculated between scores at Test 1 and at Test 2 confirmed that the reliability of the Type-Token ratio may be low when used with this type of (possibly tiny) sample of speech output. In any case, Maher (1972) reminds us that abnormalities in Type-Token ratios are simply statistical indices of abnormality (analogous to high body temperature in medicine) which have limited potential for helping us understand the nature of the underlying problems.

The other measures suggest that there are few interesting differences between different groups of subjects with dementia on this particular rather artificial test of speech production.

Comprehension deficit.

As previously stated, impairments of comprehension are known to occur in dementia. Semple et al (1982), for example, noted considerable impairment of auditory comprehension, as judged by performance on the Token Test, in about half of a small sample of subjects with DAT. A brief and practically-oriented analysis of the assessment of comprehension deficit in this study will be presented, based primarily on performance on the Token Test. This is considered one of the most sensitive existing clinical tests for the detection of impairments of auditory comprehension; its main applications have been with aphasic subjects. Of the several existing versions of the Token Test, the one chosen was that described by De Renzi (1979). It consists of 36 instructions of increasing length and complexity which the subject must try to carry out upon an array of coloured tokens placed in front of him. The subject is allowed a second try at any failed item and scores half points if this try is correct. The administration and scoring of the test was summarised in Chapter 2, and appears in detail in Appendix 1. As this test was part of Full testing, subject details are exactly as described in previous chapters concerning subjects who completed Full testing.

First it is worth considering the numbers of subjects in the various diagnostic categories showing given levels of comprehension deficit as defined by De Renzi: These are shown in Table 3.

It can be seen that some impairment of comprehension, as judged by Token Test performance, is common in the DAT and MID groups. Even the mean scores in the DAT and MID groups are in the range indicating comprehension deficit as defined by De Renzi, i.e. less

Table 3 Numbers and percentages of subjects in each diagnostic category showing given levels of comprehension deficit as defined by De Renzi.

	DAT		MID		Other	
	n	(%)	n	(%)	n	(%)
none (36-29)	25	(43.1)	31	(53.4)	47	(81.0)
mild (25-28)	8	(13.8)	11	(19.0)	4	(6.9)
moderate (17-24)	17	(29.3)	8	(13.8)	4	(6.9)
severe (9-16)	6	(10.3)	5	(8.6)	3	(5.2)
very severe (0-8)	2	(3.4)	3	(5.2)	0	(0.0)
Total showing comprehension def.	33	(56.9)	27	(46.6)	11	(19.0)

than 29, and the mean for the Other group is only just above the cut-off. The figures are 25.2 (SD 7.6) in the DAT group, 26.7 (SD 8.1) in the MID group, and 30.9 (SD 6.1) in the Other group.

On a practical level it seems that testing clinically for comprehension deficit may be a sensible procedure, if only to gauge how easily a person may understand other test instructions (as well as ideally helping to provide information and advice to carers and relatives about the person's difficulties). An additional indication of the usefulness of the Token Test may be that it shows one of the highest correlations with overall Full score of any individual test.

If some brief clinical assessment of comprehension difficulties is required, the question arises as to which method might be the most appropriate. The 36-item Token Test was therefore compared to a 16-item version of itself (as described by Spellacy and Spreen, 1969) which uses pass-fail scoring with no second attempts. Score on this 16-item version could be computed from records of

performance on the 36-item one, so numbers of cases are as before. In addition, the 36- and 16-item versions were compared with another brief test of a type commonly used in aphasia examinations, previously referred to here as the 'Supplementary Dysphasia' test. This consisted of 12 items requiring the manipulation of three common objects to increasingly compex instructions. Details of this test again appear in Appendix 1. It was administered to only a proportion of subjects and so comparisons are based on relatively small numbers. Table 4 shows inter-correlations between the two versions of the Token test, the Supplementary Dysphasia test, overall Full score, and age in each diagnostic group and in the three combined.

Table 4 Inter-correlations between the 36-item Token test, the 16-item version, the Supplementary Dysphasia test, overall Full score, and age in each diagnostic group and in the three combined. Pearson product-moment correlations; ns=not significant at .05 level, 2-tailed test.

	16-item Token T.	Supp. Dysph.	Full score	Age
All (n=174)				
36-item Token	.96	.89 a	.81	-.34
16-item Token		.83 a	.79	-.33
DAT (n=58)				
36-item Token	.95	.87 b	.88	-.14 ns
16-item Token		.82 b	.81	-.11 ns
MID (n=58)				
36-item Token	.97	.91 c	.75	-.22 ns
16-item Token		.84 c	.73	-.20 ns
Other (n=58)				
36-item Token	.97	.90 d	.74	-.24 ns
16-item Token		.87 d	.72	-.22 ns

a: n=54, b: n=24, c: n=17, d: n=13.

The correlations between 36-item Token Test scores and Supplementary Dysphasia scores are high, but not as high as those between the 36-item and 16-item versions of the Token Test. Either

version of the Token Test would also seem to be more sensitive to mild impairments than the Supplementary Dysphasia in that several subjects scored maximum points on Supplementary Dysphasia but less than maximum on Token Test, while the reverse did not occur. (Another way of putting this is of course simply that the Token Test is more difficult; one might therefore also expect the Token Test to be better able to detect change over time in mildly impaired subjects because of the absence of a ceiling effect.) The Token Test also has advantages over tests like the Supplementary Dysphasia test used here in terms of minimising the contribution of agnosic or apraxic problems (De Renzi, 1979). Finally, the 16-item Token Test is scarcely longer than common tests of similar style to the Supplementary Dysphasia test: it is shorter than the one used here since the 12-item Supplementary Dysphasia allowed second attempts whereas the 16-item Token scoring did not. The Token Test would therefore seem to be the test of choice.

A regression analysis on all 174 subjects indicated that 16-item Token score = (0.564 x 36-item Token score) - 4.8

There are a number of possible explanations for the high correlation between Token Test and overall Full scores in dementia. Poor auditory verbal comprehension might affect understanding of all test instructions and hence interfere with all performances: this is implausible in view of the correlations between the Token Test and performance tests (e.g. Copying Designs) where verbal instructions are minimal and easily mimed, and in view of the good performance of some subjects on these and other tests despite very poor Token Test performance. Alternatively, it may be that all abilities tend to decline in dementia and that auditory

comprehension, as one of these abilities, declines apace: if this is the case then auditory comprehension may be seen as a 'core' or highly representative ability since its correlation with Full score is higher than that shown by most other tests. Thirdly, perhaps in dementia the Token Test is not a measure (or at least not only a measure) of auditory comprehension, but rather a measure of a combination of abilities including attention, concentration, immediate verbal memory, and so on: hence it correlates highly with an overall measure (Full score) because it can be affected by a number of different impairments, acting singly or in combination, at least one of which is likely to be present in dementia. In any type of subject group the separation of auditory comprehension and immediate verbal memory must be partly one of semantics or definition; it cannot be attempted on the basis of the data available here. As regards concentration, it was hoped that allowing second chances at failed items would reduce the contribution to the test score of momentary lapses of attention or concentration: the efficacy of this cannot be judged, but it is notable that correlations between Full score and the 16-item Token test (where no second chances were credited) are little lower than those between Full score and the 36-item version. How - ever, success at the second attempt on the 36-item version received only half the score awarded to initial success. If concentration problems were of paramount importance here, one might expect that, say, Digit Span would correlate as highly with Full score as does the Token Test, and that Digit Span would correlate very highly with the Token Test itself. This is not the case. Albert (1980) comments that the interaction of perseveration and memory disorders

(among other things) and the consequent 'incoherence of thought' make comprehension difficult to evaluate in dementia. It is true that some subjects in this study showed perseverative tendencies on this test; unfortunately the specific nature of errors was not consistently recorded, so that no formal error analysis is possible.

A previous analysis concerning which of all the tests used in this study were most sensitive to decline over 10 months showed that the Token Test seemed to perform well in this role. Table 5 shows correlations between drop in performance on the 36-item Token test, on the 16-item version, and on overall Full score in each diagnostic group and in the three combined.

Table 5 Intercorrelations between drop in performance on the 36-item Token test, on the 16-item version, and on overall Full score in each diagnostic group and in the three combined. Pearson product-moment correlations; ns=not significant at .05 level, 2-tailed test.

	16-item Token drop	Full score drop
All (n=54)		
36-item Token drop	.81	.56
16-item Token drop		.41
DAT (n=27)		
36-item Token drop	.76	.62
16-item Token drop		.41
MID (n=14)		
36-item Token drop	.95	.13 ns
16-item Token drop		.19 ns
Other (n=13)		
36-item Token drop	.89	.54 ns
16-item Token drop		.59

Drop on the 16-item version is representative of that on the 36-item version, but both are at best only moderately representative of the drop in Full score. Test-retest reliabilities (Pearson product-moment correlations) over the 10 months in subjects who

were retested were as follows:

	All (n=54)	DAT (n=27)	MID (n=14)	Other (n=13)
36-item Token	.91	.89	.94	.88
16-item Token	.84	.79	.87	.93

These are quite acceptable considering the length of time between testings.

It seems that a very short version of the Token Test, with only 16 items presented only once each, may be a useful tool in both the clinical and research evaluation of impairments of functioning in dementia.

Other tests classified as language tests and used in the main testing are: (1) Writing and Reading, in terms of both overall score and each component part (Writing name, writing a brief sentence and some numbers, reading the 'quick brown fox...' sentence and then writing that sentence); (2) Automatic Speech, in terms of both overall score and each of the two component parts (Counting from 1 to 20 and reciting the alphabet); (3) Sentence Repetition; and (4) Digit Span. Analyses of mean test scores broken down by diagnostic group and by age groups within diagnostic groups, followed by appropriate correlational analyses, produced no remarkable findings or trends which could not have been predicted from more general aspects of test performance in the various groups (such as patterns of factor scores or of overall Full scores): details of these analyses will therefore not be presented. Data on some of these tests, particularly Sentence Repetition and Digit Span, are insufficiently rich or detailed (in terms of potential for comparing one aspect or component of performance on the test against another) to merit further consideration. The brief ratings

of language characteristics completed after testing were similarly uninteresting.

CHAPTER 8

Further Analysis of Other Abilities.

Psychomotor performance.

Changes in psychomotor functioning are known to occur in normal ageing as well as in dementia. Psychomotor performance has more than one component, including a 'high-level' decision or control component as well as a more 'low-level' motor or movement time component. Psychomotor slowing with normal ageing is regarded as primarily a central, high-level phenomenon: in dementia the lower-level motor component may be more important (Miller, 1977), not simply as a result of slower peripheral nerve conduction.

Gilleard (1979, 1982) suggests that the Gibson Spiral Maze (GSM; Gibson, 1965) offers a useful measure of changes in psychomotor speed and accuracy in both normal and abnormal ageing of the central nervous system. It is included, with minor procedural modifications, in the widely used CAPE. Gilleard (1982) calculated correlations between psychomotor speed and psychomotor accuracy in different age groups of normal subjects (speed and accuracy being measured by GSM time scores and log-transformed GSM error scores respectively: log-transformation of error scores is commonly used to reduce the degree of skewness of the distributions of scores). In young subjects there was a strong negative correlation, in middle-aged subjects a weak one, and in old subjects none at all. From this he concluded that the age-related decline in speed and in accuracy were independent of each other and that different factors might be responsible for each. Alexander (1971) combined the GSM time and error scores in an 'Efficiency score' where Efficiency = raw time score x log(error score + 1). He found a general loss of this efficiency in old people with 'chronic brain syndrome' compared to normals and psychiatric controls. (The groups also

differed significantly on the time measure alone but not on error or log-transformed error score alone; they also, incidentally, differed significantly on Kendrick's Digit Copying task). Gilleard (1979) found significant differences on all three measures (i.e. time, log-tranformed errors, and efficiency) between normal controls, patients with senile or arteriosclerotic dementia, patients with unipolar depression, and patients with bipolar depression. The pattern of results agreed with previous studies (Alexander, 1971; Mayo, 1966; Blackburn, 1975) in showing a general impairment of GSM efficiency in dementia with a more specific impairment of the speed component in depression (particularly bipolar depression).

Alexander (1971) also computed a 'Strategy score' where Strategy = log(error score + 1) / raw time score. This was to give some indication of whether a subject's style of performance tended towards quick and careless or slow and sure (with the latter producing the lower strategy score), though he did not find significant differences between subject groups on this measure. One might expect to find disinhibited, impulsive subjects adopting a quick and careless approach (and this indeed was the type of thing the GSM was originally intended to assess in the field of delinquency).

The present study attempts to investigate relationships between these various parameters of psychomotor performance in different groups of impaired elderly subjects using a shortened but otherwise standard GSM, a shortened and simplified GSM (one with all the circular obstacles removed), and another simple psychomotor task involving filling a grid of boxes with marks like a number one for

30 seconds. Task complexity therefore varies. The administration of the tests was summarised in Chapter 2 and is described in detail in Appendix 1. As in previous work with elderly subjects, the time stresses used in Gibson's original instructions were omitted. Time stresses would be expected to lead to less conservative performance, producing lower time and higher error scores. Copies of the materials appear in Appendix 2.

As these tests were part of Short testing, subject details are as described in previous chapters concerning subjects who completed Short testing, except that here only subjects who completed both mazes will be considered: different aspects of performance cannot be assessed in those who failed to complete the mazes. One or both mazes were not completed by 14 DAT subjects, 11 MID subjects, and 10 Other subjects. This leaves 60 DAT subjects (18 male, 42 female; mean age 76.1 years, SD 9.4 years), 63 MID subjects (36 M, 27 F; mean age 69.9, SD 12.4), and 57 Other subjects (30 M, 27 F; mean age 56.8, SD 12.8). Subjects who completed the first maze almost always also managed to complete the second, so there is very little wastage of data concerning the first maze.

Alexander's efficiency and strategy measures will be used in addition to the time, error, and log-transformed error scores. (A measure akin to his efficiency measure was computed for previous analyses, but using the time score multiplied by the log of errors plus 10 rather than plus 1. In fact this produces score distributions closer to normal, which was desirable for purposes of factor analysis and the like, but here Alexander's version is used for the sake of consistency with the published work.) Maze 1 is the simpler maze with the obstacles removed, Maze 2 the more complex

one. Table 1 shows mean scores on the different measures for Mazes 1 and 2, and raw score on Box-filling, in subjects completing both mazes, broken down by diagnostic group.

Table 1 Mean time, raw error, log-transformed error, efficiency, and strategy scores for Spiral Mazes 1 and 2, and raw score on Box-filling, in subjects completing both mazes, broken down by diagnostic group. Standard deviations in parentheses. Significance of one way Anova shown at right of each row of three scores.

		DAT (n=60)	MID (n=63)	Other (n=57)	Sig
Maze 1	Time	37.6 (18.4)	39.4 (18.6)	30.7 (14.3)	.02
	Raw Errors	4.9 (10.0)	4.3 (7.9)	5.3 (9.3)	ns
	Log Errors	.454 (.480)	.458 (.460)	.471 (.504)	ns
	Efficiency	18.1 (26.5)	19.0 (23.3)	15.0 (19.8)	ns
	Strategy	.014 (.017)	.013 (.014)	.018 (.021)	ns
Maze 2	Time	71.7 (39.2)	78.8 (41.3)	58.7 (25.9)	.02
	Raw Errors	17.5 (17.3)	14.3 (13.4)	12.4 (15.3)	ns
	Log Errors	1.032 (.507)	.983 (.467)	.875 (.498)	ns
	Efficiency	76.1 (69.2)	79.1 (58.7)	54.0 (45.4)	.05
	Strategy	.018 (.014)	.015 (.010)	.017 (.013)	ns
Box-filling	Raw	37.1 (17.6)	33.9 (16.0)	46.8 (16.3)	.001

The groups differ significantly on the time measure on each maze and on performance on Box-filling: in each case the differences are in the direction of quickest performance in the Other group, slowest in the MID group, with performances in the DAT group being

marginally quicker than those in the MID group. The efficiency measure in Maze 2 shows significant differences of a similar nature. (Using an efficiency score equal to the time score multiplied by the log of errors plus 10 rather than plus 1, as described above, the groups do differ significantly in efficiency on both mazes.) Error scores show no significant differences between groups, and the strategy scores indicate no consistent differences in style of performance in different groups.

The various measures were analysed broken down by age group as in previous analyses: details will not be presented as a simple correlational analysis seems to demonstrate the main points. Pearson product-moment correlations between scores on the different measures for Mazes 1 and 2, raw score on Box-filling, age, and overall Short score were calculated for each diagnostic group and for the three combined.

In the DAT, Other, and combined groups there were small but often significant correlations, in the expected directions, between age and the time, error, and efficiency measures; in the MID group such correlations were close to zero. Correlations between age and strategy were small but mostly positive (indicating a trend for older subjects to be more quick and careless rather than slow and sure): the correlations reached significance only for Maze 1 in the DAT group and Maze 2 in the Other group.

In all groups there were moderate correlations, in the expected direction, between overall Short score and the time, error, and efficiency measures, as one would expect given the contribution of the maze measures to the Short score. Small negative correlations found between Short score and strategy merely indicate that more

impaired subjects are more likely to have a relatively quick and careless style.

The correlation between time and log-transformed error score was small and non-significant for both mazes in each of the three groups, consistent with Gilleard's (1982) finding that in elderly subjects the two measures are largely independent and do not show the inverse relationship (i.e. negative correlation) seen in younger subjects. Here, in fact, the correlations were positive: this presumably bears some relationship to general level of impairment, with more impaired subjects tending to be both slower and less accurate than relatively able subjects. There were no consistent patterns of change in the size of correlations between time and log-transformed error scores on each maze in different age groups within diagnostic groups. The absolute sizes of correlations were small, and almost none were significant. Other correlations calculated produced no surprises bearing in mind the methods of calculation of the various measures.

All the measures were analysed broken down by sex in each diagnostic group: few significant differences emerged even at the lax criterion level of .10. There were no significant sex differences in the DAT group. In the MID group, males had higher strategy scores (i.e. were less cautious) than females on Maze 1; males also scored more highly on Box-Filling. In the Other group, males made fewer errors on Maze 2 than did females, and were also more cautious on this maze: these findings are of no importance in view of the mixed nature of the group. The measures were similarly broken down by drug use (psychoactive medication vs none) in each group. There were no significant differences in the DAT or MID

groups. In the Other group, subjects on medication were slower on both mazes and less efficient on the Maze 2: again this may well be explicable in terms of a confounding of drug use and diagnostic category rather than a simple effect of drug use impairing performance.

Changes over time between Tests 1 and 2 will now be considered in subjects who completed both mazes on both test occasions. (A pattern of completion at Test 1 followed by non-completion at Test 2 may indicate decline, but cannot help elucidate the nature of that decline.) Discounting subjects who failed to complete one or both mazes on either occasion, we have 29 DAT subjects (10 male, 19 female; mean age 73.9 years, SD 8.7 years), 18 MID subjects (11 M, 7 F; mean age 71.7, SD 8.0), and 16 Other subjects (5 M, 11 F; mean age 61.3, SD 12.7). Table 2 shows mean scores on the different measures for Mazes 1 and 2, and raw score on Box-filling, at Test 1 and at Test 2 in subjects completing both mazes on both test occasions, broken down by diagnostic group.

On Maze 1 the DAT subjects have not become slower over time but have become significantly less accurate, and there is consequently a significant deterioration in efficiency score. Because of the increase in errors with no increase in time, a significant difference appears in the strategy measure in the direction of subjects becoming relatively more quick and careless with the progression of time (and presumably of their dementia). On Maze 2 the only significant change is in efficiency, comprising trends to be both slower and less accurate. Box-filling too shows a significant deterioration. Trends are very similar in the MID group but here, partly because of smaller changes and partly because of

Table 2 Mean time, raw error, log-transformed error, efficiency, and strategy scores for Spiral Mazes 1 and 2, and raw score on Box-filling, at Test 1 and at Test 2 in subjects completing both mazes on both test occasions, broken down by diagnostic group. Standard deviations in parentheses. Significance of correlated t test shown beside each pair of scores: + .10 level, * .05 level, ** .01 level, 2-tailed test.

	DAT (n=29)		MID (n=18)		Other (n=16)	
	t 1	t 2	t 1	t 2	t 1	t 2
Maze1 Time	37.3 (18.2)	37.4 (16.2)	46.1 (21.0)	45.0 (22.0)	29.7 (8.5)	31.4 (16.5)
Raw Errors	3.9 (12.0)	7.7 (*) (10.7)	3.6 (5.6)	5.1 (9.4)	4.6 (10.7)	8.0 (14.3)
Log Errors	.335 (.436)	.653 ** (.511)	.383 (.477)	.513 (.470)	.373 (.501)	.519 (.609)
Efficiency	14.7 (31.4)	25.7 * (27.4)	21.8 (30.7)	25.3 (33.0)	10.6 (14.6)	19.4 (33.9)
Strategy	.009 (.011)	.020 ** (.017)	.007 (.009)	.012 * (.010)	.014 (.018)	.018 (.022)
Maze2 Time	69.9 (37.4)	76.2 (42.7)	84.7 (34.2)	88.0 (34.9)	54.1 (19.2)	57.5 (25.1)
Raw Errors	15.7 (18.7)	20.1 (19.9)	12.7 (12.9)	15.1 (15.3)	15.5 (20.1)	18.4 (18.4)
Log Errors	.946 (.538)	1.080 (.538)	.936 (.465)	1.049 (.381)	.895 (.585)	1.080 * (.454)
Efficiency	69.3 (74.0)	84.5 * (80.7)	85.4 (65.9)	96.3 (60.8)	47.5 (34.9)	61.3 (43.0)
Strategy	.017 (.015)	.018 (.016)	.012 (.008)	.013 (.006)	.020 (.020)	.022 (.012)
Box-filling Raw	39.8 (16.6)	34.8 * (15.4)	29.2 (14.8)	30.2 (15.3)	44.4 (17.1)	44.1 (17.1)

smaller subject numbers, only the change in strategy score on Maze 1 reaches significance. The only difference between the DAT and MID groups in terms of the trends is that the MID subjects show no decline on Box-filling. Trends in the Other group are unremarkable and similar to those in the MID group.

Test-retest reliabilities of all measures (using Pearson product-moment correlations in all subjects completing both mazes on both occasions) were acceptable apart from that of the Maze 1 strategy score, with a correlation of only 0.43.

No striking differences between groups have emerged in this study. A number of factors might affect performance on these tests as well as basic psychomotor abilities (which presumably involve various functions including visual perception, hand-eye coordination, motor control and proprioception, motor speed, some kind of mental speed, and so on). Strategy - relatively quick and careless as against slow and sure - has already been mentioned: this may reflect premorbid personality as well as effects of the dementia itself. Other aspects of personality and motivation may also be relevant. Another important factor might be memory since in the mazes subjects have to remember (even allowing for the use of prompts and reminders) the fact that they are supposed to avoid the obstacles (where present) and the black walls. A casual observation was that some subjects made many errors because they did not care that they were sailing straight through obstacles, while others did so because they seemed to forget that they were supposed to avoid them. The limited results of the present study provide few clues as to the nature and importance of such factors, but they may have contributed to the lack of clarity of patterns of performance found.

Correlations between the three psychomotor test scores and other tests appear in Appendix 3, and are generally not very high. Scores on the three tests tended to come out as constituting a separate factor in factor analyses of scores in Full testing. These

observations suggests that slowing itself is only one component of the cognitive impairment seen in dementia, and not necessarily a major one as has been suggested in the past.

Motor apraxia.

Motor apraxia was not routinely assessed, for practical reasons described in Chapter 2, despite its reputation as a relatively common feature of dementia; but it was later decided to assess subsamples of subjects using Kertesz & Hooper's (1982) test. This particular test was chosen because it is reasonably brief, does not penalise the subject for requiring a demonstration of the required action (thus minimising the possible contribution of receptive dysphasia), provides fairly clear instructions concerning the scoring of performances, and is reported to have high inter-rater reliability. The test consists of 20 items, each scored 0, 1, 2, or 3 according to quality of performance: 3 represents normal performance while 0 is applied to completely unrelated or unrecognisable attempts at performing the required actions. There are four sections with 5 items in each: 'Facial' (e.g. whistle), 'Upper limb' (e.g. wave goodbye), 'Transitive' (e.g. pretend to use a spoon to eat with), and 'Complex' (e.g. pretend to light a cigarette). Certain items permit the use of actual objects if the action is not performed adequately in mime. The complete test appears in Appendix 1.

The test was carried out with 51 subjects, diagnosed in the usual way, all of whom had just completed Full testing as part of the main study: 25 DAT subjects (8 male, 17 female; mean age 74.6 years, SD 8.1, range 54-85), 23 MID subjects (14 M, 9 F; mean age 75.3 years, SD 10.2, range 44-92), and 3 subjects who turned out to

have Other diagnoses and who will be considered no further. 8 of the DAT subjects were on psychoactive medication (hypnotic 3, major tranquilliser 4, two of these 1) and 12 of the MID subjects (anti-depressant 1, major tranquilliser 8, two drugs 1, and 'other' psychoactive medication 2). The average age of the DAT subjects is clearly not significantly different from that of the MID subjects. The subject samples should not be biased in any particular way since the subjects were basically the last 25 DAT subjects and the last 23 MID subjects to receive Full testing, after the author had decided that testing for apraxia might be worth while. Table 3 shows mean scores of the two groups on each of the four sections of the test and on the test as a whole.

Table 3 Mean scores of the two groups on each of the four sections of the test and on the test as a whole. (Standard deviations in parentheses; significance of the difference between groups by independent t test, 2-tailed test, shown below.)

	Facial	Upper limb	Transitive	Complex	Total
DAT (n=25)	13.0 (1.7)	13.2 (1.7)	11.1 (3.6)	9.7 (3.4)	46.9 (9.3)
MID (n=23)	13.7 (1.3)	13.9 (1.1)	12.7 (1.3)	11.2 (2.3)	51.5 (4.9)
Sig. level	.10	.10	.05	.10	.05

The best possible score on each section of the test is 15, and on the test as a whole 60. Scores considered apraxic are those less than or equal to 14, 12, 12, 11.7, and 49.7 in the respective categories. It can readily be seen from the table that the subjects tested were on average not very apraxic; but some individuals were (as the standard deviations suggest). The DAT subjects are significantly worse than the MID subjects on all sections at a lax criterion level of .10. It might seem that motor apraxia is more

common in DAT than in MID; but DAT subjects in this study were worse than MID subjects on many tests and so this finding may well reflect general differences in level of functioning rather than specific differences to do with apraxia.

Appropriate Pearson product-moment correlations were calculated. In the DAT group, scores on the four sections and score on the whole test correlated significantly with overall Full score but not with age. Intercorrelations between the five scores (total and the four subsection scores) were all highly significant. In the MID group the pattern of correlations was similar except that the intercorrelations between apraxia test scores were rather lower (occasionally failing to reach significance).

Analysis of scores broken down by sex and by drug use (psychoactive medication vs none) in each group and in the two combined produced few notable results. In both groups lumped together, males were significantly better than females on the Upper limb section; such a trend existed in each group separately, though in the DAT group it did not reach significance. The only trend in the drug analysis was for DAT subjects on medication to be slightly (non-significantly) worse on the Upper limb and Transitive sections than were those on no medication.

An analysis of which other abilities motor apraxia correlates with most strongly may be of interest despite the unexciting group differences described above. Kertesz & Hooper (1982) found relationships between motor apraxia (as assessed by this test) and tests of aphasia in a large group consisting largely of dysphasic stroke patients. They found that degree of apraxia correlated best with degree of comprehension deficit and next best with an overall

severity of aphasia score (though Broca's aphasics tended to have the most severe apraxia, including Facial apraxia). Of the non-verbal tests used, scores on a drawing test correlated most highly with apraxia scores.

Here, in the DAT and MID groups lumped together, the total apraxia score correlated significantly with almost every test comprising Full testing. Correlations over .7 occurred with the Token Test, Reading and Writing, Miscellaneous Visuo Tasks, and Object Recognition. In DAT subjects alone, total apraxia score again correlated significantly with most other tests; correlations over .7 occurred with Automatic Speech, the Token Test, Reading and Writing, Sentence Repetition, Sentence Production, Arithmetic, Misc Visuo Tasks, Object Recognition, and Nominal Ability. In MID subjects alone, total apraxia score once again correlated significantly with most other tests, but correlations over .7 occurred only with Misc Visuo Tasks and Object Recognition. Correlations with factor scores described previously were also calculated. In the DAT group and in all the subjects combined, apraxia score correlated most highly with Factor 2 ('parietal') and Factor 4 ('object recognition and nominal ability') from Full testing, and with Sh. factor 2 (the 'verbal' one) from Short testing. In the MID group there was no very high correlation with any factor.

It seems that although performance on the apraxia test is related to overall levels of test performance in both subject groups, it is particularly strongly associated with tests of language function in the DAT group. Such an association with severity of language disturbance parallels Kertesz & Hooper's

findings in aphasia. Separate testing for apraxia seems justified in dementia on clinical grounds.

Spatial span.

Corsi's block-tapping test (Milner, 1971) provides a measure of 'spatial span' by asking subjects to touch a series of blocks in an irregular array in the same order as they have been touched by the examiner, with the sequences lengthening progressively as in the usual verbal Digit Span. Using this test, Grossi et al (1977, quoted in Cantone et al, 1978) found a significant difference between spatial span and ordinary verbal digit span in DAT but not in a group of chronic choreics. Cantone et al (1978) followed this up in subjects with DAT, MID, 'Senile dementia', 'Simple cerebral atrophy', and Huntington's chorea (and normal controls). Their subject classifications would now be considered rather unusual, but they found that all patient groups had lower verbal and spatial spans than did controls. A ratio of verbal: spatial span was calculated, and on this only DAT and Senile dementia groups differed significantly from controls. Relevant ratios in the present context were as follows:

Controls 1.09 (SD 0.20, n=90); DAT 3.29 (SD 1.70, n=21); Senile dementia 2.00 (SD 0.50, n=7); and MID 1.46 (SD 0.83, n=25).

The ratios indicate that DAT subjects had much lower spatial spans than verbal ones, whereas controls had spatial spans almost as long as their verbal spans. For some reason (perhaps deliberately to provide roughly equal verbal and spatial spans in control subjects) the authors tested spatial span using five trials of any given length, judging that length 'passed' if the subject performed at least three of the five sequences correctly; with verbal span they

adhered to the usual procedure with a maximum of two trials of any given length, only one of which need be performed correctly. Corsi and Milner both advocate the use of a two-trial procedure on spatial span analogous to that used in verbal digit span.

Cantone et al discuss the possible clinical utility of scores on the two tests and of the ratio of one to the other. Samples of DAT and MID subjects were therefore tested on a two-trial version of the Corsi test, ordinary Digit Span forward having been given as a routine part of testing in the main study. Precise details of the test and materials appear in Appendices 1 and 2: the subject is shown a board with 9 blocks attached to it in an irregular pattern and numbered on the side facing the tester. The task is explained to the subject, and the tester taps out increasingly long sequences (using the same actual numbers as in the verbal Digit Span test), allowing a maximum of two sequences of any given length. The practice of adding .4 to the raw span score for correct performance on the first trial of any given length as used in Digit Span in the main study was deleted.

The test was carried out with 50 subjects, diagnosed in the usual way, all of whom had just completed Full testing as part of the main study: 24 DAT subjects (8 male, 16 female; mean age 75.4 years, SD 7.0, range 56-85), 23 MID subjects (14 M, 9 F; mean age 75.3 years, SD 10.2, range 44-92), and 3 subjects who turned out to have Other diagnoses and who will be considered no further. 8 of the DAT subjects were on psychoactive medication (hypnotic 3, major tranquilliser 4, two of these 1) and 12 of the MID subjects (anti-depressant 1, major tranquilliser 8, two drugs 1, and 'other' psychoactive medication 2). The average age of the DAT subjects is

clearly not significantly different from that of the MID subjects. The subject groups are almost identical to those receiving the apraxia test previously described and again should not be biased in any particular way.

Three scores were computed from the two spans for each subject:

(1) Span Difference = Verbal span - Spatial span.

(2) Abs. Span Difference = the absolute value of (1), i.e. ignoring any negative sign (since a large difference between spans may be interesting regardless of the direction of the difference; MID subjects might be expected to show larger absolute difference scores, owing to the nature of their pathology, than might DAT subjects).

(3) Span Ratio = Verbal span / Spatial span.

Table 4 shows the mean raw spans and mean values of the three computed variables in each diagnostic group.

Table 4 Mean values of Verbal span, Spatial span, and three computed variables in each diagnostic group. (Standard deviations in parentheses; significance of the difference between groups by independent t test, 2-tailed test, shown below.)

	Verbal Span	Spatial Span	Span Difference	Abs.Span Difference	Span Ratio
DAT (n=24)	5.5 (1.1)	3.5 (1.2)	2.0 (1.3)	2.0 (1.2)	1.80 (0.89)
MID (n=23)	5.4 (1.2)	4.2 (1.1)	1.2 (1.5)	1.7 (1.0)	1.39 (0.52)
Sig. level	ns	.07	.08	ns	.07

Mean Verbal Span is almost identical in the two groups but, at a lax criterion level of .10, the DAT subjects are significantly worse than the MID subjects on Spatial Span. This leads to significantly higher Span Difference and Span Ratio scores in the DAT group (again at a lax criterion level). The Absolute span

difference is also higher in the DAT group, but not significantly so. Judging from these small samples, it would seem that poor spatial span compared to verbal span is more characteristic of DAT than of MID. DAT subjects in this study were worse than MID subjects on many tests, but interpretation of the findings in terms of general differences in level of functioning is implausible: ceiling effects are impossible in the type of test used, so there is no question of the lack of difference between groups on Verbal Span being attributable to this. Similarly there are no signs of floor effects.

Pearson product-moment correlations between Verbal Span, Spatial Span, difference and ratio measures, age, and overall Full score were calculated in each diagnostic group and in the two combined. None of the measures correlated significantly with age. Verbal and Spatial spans correlated with overall Full score. Intercorrelations between test measures were in expected directions; the correlations between Verbal and Spatial spans were weak, suggesting considerable individual variation (especially in the MID group) in the relative levels of performance on the two tests. The correlation between Span Difference and Absolute Span Difference was high in DAT and low in MID, suggesting that Verbal Span is generally higher than Spatial Span in DAT but that this is not so consistently the case in MID. Spatial Span correlated more highly with Full score than did Verbal Span. None of the measures correlated significantly with estimated premorbid IQ.

The five measures were analysed broken down by sex in each diagnostic group and in the two combined. There were no significant sex differences, indicating that the group differences are not

explicable in terms of the greater proportion of males (who might generally have better spatial compared to verbal skills than females) in the MID group. The measures were similarly broken down by drug use (psychoactive medication vs none). In all subjects lumped together, those on drugs had a significantly lower mean Span Difference than did those on none. In the DAT group those on drugs had a lower Verbal Span than the others, while in the MID group those on drugs had a higher Spatial Span and a lower Span Ratio than did those on none. The proportion of subjects taking drugs was slightly higher in the MID than in the DAT group (12/23 compared to 8/24): therefore drug use may have contributed to the group differences found, though it seems unlikely that it can entirely account for them. It is quite plausible that there is no causal relationship between drug use and scores on the various measures here: as previously discussed, the reasons for subjects being on drugs may in part depend on aspects of their cognitive state. Intuitively speaking, it would be surprising if drug use could lower Verbal but not Spatial span in one group while raising Spatial but not Verbal span in the other. The numbers of subjects are of course too small to permit breakdown of drug types.

In the DAT and MID groups lumped together, Spatial Span (and Verbal Span) correlated significantly with many of the tests in Full testing. Correlation over .7 occurred between Spatial Span and Copying Designs. In DAT subjects alone, Spatial Span again correlated significantly with many other tests; correlations over .7 occurred with the Token Test, Reading and Writing, Copying Designs, and Porteus Mazes. In MID subjects alone, Spatial Span once again correlated significantly with some other tests, but

there were no correlations over .7. (Verbal Span correlated over .7 with no other test in either group or in the two combined.) It seems that, at least in MID subjects, Spatial Span measures an ability (or abilities) which is reasonably independent of abilities assessed by other tests used here. One would imagine that the abilities required for success on the test would involve a mixture of visuospatial perception and immediate non-verbal memory, though this study cannot identify the necessary components. The correlations in DAT subjects between Spatial Span and the Token Test and Reading and Writing perhaps represent the influence of some common 'parietal lobe' contribution to these tests.

As suggested above, the differences between DAT and MID subjects are not easily attributable to artefact. The most parsimonious interpretation of the results is simply that the DAT subjects were more spatially impaired. The study lends a little support to Cantone et al's (1978) claim that the verbal-spatial span combination could be a useful tool in discrimination of basic cognitive disabilities in dementia.

Drawing ability.

Moore & Wyke (1984) asked DAT subjects to draw a house, a cube, a face, and a clock spontaneously and then to copy drawings and models of a house and a cube. Impairments were present on all types of drawing task, and level of performance was related to severity of dementia. The drawing difficulties were not particularly similar to those seen in patients with focal lesions or to those seen in small children. Many subjects in this study were asked to draw a clock face and a person spontaneously in addition to completing other visuospatial or constructional tasks in the main study such

as the copying of paper-and-pencil designs, Block Design, Miscellaneous Visuospatial Tasks and so on. The drawing tasks differ from all the other tests (apart from constructing a square with sticks) in that they are spontaneous constructional tasks where no model is provided.

The numbers of subjects asked to draw a clock face, all of whom had just completed Full testing as part of the main study, were 52 in the DAT group, 47 in the MID group, and 50 in the Other group. The numbers subsequently asked to draw a person as well were 32, 18, and 12 respectively. Subject selection depended partly on whether the author thought that a subject was willing to tolerate extra tests after having already completed a considerable number, so the subject groups may not be representative of the whole groups. Details of test procedure and scoring methods (based on the completeness and accuracy of the drawings) are given in Appendix 1.

Analyses of the mean test scores broken down by diagnostic group and by age groups within diagnostic groups showed no remarkable findings or trends which could not have been predicted from more general aspects of test performance in the various groups (such as patterns of factor scores or of overall Full scores): details of the analyses will therefore not be presented.

In all 149 subjects lumped together, score on drawing a clock face correlated significantly with many other tests comprising Full testing, and over .7 with Copying Designs, Memory for Designs, Block Design, and Miscellaneous Visuo Tasks. In DAT subjects alone, scores again correlated significantly with many other tests, and over .7 with Memory for Designs and Block Design. In MID subjects alone, scores once more correlated significantly with many other

tests, and over .7 with Copying Designs and Memory for Designs. In Other subjects alone, scores correlated over .7 with Memory for Designs and Block Design as well as correlating significantly with other tests. Scores on drawing a person correlated over .7 with Copying Designs and drawing a clock in all 62 subjects attempting it; with Copying Designs and drawing a clock in the DAT subjects; with Copying Designs, Memory for Designs, Block Design, the Weigl test, Spiral Maze 2, and drawing a clock in the MID subjects; and only with drawing a clock in the Other subjects.

Performance on these simple spontaneous drawing tasks is related to overall levels of test performance in all subject groups, and is so strongly associated with other tests of visuospatial and constructional ability that one may wonder whether in fact spontaneous drawing measures anything different or unique in these subject populations. The fact that there are some high correlations between scores on spontaneous drawing and Memory for Designs might suggest that spontaneous drawing is tapping some ability to remember or visualise the appearance of objects. To some extent of course it must do but, as previous analyses have shown, the Memory for Designs test is itself a test of constructional or perceptual ability as well as of memory. More detailed testing of the various abilities would be required to investigate the relationships between these abilities: data from the present study are insufficiently detailed. As with many other single tests, drawing ability may not be as specific or as interesting a feature in dementia as has sometimes been suggested.

The other tests classified as visuospatial or constructional tests are: (1) Block Design; (2) Miscellaneous Visuo. Tasks, in

terms of both overall score and each component part (Square with Sticks, Multiple-choice Block Designs, Yerkes test, Time-telling, Poppelreuter figures, recognition of hatched pictures); and (3) Copying Designs (and its relationship to Memory for Designs). Again, data from these tests were analysed in terms of mean scores broken down by diagnostic group and by age groups within diagnostic groups, followed by appropriate correlational analyses. Again no remarkable findings or trends emerged which would not have been predictable from more general aspects of test performance in the various groups (such as patterns of factor scores or of overall Full scores).

Identical comments apply to the other tests used in the main study which have not already been discussed. These are: (1) Porteus Mazes; (2) Weigl Test; and (3) Arithmetic, both in terms of total score and the two component parts (WAIS Arithmetic items and straightforward mental calculations not presented in a problem-style format).

Here again, as with several of the tests referred to in this thesis, it seems to be the case that particular impairments or performances on isolated clinical tests in dementia may become considerably less interesting and noteworthy when considered in the broader context of overall levels or patterns of impairment.

CHAPTER 9

Concluding Comments.

The main findings of the studies presented can be briefly reviewed as follows.

The main initial study, presented in Chapter 2, showed slight average differences between the patterns of cognitive impairment in DAT compared with MID which are of questionable significance and which can have no differential diagnostic power in the individual case. However it also showed differences between different age groups within diagnostic groups which are probably not entirely explicable in terms of artefacts involving overall level of impairment, duration of condition, estimated premorbid IQ, use of medication, gender, or aspects of methodology or subject selection. The most plausible hypothesis concerning the age-related differences within DAT is that at least some of these are related to the various recently discovered structural and chemical differences between subject groups within DAT, as described in the introduction. The fact that the age-related patterns in MID were not entirely dissimilar invites caution in attributing the DAT differences to the existence of different sub-types of DAT, but the study lends support to other clinical studies suggesting a degree of non-random heterogeneity within DAT - non-random in the sense that patterns bear some demonstrable relationship to another variable (in this case, as in some other studies, age at testing). This perhaps deserves to be taken into consideration in studies concerning the aetiology, treatment, or management of DAT: different types of subjects may differ in important ways as regards aetiology or response to a given intervention, and issues may be clouded if heterogeneous samples are used.

No differences were found between diagnostic groups on two

measures of how focal their neuropsychological impairments tend to be (one measure being the computed focality score, the other being based on sizes of intertest correlations). On the first measure, no differences were found between subgroups within diagnostic groups (while the second was not applied to subgroups). Hence there was no support for the idea that certain groups of subjects with dementia are likely to show more focal patterns of performance (i.e. any focal pattern as opposed to one focal pattern in particular) than are others.

From a clinical point of view, perhaps the most interesting results from the initial study concerned inter-individual variability. Despite the identification of different average patterns of impairment in subgroups within diagnostic groups, it became clear that there is considerable variation between individuals within any particular subgroup. A patient's diagnostic category will therefore tell us little about the details of the impairments he is likely to show, and level of performance in one area of functioning will not accurately predict level of functioning in another. A strong case can be made for idiographic assessment (however that assessment is accomplished); individuals must be considered on their merits, and interventions tailored accordingly.

Factor structures of abilities did not differ markedly between different diagnostic categories, but in no group was the pattern of performance describable by a single factor: cognitive impairment in dementia cannot be seen as unidimensional.

Subjects retested in the follow-up study described in Chapter 3 seemed in general to be reasonably representative of those

initially assessed. Decline over the 10 months was shown to be demonstrable on testing. The extent of the decline showed considerable variation even within diagnostic groups, and was only rather weakly predictable using information concerning subjects' test performances or personal characteristics at initial assessment. It was not strongly related to age, level of impairment, or the relative severity of 'parietal' impairments. The levels of statistical prediction achieved, even with this post hoc analysis on specific samples of patients, were well below any that would allow reliable predictions in individual cases.

Some tests registered decline while others did not: memory tests of the kind traditionally used did not seem to be the most successful detectors of change in these groups of subjects with established dementia, though this does not cast doubt on their possible superiority in detecting deterioration in the earliest stages of dementia where diagnosis may be in doubt.

Differential decline in different Factors could not however be explained entirely in terms of differential test sensitivity (or floor and ceiling effects and so on): the fact that some abilities showed significant decline while others did not lends further support to the impression gained from factor analyses that decline in dementia cannot be seen as a unidimensional process.

No changes over the 10 months occurred on two measures of how focal the subjects' impairments seemed to be (i.e. the computed focality scores and the mean sizes of intertest correlations). Taken with the original analyses of focality scores, this provided no evidence for de-differentiation of cognitive abilities in dementia with time and progression. This implies that typical

global patterns of impairment can no more easily be assumed in advanced dementia than they can in the early stages, and that identification of strengths and weaknesses still requires individual consideration.

Despite the 10-month gap between assessments, test-retest reliabilities were found to be adequate, suggesting that the initial test performances found are quite likely to be dependable.

The study presented in Chapter 4 confirmed that relationships between neuropsychological test performance and everyday functioning do exist in dementia. The relationships found were not strong enough to allow prediction of everyday impairments on the basis of test performances, but their existence is important as regards other stated goals of neuropsychological asessment such as the explanation of observed behavioural deficits: if tests show no association with these deficits then such an explanatory link becomes tenuous. In this respect the presence of certain specific associations between particular test performances and particular rated behavioural deficits is encouraging. Behavioural ratings made on the basis of a single test session (or pair or triplet of sessions where testing was split up) also showed meaningful relationships with everyday behaviour as assessed by independent raters.

Rating scale measures seemed to be at least as sensitive to decline over 10 months as were test measures, but this may in part reflect the contribution of purely physical deterioration to score changes on the particular rating scale used.

The study of short term day-to-day variability in cognitive functioning in DAT and MID subjects contained in Chapter 5

indicated that a small amount of variability was present in both groups, but that the degree of variability was no greater in the MID than in the DAT subjects. Subject numbers were small, but the notion that day-to-day fluctuation is a characteristic or universal feature of MID rather than of DAT is not supported. The study suggests that isolated assessments of MID subjects will not necessarily be unreliable, and also provides further evidence of the adequate reliability of tests used in the main study.

The study concerning inhibition of competing responses in recognition memory in DAT, MID, and Korsakov's syndrome in Chapter 6 showed almost no differences between groups, but did suggest that the phenomenon was occurring with some types of material. This could have implications as regards environmental manipulations, particularly in view of the fact that the 'competing responses' here were externally provided: a failure of inhibition of purely internally-generated competing responses (as is presumed to be demonstrated by the results of experiments using letter-cueing to improve performance on word recall tests) can be considered harder to tackle in a practical way.

The study of encoding preferences in DAT, MID, and depression showed merely that the depressed subjects tested had milder memory difficulties than the other two groups and that the DAT subjects were more likely than the others to fail to process the material in any significant way at all, perhaps as a result of attention failures. A trite generalised conclusion might be that subects with moderately severe DAT might benefit particularly from strategies or interventions designed to attract, direct, or focus attention.

Depressed subjects were again found to have milder memory

difficulties than DAT and MID subjects in a signal detection analysis of recognition memory: significant group differences were found only on the sensitivity measure and not on the response criterion measure, though the pattern of learning over time may have been qualitatively different in the depressed group compared with the other groups.

Analyses of performance on a paragraph recall test, a face-name learning test, a brief picture-recognition memory test, and an orientation questionnaire (all from the main study) in DAT, MID, and Other conditions produced no remarkable findings or group differences, but of course this does not deny the potential value of such approaches in clinical work with individuals.

The study of recognition and nominal ability appearing in Chapter 7 showed no striking differences between DAT, MID, and Other groups but showed some interesting relationships between recognition and naming which are difficult to explain away entirely in terms of artefact and which suggest that a traditional model of object naming in dementia involving discrete and sequential stages of recognition and name-finding may not be tenable. The results of the study also suggest that nominal ability is not a particularly successful index of severity of dementia, and that many other tests are better.

A study of sentence production in DAT, MID, and Other conditions produced no notable findings, perhaps as a result of the limited nature of the test used and the dubious reliability of some of the measures of language production derived from it.

The analyses of comprehension deficit in DAT, MID, and Other conditions suggested simply that a very short version of the Token

Test may be a useful and practical tool in the assessment of dementia, even though it may not necessarily be measuring exactly the same abilities as it measures in, say, dysphasic stroke patients.

No clear results or patterns emerged from the study described in Chapter 8 concerning psychomotor performance in DAT, MID, and Other conditions. This may well reflect authorial misjudgement in assuming that such a complex and multifactorial ability or set of abilities could be adequately investigated using such a basic methodology and limited range of tests. However, the results of correlational and factor analyses did suggest that slowing itself is unlikely to be a principal contributor to other aspects of neuropsychological impairment in dementia.

A study of motor apraxia showed no differences between DAT and MID which might not have resulted from a general difference in overall level of impairment. The results from the DAT group confirmed the correlational association between motor apraxia and severity of language disturbance which is known to exist in dysphasic stroke patients.

A study of spatial block span compared to verbal digit span showed some differences between DAT and MID which are probably interpretable in terms of spatial disturbance being more common or severe in DAT (or at least in this particular DAT sample).

Drawing ability was found to be strongly related to other visuospatial abilities in DAT and MID and to show no interesting group differnces.

Analyses of other tests used in the main study produced no remarkable results. Sometimes this was partly because of the

limited nature of the data available from the tests; sometimes, however, the lack of notable findings resulted from the opportunity to view performances and group differences in the context of other areas of functioning or overall levels of severity of dementia. The point was reiterated that the significance of particular test deficits can be difficult to judge when viewed in isolation.

There are a number of ways in which this study might, in retrospect, have been more fruitfully conducted.

The availability of data from the Other subjects has been useful in some ways, but it is not difficult to argue that the time spent in assessing them would have been better spent in assessing more DAT and MID subjects (particularly at the follow-up stage, though reasons for the disappointing follow-up rate were largely beyond the author's control and almost entirely unrelated to time spent in the assessment of Other subjects).

More attention could have been given to gauging the stage or severity of dementia using operational criteria as described in papers reviewed in the introduction under the heading of 'staging of dementia', rather than having to adopt a retrospective and crude classification based on Orientation score or overall Full score.

The collection of detailed radiological data would have been desirable considering recent findings concerning inter-relationships between regional density of the brain on CT scan, neuropsychological impairments, and age of onset as described in the introduction. There would have been insurmountable practical difficulties in collecting such data on all subjects assessed in this study; but the decision not to pursue details of CT scanning in those subjects where it was performed (about half) because such

subjects would be a selected and unrepresentative group may have been badly taken. Similarly, post-mortem neuropathological data would be of interest: no plans to collect such data have been made, again because of major practical and other difficulties. In due course the author hopes to carry out an appropriate follow-up study concerning relationships between mortality and aspects of initial neuropsychological performance (and decline over 10 months).

In view of the evidence relating to the possible existence of familial forms of dementia with characteristic features, as described in the introduction, time spent in trying to gather data on family history might have been well spent, though this kind of data is notoriously difficult to collect reliably.

The collection of some normal control data in the main studies, despite the reservations expressed in the introduction, might have been helpful in a crude scaling of the severity of impairments on particular tests in dementia even if the control sample had not been a truly representative one. Adoption of Miller's (1977) suggested methodology to assess the degree of similarity between DAT and normal ageing, i.e. comparisons between DAT subjects and older controls matched on a comparison performance variable rather than on age, would have been most interesting given time.

Detailed experimental analyses of many abilities (for example visuospatial abilities) are lacking in this investigation: this is regrettable in view of the fact that understanding of impairments of such abilities in dementia is at present so limited (compared to understanding of, say, memory impairments). The limitation the author is most conscious of, however, concerns the lack of studies which not only describe the nature of certain impairments but

demonstrate how the impairments can best be minimised or ameliorated or managed. Miller (1981a, 1984) emphasises the importance of research which bears on management issues in dementia, particularly work elucidating the conditions under which the performance of neuropsychologically impaired subjects can be optimised.

Regrettably the studies presented have included no direct experimental manipulations which accomplish this; as usual the plea in mitigation must be limit of time. It is hoped that some of the descriptive information might still contribute to the larger task by clarifying some of the issues concerning neuropsychological assessment in dementia.

The time-consuming conduction and analysis of the studies described here have produced some interesting findings. These take only a drop from the ocean of what remains to be learned about the nature of cognitive impairments in the dementias and how these might best be assessed and managed. It seems entirely justified to conclude with the cliche that the area requires further research.

REFERENCES

Aggernaes A, Myschatzky A, Paikin H, & Vitger J (1975) Empirical investigations on the reliability and the value for differential diagnosis of 21 clinical symptoms of disturbed states of consciousness. Acta Psychiatrica Scandanavica

Albert M L (1980) Language in normal and dementing elderly. In: Obler L K & Albert M L (Eds) Language and Communication in the Elderly. Lexington Books, Lexington, Massachusetts

Albert M S (1981) Geriatric neuropsychology. Journal of Consulting and Clinical Psychology 49: 835-850

Albert M S, Butters N, & Levin J (1979) Temporal gradients in the retrograde amnesia of patients with alcoholic Korsakoff's disease. Archives of Neurology 36: 211-216

Alexander D A (1971) Two tests of psychomotor function in detection of organic cerebral damage in elderly psychiatric patients. Perceptual and Motor Skills 33: 1291-1297

Appell J, Kertesz A, & Fisman M (1982) A study of language functioning in Alzheimer patients. Brain and Language 17: 73-91

Baddeley A D (1975) Theories of Amnesia. In: Kennedy A & Wilkes A (Eds) Studies in Long Term Memory. Wiley, New York

Baddeley A D (1978) The trouble with levels: a re-examination of Craik & Lockhart's framework for memory research. Psychological Review 85: 139-152

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

Baddeley A D & Warrington E K (1970) Amnesia and the distinction between long- and short-term memory. Journal of Verbal Learning and Verbal Behaviour 9: 176-189

Baddeley A D & Warrington E K (1973) Memory coding and amnesia. Neuropsychologia 11: 159-165

Ballinger B R, Reid A H, & Heather B B (1982) Cluster analysis of symptoms in elderly demented patients. British Journal of Psychiatry 140: 257-262

Baltes P B (1968) Longitudinal and cross-sectional sequences in the study of age and generational effects. Human Development 11: 145-171

Baltes P B & Willis S L (1979) The critical importance of appropriate methodology in the study of ageing: the sample case of psychometric intelligence. In: Hoffmeister F & Muller C (Eds) Brain Function in Old Age. Springer-Verlag, Berlin

Barker M G & Lawson J S (1968) Nominal aphasia in dementia. British Journal of Psychiatry 114: 1351-1356

Bayles K A (1982) Language function in senile dementia. Brain and Language 16: 265-280

Bayles K A & Tomoeda C K (1983) Confrontation naming impairment in dementia. Brain and Language 19: 98-114

Benton A L (1968) Differential behavioural effects in frontal lobe disease. Neuropsychologia 6: 53-60

Blessed G, Tomlinson B E, & Roth M (1968) The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. British Journal of Psychiatry 114: 797-811

Bondareff W, Mountjoy C Q, & Roth M (1982) Loss of neurons of origin of the adrenergic projection to cerebral cortex (nucleus locus ceruleus) in senile dementia. Neurology 32: 164-168

Botwinick J (1978) Ageing and Behaviour. Springer, New York

Botwinick J & Birren J E (1951) Differential decline in the Wechsler-Bellevue subtests in the senile psychoses. Journal of Gerontology 6: 365-368

Bowen D M, White P, Spillane J A, Goodhardt M J, Curzon G, Iwangoff P, Meier-Ruge W, & Davison A N (1979) Accelerated ageing or selective neuronal loss as an important cause of dementia? Lancet 1: 11-14

Bower G H (1972) Mental imagery and associative learning. In: Gregg L (Ed) Cognition in Learning and Memory. Wiley, New York

Breitner J C S & Folstein M F (1984) Familial Alzheimer dementia: a prevalent disorder with specific clinical features. Psychological Medicine 14: 63-80

Brody M B (1942) The measurement of dementia. Journal of Mental Science 88: 317-327

Bromley D B (1974) The Psychology of Human Ageing (2nd Edition). Penguin, London

Brown G G (1980) The persistence of amnesia: data, theory and methods. Journal of Clinical Neuropsychology 2: 343-353

Brown J (1976) An analysis of recognition and recall and of problems in their comparison. In: Brown J (Ed) Recall and Recognition. Wiley, London

Brull J, Wertheimer J, & Haller E (1979) Evolutive profiles in senile dementia: a psychological and neuropsychological longitudinal study. In: Hoffmeister F & Muller C (Eds) Brain

Function in Old Age. Springer-Verlag, New York
Burt C (1921) Mental and Scholastic Tests. King, London

Butler R (1978) Alzheimer's disease-senile dementia and related disorders: the role of the National Institute on Aging. In: Katzman R, Terry R D, & Bick K L (Eds) Alzheimer's disease: Senile Dementia and Related Disorders. Aging Series, Vol 7. Raven Press, New York

Butters N & Cermak L S (1980) Alcoholic Korsakoff's Syndrome: An Information-Processing Approach to Amnesia. Academic Press, London

Cantone G, Orsini A, Grossi D, & De Michele G (1978) Verbal and spatial memory span in dementia. Acta Neurologica (Napoli) 33: 175-183

Cattell R B (1963) The theory of crystalline and fluid intelligence. Journal of Educational Psychology 54: 1-22

Cermak L S, Butters N, & Moreines J (1974) Some analyses of the verbal encoding deficit of alcoholic Korsakoff patients. Brain and Language 1: 141-150

Cermak L S, Reale L, & De Luca D (1977) Korsakoff patients non-verbal vs. verbal memory: effects of interference and mediation on rate of information loss. Neuropsychologia 15: 303-310

Cermak L S, Reale L, & Baker E (1978) Alcoholic Korsakoff patients' retrieval from semantic memory. Brain and Language 5: 215-226

Colgan J (1985) Regional density and survival in senile dementia: an interim report on a prospective computed tomography study. British Journal of Psychiatry 147: 63-66

Constantinidis J (1978) Is Alzheimer's disease a major form of senile dementia? clinical, anatomical, and genetic data. In: Katzman R, Terry R D, & Bick K L (Eds) Alzheimer's disease: Senile Dementia and Related Disorders. Aging Series, Vol 7. Raven Press, New York

Corkin S (1982) Some relationships between global amnesias and the memory impairments in Alzheimer's disease. In: Corkin S, Davis K L, Growdon J H, Usdin E, & Wurtman R J (Eds) Alzheimer's Disease: A Report of Progress in Research. Aging Series, Vol 19. Raven Press, New York

Craik F I M & Lockhart R S (1972) Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behaviour 11: 671-684

Davis P E & Mumford S J (1984) Cued recall and the nature of the memory disorder in dementia. British Journal of Psychiatry 144: 383-386

De Leon M J, Potegal M, & Gurland B (1984) Wandering and parietal signs in senile dementia of Alzheimer's type. Neuropsychobiology

11: 155-157

De Renzi E (1979) A shortened version of the Token Test. In: Boller F & Dennis M (Eds) Auditory Comprehension: Clinical and Experimental Studies with the Token Test. Academic Press, London

Dixon J C (1965) Cognitive structure in senile conditions with some suggestions for developing a brief screening test of mental status. Journal of Gerontology 20: 41-49

Drachman D A, Fleming P, & Glosser G (1982) The Multidimensional Assessment for Dementia scales. In: Corkin S, Davis K L, Growdon J H, Usdin E, & Wurtman R J (Eds) Alzheimer's Disease: A Report of Progress in Research. Aging Series, Vol 19. Raven Press, New York

Erickson R C & Scott M L (1977) Clinical memory testing: a review. Psychological Bulletin 84: 1130-1149

Eysenck M W (1979) Depth, elaboration, and distinctiveness. In: Cermak L S & Craik F I M (Eds) Levels of Processing in Human Memory. Lawrence Erlbaum, Hillsdale New Jersey

Farmer P M, Peck A, & Terry R D (1976) Correlations among neuritic plaques, neurofibrillary tangles and the severity of senile dementia (abstract). Journal of Neuropathology and Experimental Neurology 35: 367

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

Fox J H, Huckman M S, & Topel J L (1975) The use of computerised tomography in senile dementia. Journal of Neurology, Neurosurgery, and Psychiatry 38: 948-953

Gaffan D (1974) Recognition impaired and association intact in the memory of monkeys after transection of the fornix. Journal of Comparative and Physiological Psychology 86: 1100-1109

Gaffan D (1976) Recognition memory in animals. In: Brown J (Ed) Recall and Recognition. Wiley, London

Gainotti G, Caltagirone C, Masullo C, & Miceli G (1980) Patterns of neuropsychologic impairment in various diagnostic groups of dementia. In: Amaducci L, Davison A N, & Antuono P (Eds) Aging of the Brain and Dementia. Aging Series, Vol 13. Raven Press, New York

Gardner H, Boller F, Moreines J, & Butters, N (1973) Retrieving information from Korsakoff patients: effects of categorised cues and references to the task. Cortex 9: 165-175

Gibson H B (1965) Manual of the Gibson Spiral Maze. University of London Press

Gilleard C J (1978) An Investigation into the Nature of Behavioural

Change in the Dementias of Old Age. Unpublished Ph.D Thesis, Department of Psychiatry, University of Leeds

Gilleard C J (1979) Psychomotor performance of elderly psychiatric patients on the Gibson Spiral Maze test. Perceptual and Motor Skills 48: 678

Gilleard C J (1981) Origins, developments, and directions in the classification of functional status in the elderly patient. Journal of Psychiatric Treatment and Evaluation 3: 251-255

Gilleard C J (1982) Effects of aging on Gibson Spiral Maze performance. Perceptual and Motor Skills 55: 1098

Gilleard C J (1984a) Assessment of cognitive impairment in the elderly: a review. In: Hanley I & Hodge J (Eds) Psychological Approaches to the Care of the Elderly. Croom Helm, London

Gilleard C J (1984b) Dementia: the clinical syndrome. In: Gilleard C J, Living With Dementia. Croom Helm, London

Gilleard C J (1984c) Assessment of behavioural impairment in the elderly: a review. In: Hanley I & Hodge J (Eds) Psychological Approaches to the Care of the Elderly. Croom Helm, London

Goodglass H & Kaplan E (1972) The Assessment of Aphasia and Related Disorders. Henry Kimpton Publishers, London

Graham F K & Kendall B S (1960) Memory-for-Designs test: revised general manual. Perceptual and Motor Skills 11: 147-188

Grufferman S (1978) Alzheimer's disease and senile dementia: one disease or two? In: Katzman R, Terry R D, & Bick K L (Eds) Alzheimer's disease: Senile Dementia and Related Disorders. Aging Series, Vol 7. Raven Press, New York

Gustafson L & Hagberg B (1975) Dementia with onset in the presenile period: a cross-sectional study. Acta Psychiatrica Scandanavica Supplement 257

Gustafson L & Nilsson L (1982) Differential diagnosis of presenile dementia on clinical grounds. Acta Psychiatrica Scandanavica 65: 194-209

Gustafson L & Risberg J (1974) Regional cerebral blood flow related to psychiatric symptoms in dementia with onset in the presenile period. Acta Psychiatrica Scandanavica 50: 516-538

Hachinski V C, Lassen N A, & Marshall J (1974) Multi-infarct dementia: a cause of mental deterioration in the elderly. Lancet 2: 207-210

Hachinski V C, Iliff L D, Zilhka E, DuBoulay G H, McAllister V L, Marshall J, Russel R W R, & Symon L (1975) Cerebral blood flow in dementia. Archives of Neurology 32: 632-637

Hagberg B & Ingvar D H (1976) Cognitive reductions in presenile dementia related to regional abnormalities of the cerebral blood flow. British Journal of Psychiatry 127: 209-221

Halstead H (1943) A psychometric study of senility. Journal of Mental Science 89: 863-873

Hanley I & Hodge J (Eds) (1984) Psychological Approaches to the Care of the Elderly. Croom Helm, London

Harkins S W, Chapman C R, & Eisdorfer C (1979) Memory loss and response bias in senescence. Journal of Gerontology 34: 66-72

Harrison M J G, Thomas D J, Du Boulay G H, & Marshall J (1979) Multi-infarct dementia. Journal of Neurological Science 40: 97-103

Heaton R K & Pendleton M G (1981) Use of neuropsychological tests to predict adult patients' everyday functioning. Journal of Consulting and Clinical Psychology 49: 807-821

Hebb D O (1949) The Organisation of Behaviour. Wiley, New York

Hecaen H & Albert M L (1978) Human Neuropsychology. Wiley, New York

Henderson A S (1986) The epidemiology of Alzheimer's disease. British Medical Bulletin 42: 3-10

Heston L L, Mastri A R, Anderson V E, & White J (1981) Dementia of the Alzheimer type: clinical genetics, natural history, and associated conditions. Archives of General Psychiatry 38: 1085-1090

Hirst W (1982) The amnesic syndrome: descriptions and explanations. Psychological Bulletin 91: 435-460

Hochhaus L (1972) A table for the calculation of d' and B. Psychological Bulletin 77: 375-376

Horel J A (1978) The neuroanatomy of amnesia: a critique of the hippocampal memory hypothesis. Brain 101: 403-445

Hubbard B M & Anderson J M (1981) A quantitative study of cerebral atrophy in old age and senile dementia. Journal of the Neurological Sciences 50: 135-145

Hughes C P, Berg L, Danziger W L, Coben L A, & Martin R (1982) A new clinical scale for the staging of dementia. British Journal of Psychiatry 140: 566-572

Hull C L (1943) Principles of Behaviour: An Introduction to Behaviour Theory. Appleton Century Crofts, New York

Huppert F A & Piercy M F (1976) Recognition memory in amnesic patients: effects of temporal context and familiarity of material. Cortex 12: 3-20

Hutchison J M & Jensen M (1980) A pragmatic evaluation of discourse communication in normal and senile elderly in a nursing home. In: Obler L K & Albert M L (Eds) Language and Communication in the Elderly. Lexington Books, Lexington, Massachusetts

Isaacson R L (1975) Memory processes and the hippocampus. In: Deutsch D & Deutsch J A (Eds) Short Term Memory. Academic Press, New York

Iversen S D (1977) Temporal lobe amnesia. In: Whitty C W M & Zangwill O L (Eds) Amnesia, 2nd edition. Butterworths, London

Jellinger K (1976) Neuropathological aspects of dementias resulting from abnormal blood and cerebrospinal fluid dynamics. Acta Neurologica Belgica 76: 83-102

Katz S, Ford A B, Moskowitz R W, Jackson B A, & Jaffe M W (1963) Studies of illness in the aged. The index of ADL: a standardised measure of biological and psychosocial function. Journal of the American Medical Association 185: 914-919

Katzman R (1976) The prevalence and malignancy of Alzheimer disease. Archives of Neurology 33: 217-218

Kaszniak A W, Fox J, Gandell D L, Garron D C, Huckman M S, & Ramsey R G (1978) Predictors of mortality in presenile and senile dementia. Annals of Neurology 3: 246-252

Kaszniak A W, Garron D C, & Fox J (1979) Differential effects of age and cerebral atrophy upon span of immediate recall and paired associate learning in older patients suspected of dementia. Cortex 15: 285-295

Kay D W K (1986) The genetics of Alzheimer's disease. British Medical Bulletin 42: 19-23

Kendrick D C (1982) Why assess the aged? a clinical psychologist's view. British Journal of Clinical Psychology 21: 47-54

Kendrick D C, Gibson A J, & Moyes I C A (1979) The revised Kendrick Battery: clinical studies. British Journal of Social and Clinical Psychology 18: 329-340

Kendrick D C & Moyes I C A (1979) Activity, depression, medication and performance on the revised Kendrick Battery. British Journal of Social and Clinical Psychology 18: 341-350

Kertesz A & Hooper P (1982) Praxis and language: the extent and variety of apraxia in aphasia. Neuropsychologia 20: 275-286

Killian G A, Holzman P S, Davis J M, & Gibbons R (1984) Effects of psychotropic medication on selected cognitive and perceptual measures. Journal of Abnormal Psychology 93: 58-70

Kinsbourne M & Wood F (1975) Short term memory processes and the

amnesic syndrome. In: Deutsch D & Deutsch J A (Eds) Short Term Memory. Academic Press, London

Kirshner H S, Webb W G, & Kelly M P (1984) The naming disorder of dementia. Neuropsychologia 22: 23-30

Kolers P A (1975) Reading a year later. Journal of Experimental Psychology: Human Learning and Memory 2: 554-565

Kral V A (1962) Senescent forgetfulness: benign or malignant. Canadian Medical Association Journal 86: 257-260

Larner S (1977) Encoding in senile dementia and elderly depressives: a preliminary study. British Journal of Social and Clinical Psychology 16: 379-390

Larsson T, Sjogren T, & Jacobson G (1963) Senile dementia: A clinical, sociomedical and genetic study. Acta Psychiatrica Scandanavica 39: Suppl 67, 1-259

Lawson J S & Barker M G (1968) The assessment of nominal dysphasia in dementia: the use of reaction time measures. British Journal of Medical Psychology 41: 411-414

Levin H S & Peters B H (1982) Report of the ad hoc committee for the classification of the severity of Alzheimer's disease. In: Corkin S, Davis K L, Growdon J H, Usdin E, & Wurtman R J (Eds) Alzheimer's Disease: A Report of Progress in Research. Aging Series, Vol 19. Raven Press, New York

Lezak M D (1976) Neuropsychological Assessment. Oxford University Press

Lhermitte F & Signoret J-L (1972) Analyse neuropsychologique et differenciation des syndromes amnesiques. Revue Neurologique 126: 164-178

Lishman W A (1978) Organic Psychiatry: The Psychological Consequences of Cerebral Disorder. Blackwell, Oxford

Liston E H (1979) The clinical phenomenology of presenile dementia: a critical review of the literature. Journal of Nervous and Mental Diseases 167: 329-336

Ludlow C (1977) Recovery from aphasia: a foundation for treatment. In: Sullivan M A & Kommers M S (Eds) Rationale for Adult Aphasic Therapy. University of Nebraska Medical Center.

Luria A R (1973) The Working Brain. Penguin, London

MacDonald C (1969) Clinical heterogeneity in senile dementia. British Journal of Psychiatry 115: 267-271

McKhann G, Drachman D, Folstein M, Katzman R, Price D, & Stadlan E M (1984) Clinical diagnosis of Alzheimer's disease: report of the

NINCDS ADRDA Work Group under the auspices of Department of Health and Human Services task force on Alzheimer's disease. Neurology 34: 939-944

Maher B (1972) The language of schizophrenia: A review and interpretation. British Journal of Psychiatry 120: 3-17

Matthews W B & Miller H (1979) Diseases of the Nervous Sysytem. Blackwell, Oxford

Mattis S, Kovner R, Gartner J, & Goldmeier E (1981) Deficits in retrieval of category exemplars in alcoholic Korsakoff patients. Neuropsychologia 19: 357-363

Mayes A & Meudell P (1981) How similar is the effect of cueing in amnesics and in normal subjects following forgetting? Cortex 17: 113-124

Meudell P & Mayes A (1981) A similarity between weak normal memory and amnesia with two and eight choice word recognition: a signal detection analysis. Cortex 17: 19-30

Miller E (1971) On the nature of the memory disorder in presenile dementia. Neuropsychologia 9: 75-78

Miller E (1972) Efficiency of coding and the short-term memory defect in presenile dementia. Neuropsychologia 10: 133-136

Miller E (1974a) Psychomotor performance in presenile dementia. Psychological Medicine 4: 65-68

Miller E (1974b) Retrieval from long term memory in presenile dementia. Paper presented to British Psychological Society Annual Conference, Bangor: Quoted in Miller E (1977) Abnormal Ageing. Wiley, London

Miller E (1975) Impaired recall and the memory disturbance in presenile dementia. British Journal of Social and Clinical Psychology 14: 73-79

Miller E (1977) Abnormal Ageing. Wiley, London

Miller E (1978) Retrieval from long-term memory in presenile dementia: two tests of an hypothesis. British Journal of Social and Clinical Psychology 17: 143-148

Miller E (1980) Cognitive assessment of the older adult. In: Birren J E & Sloane R B (Eds) Handbook of Mental Health and Aging. Prentice-Hall, Englewood Cliffs, New Jersey

Miller E (1981a) The nature of the cognitive deficit in senile dementia. In: Miller N E & Cohen G D (Eds) Clinical Aspects of Alzheimer's Disease & Senile Dementia. Aging Series, Vol 15. Raven Press, New York

Miller E (1981b) The differential psychological evaluation. In: Miller N E & Cohen G D (Eds) Clinical Aspects of Alzheimer's Disease & Senile Dementia. Aging Series, Vol 15. Raven Press, New York

Miller E (1984a) Psychological aspects of dementia. In: Pearce J M S, Dementia: A Clinical Approach. Blackwell, London.

Miller E (1984b) Recovery and Management of Neuropsychological Impairments. Wiley, Chichester

Miller E (1984c) Verbal fluency as a function of a measure of verbal intelligence and in relation to different types of cerebral pathology. British Journal of Clinical Psychology 23: 53-57

Miller E & Hague F (1975) Some characteristics of verbal behaviour in presenile dementia. Psychological Medicine 5: 255-259

Miller E & Lewis P (1977) Recognition memory in elderly patients with depression and dementia: a signal detection analysis. Journal of Abnormal Psychology 86: 84-86

Miller W R (1975) Psychological deficit in depression. Psychological Bulletin 82: 238-260

Milner B (1968) Disorders of memory after brain lesions in man. Neuropsychologia 6: 175-179

Milner B (1970) Memory and the medial temporal regions of the brain. In: Pribram K H & Broadbent D E (Eds) Biology of Memory. Academic Press, New York

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

Moore V & Wyke M A (1984) Drawing disability in patients with senile dementia. Psychological Medicine 14: 97-105

Morris J C, Cole M, Banker B Q, & Wright D (1984) Hereditary dysphasic dementia and the Pick-Alzheimer spectrum. Annals of Neurology 16: 455-466

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

Morris R G (1984) Dementia and the function of the articulatory loop system. Cognitive Neuropsychology 1: 143-157

Moscovitch M (1982) A neuropsychological approach to perception and memory in normal and pathological ageing. In: Craik F I M & Trehub S (Eds) Aging and Cognitive Processes (Advances in the Study of Communication and Affect, Vol 8). Plenum Press, New York

Mountjoy C Q, Rossor M N, Iversen L L, & Roth M (1984) Correlation of cortical cholinergic and GABA deficits with quantitative neuropathological findings in senile dementia. Brain 107: 507-518

Naguib M & Levy R (1982) Prediction of outcome in senile dementia: a computed tomography study. British Journal of Psychiatry 140: 263-267

Neary D, Snowden J S, Bowen D M, Sims N R, Mann D M A, Benton J S, Northen B, Yates P O, & Davison A N (1986) Neuropsychological syndromes in presenile dementia due to cerebral atrophy. Journal of Neurology, Neurosurgery, and Psychiatry 49: 163-174

Nelson H E (1982) National Adult Reading Test (NART) Manual. NFER-Nelson, Windsor, Berkshire

Norman D A & Rumelhart D E (1970) A system for perception and memory. In: Norman D A (Ed) Models of Human Memory. Academic Press, New York

Nott R N & Fleminger J J (1975) Presenile dementia: the difficulties of early diagnosis. Acta Psychiatrica Scandanavica 51: 210-217

O'Keefe J & Nadel L (1978) The Hippocampus as a Cognitive Map. Oxford University Press

O'Neill P M & Calhoun K S (1975) Sensory deficits and behavioral deterioration in senescence. Journal of Abnormal Psychology 84: 579-582

Paivio A (1971) Imagery and Verbal Processes. Holt Rinehard & Winston, New York

Palermo D S & Jenkins J J (1964) Word-Association Norms: Grade School through College. University of Minnesota Press, Minneapolis

Papez J (1937) A proposed mechanism of emotion. Archives of Neurology 38: 725-743

Pattie A H & Gilleard C J (1979) Clifton Assessment Procedures for the Elderly. Hodder and Stoughton, Sevenoaks, Kent

Pearce J M S (1984) Dementia: A Clinical Approach. Blackwell, Oxford

Pearce J & Miller E (1973) Clinical Aspects of Dementia. Balliere Tindall, London

Perez F I, Rivera V M, Meyer J S, Gay J R A, Taylor R L, & Mathew N T (1975) Analysis of intellectual and cognitive performance in patients with multi-infarct dementia, vertebro-basilar insufficiency with dementia, and Alzheimer's disease. Journal of Neurology, Neurosurgery, and Psychiatry 38: 533-540

Perlmutter M & Mitchell D B (1982) The appearance and disappearance of age differences in adult memory. In: Craik F I M & Trehub S (Eds) Aging and Cognitive Processes (Advances in the Study of Communication and Affect, Vol. 8). Plenum Press, New York

Perry E K, Tomlinson B E, Blessed G, Bergmann K, Gibson P H, & Perry R H (1978) Correlation of cholinergic abnormalities with senile plaques and mental test scores. British Medical Journal 2: 1457-1459

Perry R H (1986) Recent advances in neuropathology. British Medical Bulletin 42: 34-41

Philadelphia Geriatric Centre Minimal Social Behaviour Scale (undated). Philadelphia Geriatric Centre

Porteus S D (1965) Porteus Maze Test: Fifty Years of Application. Pacific Books, Palo Alto, California

Pryse-Phillips W & Murray T J (1982) Essential Neurology (2nd Edition). Medical Examination Publishing Co, New York

Raven J C (1961) The Crichton Block Design Test of Spatial Orientation. Unpublished, Department of Psychological Research, Crichton Royal Hospital, Dumfries

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

Renvoize E B, Gaskell R K, & Klar H M (1985) Results of investigations in 150 demented patients consecutively admitted to a psychiatric hospital. British Journal of Psychiatry 147: 204-205

Richter J A, Perry E K, & Tomlinson B E (1980) Acetylcholine and choline levels in post-mortem human brain tissue: preliminary observations in Alzheimer's disease. Life Sciences 26: 1683-1689

Rochford G (1971) A study of naming errors in dysphasic and in demented patients. Neuropsychologia 9: 437-443

Rochford G & Williams M (1964) The measurement of language disorders. Speech Pathology and Therapy 7: 3-21

Rochford G & Williams M (1965) Studies in the development and breakdown of the use of names; Part IV: The effects of word frequency. Journal of Neurology, Neurosurgery, and Psychiatry 28: 407-413

Ron M A, Toone B K, Garralda M E, & Lishman W A (1979) Diagnostic accuracy in presenile dementia. British Journal of Psychiatry 134: 161-168

Rose F C & Symonds C P (1960) Persistent memory defect following encephalitis. Brain 83: 195-212

Rosen W G (1980) Verbal fluency in aging and dementia. Journal of Clinical Neuropsychology 2: 135-146

Rosen W G & Mohs R C (1982) Evolution of cognitive decline in dementia. In: Corkin S, Davis K L, Growdon J H, Usdin E, & Wurtman R J (Eds) Alzheimer's Disease: A Report of Progress in Research. Aging Series, Vol 19. Raven Press, New York

Rosen W G, Terry R D, Fuld P A, Katzman D, & Peck A (1980) Pathological verification of ischemic score in differentiation of dementias. Annals of Neurology 7: 486-488

Rossor M N, Iversen L L, Johnson A J, Mountjoy C Q, & Roth M (1981) Cholinergic deficit in frontal cerebral cortex in Alzheimer's disease is age dependent. Lancet 2: 1422

Rossor M N, Garrett N J, Johnson A J, Mountjoy C Q, Roth M, & Iversen L L (1982) A post-mortem study of the cholinergic and GABA systems in senile dementia. Brain 105: 313-330

Rossor M N, Iversen L L, Reynolds G P, Mountjoy C Q, & Roth M (1984) Neurochemical characteristics of early and late onset types of Alzheimer's disease. British Medical Journal 288: 961-964

Rossor M & Iversen L L (1986) Non-cholinergic neurotransmitter abnormalities in Alzheimer's disease. British Medical Bulletin 42: 70-74

Roth M (1971) Classification and aetiology in mental disorders of old age: some recent developments. In: Kay D W K & Walk A (Eds) Recent Developments in Psychogeriatrics. British Journal of Psychiatry Special Publication No. 6. Headley, Kent

Roth M (1986) The association of clinical and neurological findings and its bearing on the classification and aetiology of Alzheimer's disease. British Medical Bulletin 42: 42-50

Russell E W (1975) A multiple scoring method for the assessment of complex memory functions. Journal of Consulting and Clinical Psychology 43: 800-809

Savage R D, Britton P G, Bolton N, and Hall E H (1973) Intellectual Functioning in the Aged. Methuen, London

Schaie K W (1967) Age changes and age differences. Gerontologist 7: 128-132

Schaie K W & Strother C R (1968) A cross-sequential study of age changes in cognitive behaviour. Psychological Bulletin 70: 671-680

Schneck M K, Reisberg B, & Ferris S H (1982) An overview of current concepts of Alzheimer's disease. American Journal of Psychiatry 139: 165-173

Schonell F (1942) Backwardness in the Basic Subjects. Oliver &

Boyd, London

Schuell H (1965) Differential Diagnosis of Aphasia with the Minnesota Test. University of Minnesota Press, Minneapolis

Seltzer B & Sherwin I (1983) A comparison of clinical features in early- and late-onset primary degenerative dementia: one entity or two? Archives of Neurology 40: 143-146

Semple S A, Smith C M, & Swash M (1982) The Alzheimer disease syndrome. In: Corkin S, Davis K L, Growdon J H, Usdin E, & Wurtman R J (Eds) Alzheimer's Disease: A Report of Progress in Research. Aging Series, Vol 19. Raven Press, New York

Shallice T (1982) Specific impairments of planning. Philosophical Transactions of the Royal Society of London 298: 199-209

Siegler I C (1980) The psychology of adult development and ageing. In: Busse E W & Blazer D G (Eds) Handbook of Geriatric Psychiatry. Van Nostrand Reinhold, New York

Siegler I C & Botwinick J (1979) A long term longitudinal study of the intellectual ability of older adults - the matter of selective subject attrition. Journal of Gerontology 34: 242-248

Simpson J A (1969) The clinical neurology of temporal lobe disorders. In: Herrington R N (Ed) Current Problems in Neuropsychiatry: Schizophrenia, Epilepsy, the Temporal Lobe. British Journal of Psychiatry Special Publication No. 4: 42-48

Skelton-Robinson M & Jones S (1984) Nominal dysphasia and the severity of senile dementia. British Journal of Psychiatry 145: 168-171

Smith J S & Kiloh L G (1981) The investigation of dementia: results in 200 consecutive admissions. Lancet 1: 824-827

Spellacy F J & Spreen O (1969) A short form of the Token Test. Cortex 5: 390:397

Squire L (1980) Specifying the defect in human amnesia: storage, retrieval and semantics. Neuropsychologia 18: 368-372

Sweet W H, Talland G A, & Ervin F R (1959) Loss of recent memory following section of the fornix. Transactions of the American Neurological Association 84: 76-82

Symonds R L (1981) Dementia as an experience. Nursing Times, 30 September: 1708-1710

Talland G A (1965) Deranged Memory: A Psychonomic Study of the Amnesic Syndrome. Academic Press, New York

Taylor R (1981) Memory defects in Korsakoff's syndrome. Unpublished M. Phil. Thesis, University of Edinburgh.

Terman L M & Merrill M A (1937) Measuring Intelligence. Harrap & Co, London

Terry R D (1978) Aging, senile dementia, and Alzheimer's disease. In: Katzman R, Terry R D, & Bick K L (Eds) Alzheimer's disease: Senile Dementia and Related Disorders. Aging Series, Vol 7. Raven Press, New York

Thorndike E L & Lorge I (1944) The Teacher's Word Book of 30,000 Words. Teacher's College Press, Columbia University, New York

Tomlinson B E, Blessed G, & Roth M (1970) Observations on the brains of demented old people. Journal of the Neurological Sciences 11: 205-242

Trapp G A, Miner G D, Zimmerman R L, Mastri A R, & Heston L L (1978) Aluminium levels in brain in Alzheimer's disease. Biological Psychiatry 13: 709-718

Tulving E (1972) Episodic and semantic memory. IN: Tulving E & Donaldson W (Eds) Organisation of Memory. Academic Press, New York

Tulving E & Thomson D M (1971) Retrieval processes in recognition memory: effects of associative context. Journal of Experimental Psychology 87: 116-124

Underwood B J (1957) Interference and forgetting. Psychological Review 64: 49-60

United States Employment Service (1965) General Aptitude Test Battery. United States Department of Labor, Washington DC

Van der Kam P, Mol F, & Wimmers M F H G (1971) Beoordelingschael voor ouden patienten (BOP). Van Loghum Slaterus, Deventer, Netherlands (transl. C Smidt)

Vygotsky L S (1978) Mind in Society: The development of Higher Psychological Processes. Harvard University Press

Warren E W & Groome D H (1983) The effect of memory impairment on verbal recognition scores and the marking of acoustic and semantic distractors. British Journal of Clinical Psychology 22: 215-216

Warrington E K (1986) Memory for facts and memory for events. British Journal of Clinical Psychology 25: 1-12

Warrington E K & Weiskrantz L (1968) A study of learning and retention in amnesic patients. Neuropsychologia 6: 283-292

Warrington E & Weiskrantz L (1970) Amnesic syndrome: consolidation or retrieval? Nature 228: 628-630

Wechsler D (1945) A standardised memory scale for clinical use. Journal of Psychology 19: 87-95

Wechsler D (1955) The Wechsler Adult Intelligence Scale Manual. Psychological Corporation, New York

Weigl E (1941) On the psychology of the so-called process of abstraction. Journal of Abnormal and Social Psychology 36: 3-33

Weingartner H, Kaye W, Smallberg S, Ebert M H, Gillin J C, & Sitaram N (1981) Memory failures in progressive idiopathic dementia. Journal of Abnormal Psychology 90: 187-196

Weingartner H, Kaye W, Smallberg S, Cohen R, Ebert M H, Gillin J C, & Gold P (1982) Determinants of memory failures in dementia. In: Corkin S, Davis K L, Growdon J H, Usdin E, & Wurtman R J (Eds) Alzheimer's Disease: A Report of Progress in Research. Aging Series, Vol 19. Raven Press, New York

Wells C E (1982) Refinements in the diagnosis of dementia. American Journal of Psychiatry 139: 621-622

Wertheimer J (1984) Evolutive variations in Alzheimer's disease. In: Proceedings of the Alpenlandisches Psychiatrie-Symposium, Seefeld-Tirol, 1st & 2nd Oct, 1982: 247-263. Schattauer Verlag, Stuttgart

Whitehead A (1973a) The pattern of WAIS performance in elderly psychiatric patients. British Journal of Social and Clinical Psychology 12: 435-436

Whitehead A (1973b) Verbal learning and memory in elderly depressives. British Journal of Psychiatry 123: 203-208

Whitehead A (1974) Factors in the learning deficit of elderly depressives. British Journal of Social and Clinical Psychology 13: 201-208

Whitehead A (1975) Recognition memory in dementia. British Journal of Social and Clinical Psychology 14: 191-194

Whitehead A (1977) Changes in cognitive functioning in elderly psychiatric patients. British Journal of Psychiatry 130: 605-608

Whitehead A (1982) Outcome in elderly psychiatric patients: implications of cognitive changes. British Journal of Clinical Psychology 21: 225-226

Wickelgren W A (1970) Multitrace strength theory. In: Norman D A (Ed) Models of Human Memory. Academic Press, New York

Wickelgren W A (1979) Chunking and consolidation: a theoretical synthesis of semantic networks, configuring in conditioning, S-R versus cognitive learning, normal forgetting, the amnesic syndrome, and the hippocampal arousal system. Psychological Review 86: 44-60

Wilson R S, Kaszniak A W, & Fox J H (1981) Remote memory in senile dementia. Cortex 17: 41-48

Wilson R S, Kaszniak A W, Bacon L D, Fox J H, & Kelly M P (1982) Facial recognition memory in dementia. Cortex 18: 329-336

Wilson R S, Bacon L D, Fox J H, & Kaszniak A W (1983) Primary memory and secondary memory in dementia of the Alzheimer type. Journal of Clinical Neuropsychology 5: 97-104

Winer B (1971) Statistical Principles in Experimental Design. McGraw-Hill, New York

Winocur G & Kinsbourne M (1978) Contextual cueing as an aid to Korsakoff amnesics. Neuropsychologia 16: 671-682

Winocur G & Weiskrantz L (1976) An investigation of paired associate learning in amnesic patients. Neuropsychologia 14: 97-110

Woods R T & Britton P G (1985) Clinical Psychology with the Elderly. Croom Helm, London

Woods R T & Piercy M (1974) A similarity between amnesic memory and normal forgetting. Neuropsychologia 12: 437-445

APPENDIX 1

Details of the test procedures.

Manual

The following data on subjects' characteristics is recorded:

Name; Sex; Date of birth; Age; Diagnosis; Score on Hachinski index; Relevant details of history of condition, including estimated duration; Any known sensory or motor impairments;

Marital status; Previous occupation; Social class; Educational history; Any psychiatric history, and previous contact with hospitals/homes; Any significant medical history, including any current or previous illnesses; Any history of alcohol abuse; Any history of ECT;

Location; Date of test; Time of test; Details of current and recent medication;

Relevant registration numbers; Any home address; G.P.'s address;

Any other relevant information, history, or results of special investigations (such as CT scan, EEG, other laboratory tests).

"Hello Mr/Mrs_. My name's _ and I'm a psychologist working around here"

Note: a Any noticeable response to my presence or words
b If S makes eye contact
c Any verbal R
d Approp. vbl R (Hello, Oh yes, etc)

....."and I was wondering if I could have a chat with you about how you're getting on"....

Note: e Any R
f Eye contact
g Vbl R
h Approp vbl R

...."How are you today?"

Note: i Vbl R
j App vbl R (Fine, Not too good, etc)

Explain the purpose of approaching S in a chatty way including: How's you're memory and concentration? As good as it used to be? (What do you have most difficulty with? if anything)....because I'm doing some research on how older people find their memory, concentration, & so on....I'm not singling you out/picking on you, I'm asking lots of people, anyone who'll agree to talk to me....This is research, you'd be helping me out, you don't have to talk to me; there's no obligation and it's not part of treatment. General chat about early life etc. Could I ask you to do some things to see how you're getting on? with explanation of the kind of things involved. (If S hesitant also explain just to see how you are today, even if you're not at your best.) (Do you have good days and bad days?)

(Modified as appropriate for repeat sessions ("Perhaps you'll remember I'm...etc) and for other subject groups.

Make sure S has as few distractions as possible and a clear, undazzled view of all visual or performance items. Note starting time. In all tests, encourage S to have a guess or have a go. Unless otherwise stated, instructions can be repeated, and appropriate gesturing and encouragement used, to ensure that S understands what is required.

ORIENTATION Note answers verbatim.

1.	What is your full name?	First name and surname
2.	How old are you?	Present or '79 next August'
3.	What is your date of birth?	Date,month,& year all req'd
4.	What is this place/Where are you now?	Name of ward or recog. that it is a ward; recog. that in own home, LA accommodation, or whatever
5.	What is the name of this hospital/ What is the address of this place?	Name of hosp, street address of LA home, no. & street of own house as appropriate
6.	What is the name of this town/ city (that you're in now etc)?	
7.	Who is Prime Minister (just now)?	Surname sufficient
8.	Who was Prime Minister before him/her/the present P.M?	" "
9.	Who is President of the United States of America (just now)?	" "
10.	What are the colours of the British flag/Union Jack?	Red, White, & Blue
11.	What day is it (today)?	Date not req'd (this can be explained once if date given)
12.	What month is it (now)?	
13.	What year is it (now)?	

SCORE: 0 No verbal R or inappropriate verbal R or appropriate but wrong verbal R
1 Correct

Total = 00 - 13

AUTOMATIC SPEECH

1. "Could you count up from 1 to 20 for me, as quickly as you can without making mistakes." (The prompt 1,2,3 may be given) Record: Time taken and no. of errors. (Spontaneous corrections not counted as errors; each omission or inversion = 1 error)

SCORE: 0 Not attempted
1 11 errors or more
2 5-10 errors
3 2-4 errors
4 1 error in 11" or more
5 1 error in 10" or less
6 0 errors in 11" or more

7 0 errors in 7-10"
8 0 errors in 6" or less

Total = 0 - 8

2. "Now can you say the alphabet - again as quickly as you can without making mistakes." (The prompt A,B,C may be given) Record: Time, errors (as above)

SCORE as Part 1

Total = 0 - 8

DIGIT SPAN

"I am going to say some numbers. Listen carefully, and when I have finished say them right after me." Say evenly at 1 per sec; drop pitch on last of each series. If S gets trial 1, move on to higher series. Trial 2 only given if trial 1 failed. (& two-digit series only given if both three-digit failed.) Record wrong answers verbatim. Stop after failure on both trials of a given length.

Series	trial 1	trial 2
(3)	5-8-2	6-9-4
(4)	6-4-3-9	7-2-8-6
(5)	4-2-7-3-1	7-5-8-3-6
(6)	6-1-9-4-7-3	3-9-2-4-8-7
(7)	5-9-1-7-4-2-8	4-1-7-9-3-8-6
(2)	2-7	1-3

SCORE is the longest sequence of digits correctly reproduced, plus .4 if achieved on the first trial.

Total = 0.0 - 7.4

VISION/VISUAL FIELD/HEARING

Vision (Show appropriate cards):
0 No impairment: can read small characters (distinguish small dots etc)
1 Some " " " large " " large " "
2 Marked " can't " " " " " " "

Hearing (based on conversation, information in notes, etc):
0 No apparent impairment
1 Some " " : somewhat hard of hearing
2 Markedly deaf
Visual Fields:
"Please look straight at my nose and point to which of my fingers you see wiggling."

Random order:
Mid-R, Mid-L, Both
Upper-R, Upper-L, Both
Lower-R, Upper-L, Both

BOX-FILLING

Place grid in front of S, holding with fingertips: "Look at these

boxes. You'll see that in the first three boxes there's a mark like a number one. All I'd like you to do is go along this first line here (point) and put a mark like these ones (point) in each box. Go as quickly as you can and don't worry about being neat, and stop here (point)." If S doesn't do it correctly, explain once more. If OK: "Right, that's the idea. Now I'd like you to do the same thing in the rest of the boxes starting here (point). When you get to the end of one line go straight on to the next until I say stop. Go as fast as you can and don't bother about neatness. OK? Go." "Keep going" if S pauses. Timing starts as soon as S puts pencil to paper. Go for 30". Note no. done (excluding practice) at 15".

SCORE is no. of adequately (loose criterion) marked boxes. If S completes grid, note time taken.

Total = 00 - 90

---

* Decide whether full testing, or only short, is feasible. If only short, procede with Supplementary Dysphasia, Memory for Sentences, Write name, Read sentence, Gibson Spiral Mazes, Porteus Mazes, Square with Sticks.

TOKEN TEST

Lay out (tester's view) thus:

W	B	R	G	Y	(small squares)
G	R	Y	B	W	(small circles)
Y	G	W	R	B	(Large squares)
G	W	Y	B	R	(Large circles)

"As you can see, there are 20 pieces here. Some of them are squares (touch lines of squares) while others are circles (indicate). Some are large (indic.), others are small (indic.). There are red ones (point to an eg as each colour is said), black, green, yellow, and white ones. Now, I'm going to ask you to touch one of these pieces: touch a circle." (If S asks which one: "Any one, just touch any circle.") Commands appear below. They are uttered distinctly and without any special prosodic emphasis (except item 34 "NO" followed by brief pause.....the white square). Spontaneous corrections accepted. If S says has forgotten instructions: "just do what you can remember".

Parts 1 to 5: If no R started in 5" (or wrong R), say command again (after "let's try that again" if wrong, returning tokens to place if S has moved them). Stop after 5 zero scores.

SCORE 1 Correct 1st time
.5 " 2nd "
0 Fail

Part 6: ALL items are given if S is eligible to start this part.
NO repetitions.

SCORE 1 Correct
0 Fail

[Errors can be recorded as follows:
DNA did not attempt
G,W,Y,B,R, to denote colours
L,S, to denote size
, , to denote shape

Hence LG ,SR means touched large green circle then small red circle. Include words like "on" where appropriate.]

Commands:

Part 1. All 20 tokens as in figure
1. Touch a circle
2. Touch a square
3. Touch a yellow piece
4. Touch a red one
5. Touch a black one
6. Touch a green one
7. Touch a white one

Part 2. The small tokens are removed (covered)
8. Touch the yellow square
9. Touch the black circle
10. Touch the green circle
11. Touch the white square

Part 3. The small tokens are replaced
12. Touch the small white circle
13. Touch the large yellow square
14. Touch the large green square
15. Touch the small black circle

Part 4. The small tokens are removed
16. Touch the red circle and the green square
17. Touch the yellow square and the black square
18. Touch the white square and the green circle
19. Touch the white circle and the red circle

Part 5. The small tokens are replaced
20. Touch the large white circle and the small green square
21. Touch the small black circle and the large yellow square
22. Touch the large green square and the large red square
23. Touch the large white square and the small green circle

Part 6. The small tokens are removed
24. Put the red circle on the green square
25. Touch the black circle with the red square
26. Touch the black circle and the red square
27. Touch the black circle or the red square
28. Put the green square away from the yellow square
29. If there is a blue circle, touch a red square
30. Put the green square next to the red circle
31. Touch the squares slowly and the circles quickly
32. Put the red circle between the yellow square and the green square
33. Touch all the circles, except the green one

34. Touch the red circle-no-the white square
35. Instead of the white square, touch the yellow circle
36. In addition to touching the yellow circle, touch the black circle

Total = 00.0 - 36.0

[If S fails very early, do
COLOUR MATCHING TASK

Use all 5 colours, large tokens only. Laid out as before. "Which is the same colour as this" pointing to each of the circles in turn.

and

SUPPLEMENTARY DYSPHASIA TEST

Lay out a Pencil, Key, Spoon in that order before S. (ie S K P to E). Ask S to name each object or, if unable to do so, to say or show how the object is used/what it is used for.

SCORE: 2 Named the object
1 Knew what it was but could not name it
0 Neither

Total 0 - 6

1. Touch the pencil
2. Pick up the key
3. Give me the spoon
4. Knock on the table
**5. Raise your left hand
6. Touch your nose
7. Touch the pencil then the key
8. Turn over the key
9. Put the key above the spoon
10. Touch the key after you've touched the spoon
**11. Touch your left ear with your right hand
12. Put the key on the other side of the pencil, then put it back

Give ALL Ss Questions 5 & 11.

Commands can be repeated once if S asks or makes no R in 5" (after S has been encouraged to try it at 1st hearing)

SCORE: 1 Correct 1st time
.5 " 2nd "
0 Fail (which includes getting half a long command right)

Total = 00.0 - 12.0]

WRITING & READING

"Could you write your name here please."
"And now could you write this: The door is open."
"And now the number 33." ('In figures/Just the actual numbers' if S starts longhand)
"And now the number 66."
Can repeat as often as S wishes.

"Now could you read out the sentence on this card.....and then write it down." ('The quick brown fox jumps over the lazy dog') Can read it for S if necessary, but note verbatim Ss attempt at reading it.

SCORE:

Name Count no. of errors: ie omitted, substituted, added, or completely illegible characters - including omitted spaces. (Count more than 9 errors as 9.) Score is 9 minus no. of errors.

Total = 0 - 9

The door is open + 33 + 66. Errors scored as above and as follows:
17-20: 0 15,16: 1 13,14: 2 11,12: 3 9,10: 4 7,8: 5
5,6: 6 3,4: 7 1,2: 8 0: 9

Total = 0 - 9

Read sentence Count no. of errors (omitted, added, substituted words etc) Score is 9 minus no. of errors.

Total = 0 - 9

Write sentence No. of errors as with writing above.
32+: 0 28-31: 1 24-27: 2 20-23: 3 16-19: 4 12-15: 5
8-11: 6 4-7: 7 1-3: 8 0: 9

Total = 0 - 9

Note: R L , Preferred Non-pref. hand.

ESTIMATED PREMORBID I.Q. (N.A.R.T. / WAIS Vocab)

Use Nelson N.A.R.T. unless S failed to read the 'Quick brown...' sentence or cannot read the introductory Burt words (card 1), in which case try WAIS Vocab (every 2nd word), noting reasons for choice. "I'd like you to read slowly down the words on these cards. Some of them are unusual so don't worry if you haven't seen some of them before: just have a go at reading each one." If S says 'Don't know', encourage to guess. Stop after 14 out of 15 wrong. If S gets less than 10, use Schonell: note no. of errors in 2nd 50. Go right back till 10/10 in one block.

[WAIS instructions, if req'd: Place list before S if this might help. "I want you to tell me the meanings of some words. Let us start with_; what does_ mean?" Say word very clearly, repeating as necessary. Point to word if list is shown, but remember that this may be a useless distraction. Stop after 4 consecutive failures (Zero Rs). Occasionally it is difficult to determine whether a S does or does not know the meaning of a word. E may then say "Tell me more about it" or "Explain more fully" or make some other equally neutral statement. Record answers verbatim.]

SCORE:

NART and Schonell: direct use of NART tables.

WAIS: Score the (up to) 20 words as in WAIS (ie 0-40)
Double it (taking into account the cut-off procedure)
Consult WAIS tables for the age-corrected scaled score, s, then: epIQ = 100-((10-s) X 5)

(eg 75 yr old's raw score of 47:- Scaled Score 12:- epIQ 110)

SENTENCE REPETITION

"I'm going to say a sentence now and I'd like you to say it back/ repeat it straight after me." Each sentence is read once, clearly and fairly slowly. Errors include omissions, substitutions, additions, changes in words or order of words, but not contractions eg 'We're' in 1.

SCORE: 1 for every correct word in order. 1 off for added words in otherwise correct sentences.

1. We are going to buy some sweets for Michael. (Words: 9)
2. Jack likes to feed the little puppies in the shed. (10)
3. At the summer camp the children get up early in the morning to go swimming. (15)
4. Yesterday we went for a ride in our car along the road that crosses the bridge. (16)
5. The aeroplane made a careful landing in the space which had been prepared for it. (15)
6. Tom Brown's dog ran quickly down the road with a huge bone in its mouth. (15)

Total = 00 - 80

SENTENCE PRODUCTION

"Now I am going to give you a word, and I want you to use it in a sentence. For example, if I say 'house', you could say 'We are building a house'. If I say 'old' you could say 'This book is old'. Give me a sentence using 'car'." If S does not give a complete sentence: "No I want a complete sentence. You would have to say (correct the sentence given). Try again: give me a sentence using 'coat'" and continue with the test words. Transcribe the responses. Score as an error any sentence that is unintelligible, incomplete, or incorrect in structure. Do not count articulatory errors, colloquialisms, or usage consistent with the educational level of the patient (sic) as errors. (Can write the word down or spell it out if S is having any trouble catching it, eg using 'won't' for 'want', 'knew' for 'new', etc.)

1. Coat
2. New
3. Want
4. Have
5. After
6. Belongs

SCORE 1 for each correct sentence.

Total 0 - 6

Also SCORE mean letter-length and mean word length of correct sentences.

Total 00 - nn, 00 - nn

PARAGRAPH RECALL

"I'm going to read you a paragraph from an old newspaper. I'd like you to listen carefully and try to remember as much as you can of it because when I'm finished I'll ask you to tell me as much as you can about it. OK? ... Here we go." Read clearly and fairly slowly.

"/ In a city / in India, / five thousand / schoolchildren / paraded in / the main / square / to celebrate / the 70th / birthday of / the Prime Minister. / While reviewing / the parade / the Prime Minister released / a number of / white doves, / the symbols / of peace. / The doves flew / over the heads of / the young marchers. / One of them, / however, / perched on top of / the Prime Minister's / head / while he took the salute. /

"Could you tell me as much as you can remember of that paragraph." If no spontaneous attempt prompt with "It was about someone's birthday celebrations." Record all answers verbatim. When S has finished (ie after 'right; anything else?'-'no') ask ALL the following questions (saying 'You got that,_[Correct answer]' if S already has).

1. Which country was the city in?
2. How many schoolchildren were there?
3. Whose birthday was it?
4. How old was he?
5. What did the Prime Minister release?
6. What did one of them perch on?

Tell S if s/he is correct, and give the correct answer if S doesn't know.

SCORE:

Spontaneous recall: 1 for each idea (in slashes) produced.

Total 00 - 27

To questions:
0 No verbal R or innapropriate verbal R or wrong R
1 Correct to question
2 Correct already, ie spontaneously

Total 00 - 12

(Delayed recall scored identically)

ARITHMETIC

Don't mention the word arithmetic. "Now let's try these." Start at no. 3. Give 2 if 3 is failed, then 1 if 2 is failed.

1. Lay out 7 sticks thus: lll llll and ask 'How many sticks are

there altogether?'
2. If you have 3 books and give one away, how many do you have left?
3. How much is four pounds and (plus) five pounds?
4. If a man spends six pence on stamps and he gives the clerk ten pence, how much change should he get back?
5. A man was paid 25 pounds by each of six customers. What is the total amount he was paid?
6. How many inches are there in two and a half feet?
7. How many chairs can you buy for 36 pounds if one chair costs 6 pounds?
8. How many hours will it take a man to walk 24 miles at the rate of three miles an hour?

Stop after 4 consecutive failures.

SCORE 1 for each correct answer.

Total = 0 - 8

Additional questions for all Ss:

9. What's 27+8 ?	(If fail, 3+4)	('plus','and','add on'.
10. What's 31-7 ?	(8-5)	'minus','take away','less'.
11. What's 7*8 ?	(3*3)	'times','multiplied by'.
12. What's 28/4 ?	(8/2)	'divided by','4 into 28'.)

SCORE 1 for each correct answer (or assumed correct answer if the more difficult one of a pair is answered correctly).

Total = 0 - 8

YES-NO PICTURE MEMORY

"I'm going to show you ten pictures now and I'd like you to look carefully at them and try to remember them; because afterwards I'll show you the same pictures mixed up with another ten and ask you which ones you've seen before." Show the 10 pictures for 4" each. "Right, that's the 10 pictures. Now this pile has the same 10 mixed up with another 10. We'll just go through the pile and I'd like you to tell me whether you've seen each one before or not. Here's the first one. Did I show you it before or not?" For each picture, if S says don't know or doesn't answer in 5", say "Just have a guess, yes or no." Repeat the requirement as appropriate to make sure S always knows what the task is supposed to be.

SCORE:

Hits = No. of correct hits, 0-10
False alarms = No. of false alarms, 0-10

Raw score = Hits minus False alarms

Total = 00 - 10

See table for estimated d' (-3.50 to 3.50)
and B (0.22 to 4.63)

FACE-NAME LEARNING

('How d'you get on remembering people's names etc.') "I'm going to show you 4 pictures of people's faces and I'll tell you their names. I'd like you to try to remember each person's name because afterwards I'll show you the pictures again and ask you what each person is called." Each card is shown for 5" and the name clearly stated, "This is _____... ...(repeated once)." After the first presentation of the four faces, a test. "What's this person's name?" If S doesn't know, doesn't start to respond in 5", or is wrong, give first name and ask for second. Always finish with a clear statement of the name while S is looking at the picture for a few secs, reminding S as appropriate to try to remember because s/he'll be asked again. After the first test, re-order the faces and test again.

1 (younger male) GEORGE WILSON
2 (older female) FIONA MURRAY
3 (old male) ALEC MCLEOD
4 (young girl) ANNIE DUNCAN

Order of presentation:

Initial: 1 2 3 4

1st test: 3 1 4 2
2nd test: 2 4 1 3
3rd test: 4 2 3 1

Delayed test: 1 2 3 4

(Record R verbatim)

SCORE: 0 Nothing correct
1 2nd name after prompt
2 One name correct spontaneously
3 Both names correct

Total = 00 - 48 (00 - 12 for each test)

[If S scores very poorly, do
FACE MATCHING TASK

Show the cards in a row with duplicate set below. Point to each of the upper line in turn: "Which is the same as this."]

SPIRAL MAZES

1. Place 1st (unspotted) maze right way up in front of S: "This is like a maze - you start with your pencil at the little arrow here (point) and you go round and round. Try to keep on the white track and not touch the thick black lines. Try to go as quickly as you can, until you come out at the end here (point)." E holds maze steady. A bluntish pencil is provided (to avoid tearing), built up if necessary. A maximum of three prompts is allowed, including only one repetition of the instruction to avoid touching the black

lines. If S stops for over 2", 'keep going' is said.

2. If S has attempted 1st maze, place 2nd (spotted) one right way up before S: "This is the same idea but this time there are obstacles in the maze - these little circles (point), which you've to miss. So again start at the little arrow here (point) and go round and round. Again try to keep on the white track and not touch the thick black lines or these small black circles (point) - go past them without touching them. Try to go as quickly as you can until you come out at the end again here (point)." (Prompts etc as above if necessary.)

In both cases, timing begins as soon as S puts the pencil on the paper to start.

SCORE:
(a) Time in seconds to complete. (Limit 3.5 min. If close to finishing, let S continue for the sake of morale.)
Total 00 - nn

(b) Errors: 1 when circle or line is touched (not penetrated) by pencil.
2 " " " " " penetrated.
2 points are scored for each inch of continued contact/ penetration if pencil remains on black line for long.

Total = 00 - nn

A maze is scored as not completed (N/C) if S completes at least 1st circle of it but gives up before finishing, and after 3 prompts (or of course if runs out of time). It is scored not attempted (N/A) if S fails to complete even 1 circle of it after 3 prompts.

PORTEUS MAZES
(Years 5,6)

E holds top of paper with fingertips so it doesn't shift, with other mazes covered by folding the paper. "This is another kind of maze. You start with the pencil here (point) and go through the maze until you come out here (point). Try not to enter any 'blind alleys', dead ends, and don't cross any of the walls of the maze - these lines here (point). Try not to lift your pencil from the paper until you've come out the end of the maze. Right, start now." Timing begins immediately assuming S understands. Directions can be repeated once if S is making no start. A maximum of three prompts (eg 'keep going' if S hesitates) is allowed. Note time to complete.

Present second maze in a similar way.

SCORE:

2 Pts for each third of a maze reached (see sample mazes); ie 6 for finishing a maze, minus points for each of the following errors:

1 pt Any blind alley entrance
1 pt A touched wall/line (or just crossed & back, realising the error (be lenient in case of poor motor control)

2 pts Right through a wall/line and continues

Plus time bonuses for each maze, only if the maze gets score of 5 or 6:

3 pts for 15" or less
1 pt for 16" to 25"

Total 00 - 18

DRAWING SHAPES

"Please draw a circle."
' " " " square.'
' " " " triangle'
If S fails, ask S to copy the appropriate card (noting whether each drawing is spontaneous or copy, and noting which is the bottom of the page).

SCORE:
(see sample drawings for help)

Circle, square, triangle, all:
0 Not a circle/square/triangle
1 Recognisable
2 Good

Total 0 - 6 for spontaneous drawing,
0 - 6 for copies.

MEMORY FOR DESIGNS and COPYING DESIGNS

"I'm going to show you a(nother) design now, and (this time) I'd like you to draw it from memory after I've taken it away. I'll show you it for about 10 seconds so have a good look at it so that you can draw it afterwards." At card 4, point out that "There are two parts here, so have a good look at both." If S omits 2nd part completely, prompt with "Can you remember anything of the second part?" Give all 4 design-memory cards (unless S is doing so badly that this is pointless). Cover previous attempts to reduce 'copy-perseveration'.

[If S makes no attempt, can show multi-choice cards and ask 'Which one of these was it?']

Then go through the 4 cards again and ask S to copy each one. Record which are memorial attempts, which copies, and note which is the bottom of the page.

SCORE:
(See scoring criteria and sample drawings)

Diamond	0 - 3
Star	0 - 3
Flags	0 - 3
Double	0-3,0-3

Total 00 - 15 for memorial attempts,
      00 - 15 for copies.

SQUARE WITH STICKS

Give S 4 sticks, and ask S to make a square from them. If fails, make a square with the other 4 and ask S to copy.

SCORE:

4 Spontaneous square in 0-8"
3 " " " 9+"
2 Copy square in 0-8"
1 " " " 9+"
0 Couldn't

Total = 0 - 4

MULTIPLE-CHOICE BLOCK DESIGN

"Look at this design on the left here (point): could you point to the one of these three (indicate) which is the same as it, and the same way round?" Instructions can be repeated, explained as necessary if makes no attempt or doesn't understand.

SCORE: 1 Correct
0 Rotated design
0 Different design

Total = 0 - 3

BLOCK DESIGN

"You see these blocks. Some are black, some yellow, and some half and half (indicating each type). I'm going to put these 4 together to make a design like the one on the card here (point). Watch me." Make the design, scramble 4 the blocks, and say "Now could you make one the same." If S fails (or runs out of time), say "Watch me again" and repeat the demonstration and request. Pass or fail, proceed to design 2. Proceed exactly as above. (Can start with 2 in a fairly capable S, and use 1 only if 2 is failed.) Pass or fail (unless S hasn't touched blocks or is obviously quite incapable of doing this task, in which case stop) proceed to design 3 - For design 3 and onwards no demonstrations are given: the card is shown with the instruction "Now make one like this." No second attempts. NB In every case provide S with only the blocks required for the design in question. Stop after 3 consecutive designs (including 1,2) are failed.

Time limits: All designs but last 60"
Last 120"
For successes, note Time.
For failures, note no. of correct joins at end of time.

SCORE:

1 & 2 1st attempt 4, 2nd 2.
3 to 6: 1-10" 7, 11-15" 6, 16-20" 5, 21"+ 4.
(7 not scored, used for clinical interest and only if 6 completed easily)

Total = 00 - 36

YERKES' TEST

"This is a drawing of a pile of blocks. How many blocks are used in making this pile?" Show appropriate cards. Stop after 2 consecutive failures.

SCORE: 1 for each correct answer.

Total = 0 - 3

TIME-TELLING

"Can you tell me what time this clock is showing?" All 3 cards.

SCORE: 1 for each correct answer.

Total = 0 - 3

PARAGRAPH RECALL (Delayed)

"Do you remember a while ago I read you a paragraph out of an old newspaper. Can you remember what it was about?" If no spont. attempts say "It was about someone's birthday celebrations." Score spontaneous attempts as before. (Record verbatim.) When S has finished ('right, anything else?'-'no') ask ALL the following questions (saying 'You got that,_[Correct answer]' if S already has).

1. Which country was the city in?
2. How many schoolchildren were there?
3. Whose birthday was it?
4. How old was he?
5. What did the Prime Minister release?
6. What did one of them perch on?

Give answers if S doesn't know; tell S if correct.

SCORE: as for immediate test.

FACE-NAME LEARNING (Delayed)

"And now d'you remember earlier I showed you pictures of 4 faces and asked you to try and remember their names. Can you remember this person's name?" and so on. Prompting as before.

SCORE: as before.

WEIGL TEST

Scatter the blocks. "I'd like you to sort these pieces into groups." (If S asks how: "any way you like, just sort them into groups.")(that

go together.) Time limit to start 15". Timing starts from end of instructions.
Record: Principle 1
Time 1
If fails, do one grouping according to Form (all circles together) and say "Can you group the rest in this way, please."
Record: Time 1*
If no start in 15" or if fails, Stop test.
After successful grouping (assisted or not) ask "Right. Why do these ones go together/Why have you grouped them in this way?" Record verbatim S's answer to this question and score whether correct identification of principle, yes or no.

"Now could you group them again, but in a different way?" Time limit to start 15".
Record: Principle 2
Time 2
If fails, do one grouping according to Colour (all blues together), or form (circles together) if S spontaneously sorted by colour before, and say "Could you group the rest in this way, please?"
Record: Time 2*
If no start in 15" or if fails, Stop test.
After successful grouping (assisted or not) ask "Right. And why do these ones go together/Why have you grouped them in this way?" Record verbatim S's answer and score whether correct identification of principle, yes or no.

SCORE:

For each principle achieved spontaneously 3; after model 2

For each principle verbally identified 1

Time bonus if achieved spontaneously 0-10" 3 after model 0-10" 1
11-20" 2
21-30" 1

Total = 00 - 14

OBJECT RECOGNITION AND NAMING

"Could you tell me what these pictures show, please?" Record S's attempts verbatim. Time limit 15" in each case. If correct, record latency of response in seconds. If fails, ask S if can tell or show me how the object is used/what it is used for and record this.

SCORE: (a) No. of objects named
(b) No. of objects recognised

Total = 00 - 23 in each case.

Naming 'proportion' later computed = (a)/(b)
[In sub-sample, after above procedure show S real object where possible (asterisked on score-sheet) and see if can name, show use, before and after handling the object.]

VISUAL PERCEPTION WITH INTERFERENCE

"Now what objects can you see in this jumbled-up drawing?" (Poppelreuter figure).

SCORE 1 each

Total = 0 - 4

"And what object is shown beneath the criss-cross lines here?" (Table). "And which words are shown under the lines here?" ('big, just, nurse').

SCORE 1 each

Total = 0 - 4

NAMING OF BODY PARTS

"Could you tell me what this is called please?" Point in turn to Thumb, (Fore)finger, Elbow, Knee, Nose, Eye, Ear. Time limit 15" in each case. If correct, note latency.

SCORE 1 each

Total = 0 - 7

Note time testing ended.

"Thank you very much Mr/Mrs_____ (etc).

Note: k Any R
l Any acknowledgement of departure (eg nod, wave, etc)
m Any vbl R
n Approp vbl R

and add this to the initial interaction ratings.

SCORE: 1 point for each letter circled.

Total = 00 - 14

LANGUAGE RATINGS (Minnesota) Circle on scoresheet.

Understanding What Is Said.
0 No observable impairment
1 Follows general discussion with only minimal difficulty
2 Follows ordinary conversation with little difficulty
3 Follows most conversation but sometimes fails to grasp essentials
4 Follows simple conversation but requires repetition
5 Follows brief statements with considerable repetition
6 Usually responds innapropriately because does not understand

Speech.
0 No observable impairment
1 Converses easily with occasional difficulty
2 Conversational speech, with mild difficulty finding words or expressing ideas
3 Some conversational speech but marked difficulty in expressing

long or complex ideas
4 Ready communication with single words and short phrases
5 Expresses needs and wishes in limited or defective manner
6 No functional speech

Dysarthria.
0 No observable impairment
1 Occasional hesitation or slurring
2 Intelligible speech with mild slurring or slowness
3 Intelligible but obviously defective speech
4 Frequent omissions and substitutions of sounds
5 Speech usually barely intelligible
6 Speech usually unintelligible

Also note specific features of language:

Para	Paraphasic errors
WF	Word finding difficulty, circumlocution
J	Jargon speech
Pers	Perseverative speech
C&S	Cliches, Stereotyped phrases
E	Echolalia

In each case score: 1 Slight
2 Marked
If not present, do not score. (eg Para 1, J 1)etc.

Session Ratings

	Score	Variation
1 Impaired consciousness	0 1 2 3 4	0 1 2
2 Apparent confusion	0 1 2 3 4	0 1 2
3 Attention	0 1 2 3 4	0 1 2
4 Concentration	0 1 2 3 4	0 1 2
5 Distractibility	0 1 2 3 4	0 1 2
6 Engagement	0 1 2 3 4	0 1 2
7 Cooperation	0 1 2 3 4	0 1 2
8 Interest in tests	0 1 2 3 4	0 1 2
9 Psychomotor slowing	0 1 2 3 4	0 1 2
10 Apparent fatigue	0 1 2 3 4	0 1 2
11 Anxiety	0 1 2 3 4	0 1 2
12 Emotional lability	0 1 2 3 4	0 1 2
13 Depression	0 1 2 3 4	0 1 2
14 Elation	0 1 2 3 4	0 1 2
15 Disinhibition	0 1 2 3 4	0 1 2
16 Confabulation	0 1 2 3 4	0 1 2
17 Perseveration	0 1 2 3 4	0 1 2
18 Distress at failure	0 1 2 3 4	0 1 2
19 Insight re own cond.	0 1 2 3 4	0 1 2
20 Perplexity	0 1 2 3 4	0 1 2
21 Paranoid features	0 1 2 3 4	0 1 2
22 Hostility/aggression	0 1 2 3 4	0 1 2
23 Suspiciousness	0 1 2 3 4	0 1 2

24 Delusions	0 1 2
25 Hallucinations	0 1 2
26 Thought disorder(not dysphasia)	0 1 2

27 Inappropr.grimaces/mannerisms	0	1	2
28 Bizarre sitting position	0	1	2
29 In constant motion	0	1	2
30 Motionless	0	1	2
31 Disarranged clothes	0	1	2
32 Drooling	0	1	2
33 Nasal mucous/food conspicous (on clothes,face)	0	1	2
34 S attempts to move away (without explanation)	0	1	2

Additional tests used on some but not all Ss who underwent full testing.

(a) Tests given after Naming Body Parts:

BOX-FILLING 2

As first presentation. 1st grid covered by folding paper.

SPONTANEOUS DRAWING

"Could you draw a picture of a clock. please?" ("Put in numbers; and hands." if necessary. Note what prompt is given.)

SCORE:

Circular face	Good	2
(or regular eg square)	Fair	1
	Bad	0

PLUS

12 Nos. (or marks) in exactly right place	5
slightly disorganised	4
Less than 12 (eg missed 1 or 2) in roughly correct places	3
12 or less in a mess (eg 1 side only)	2
Just a few stuck in anywhere	1
No numbers (or appropriate marks)	0

PLUS

Hands	Two, good	2
	Rough or just one	1
	No	0

Total = 0 - 9

"Could you now draw a man?"

SCORE:

Head	1			
Body	1		Symmetry/proportion	
2 arms	1			
2 legs	1		Good	2
Hands	.5	PLUS	Fair	1
Feet	.5		Bad	0
Eyes	.5			
Nose	.5			
Ears	.5			
Hair	.5			

Total = 0 - 9

(b) Tests given at another test session (or after second half of main testing, if this was split):

EXTENDED YES-NO PICTURE MEMORY

"I'm going to show you a pile of pictures/designs now. There's a hundred in all. The first twenty are all different, but after that some of the pictures/designs will be repeats - in other words you'll have seen some of them before. After the first twenty, then, I'll ask you if you've seen each one before and you just have to say yes or no. OK? So I'll just start turning over the cards, and you just have a good look at each picture/design and try to remember each one." Turn over every 3 seconds. Once R is required, ask S to guess if unsure. Use Pictures first.

SCORE:
Pictures:
Raw Score = Proportion of hits minus Proportion of false alarms (0.00 to 1.00)

See Tables for d' and B.

Repeat for Designs

KERTESZ AND HOOPER APRAXIA TEST

S is asked to carry out each movement. If no R or only amorphous/ approximate R is given, S is then shown the movement by E and is verbally and gesturally encouraged to imitate. (To reduce effects of comprehension deficit, S is credited with the best score of these 2 methods, if the second is req'd. Ie S is not penalised for requiring a demonstration.)

SCORE: 3 Good, standard performance
2 Impaired but recognisable performance
1 Poor but approximate performance

If S gets 0 on the * items, E gives actual object and asks S to use it. If doesn't manage, E demonstrates with the object and S is then encouraged as before to imitate with the object.

SCORE: 1 For adequate performance.

0 No performance, unrecognisable or unrelated gesturing or completely erroneous use of objects.

Total = 00 - 60

Facial	1	Put out your tongue
	2	Close your eyes
	3	Whistle
	*4	Sniff a flower
	*5	Blow out a match
Upper Limb	1	Make a fist
	2	Salute
	3	Wave goodbye
	4	Scratch your head
	5	Snap your fingers
Trans-itive	*1	Use a comb
	*2	Use a toothbrush
	*3	Use a spoon to eat
	*4	Use a hammer
	*5	Use a key
Complex	1	Pretend to drive a car
	2	Pretend to knock at the door
	*3	Pretend to fold a paper
	4	Pretend to light a cigarette
	5	Pretend to play the piano

SPATIAL DIGIT SPAN

Using Corsi board, same kind of procedure as for normal digit span. (Again no backwards span here.)

"You can see I have here a board with some small wooden blocks on it. I'm going to touch some of the blocks with my pencil/finger like this (demonstrate) and then I'd like you to touch the same blocks as I touched in the same order. OK?" Repeat or elaborate as necessary for S to understand the task requirement. Record the blocks S touches, in order.

Series	Trial 1	Trial 2
(2)	2-7	1-3
(3)	5-8-2	6-9-4
(4)	6-4-3-9	7-2-8-6
(5)	4-2-7-3-1	7-5-8-3-6
(6)	6-1-9-4-7-3	3-9-2-4-8-7
(7)	5-9-1-7-4-2-8	4-1-7-9-3-8-6
(1)	3	6

SCORE is longest sequence correctly performed, plus .4 if achieved on first trial.

0 Missed block (even on a single-block trial. Eg put finger between blocks etc.)

ENCODING MEMORY TEST

"I'm going to give you a list of ten words. Each word is written on a card; I'll show you the cards and I'll also say each word out loud. I want you to concentrate on the words and try to remember them, because then I'll ask you to pick which words I gave you from a choice. OK? So here are the words - you've just to concentrate on them and try to remember them."

Show at the rate of 1 card per 2 seconds. Turn over each card and as it lands say clearly the word showing. After the 10 words, about 30 seconds of chat to prevent rehearsal. Then show each of the 4-choice cards and ask "Which one of these did I give you?". Ask S to guess if does not know.

Repeat with second list.

		Target word	Semantic distractor	Acoustic distractor	Unrelated distractor
A.	1	CHAIR	SEAT	SHARE	HAND
	2	BLACK	DARK	SLACK	WISH
	3	FOOT	SHOE	PUT	HATCH
	4	SHORT	SMALL	PORT	TOOTH
	5	COLD	ICE	BOLD	ROUGH
	6	EARTH	DIRT	WORTH	LANE
	7	LAMP	LIGHT	STAMP	HEALTH
	8	SLEEP	REST	HEAP	SOUR
	9	HARD	STONE	YARD	JOY
	10	SWIFT	FAST	LIFT	EYES
B.	1	BREAD	FOOD	WED	CAT
	2	BATH	WASH	PATH	HELP
	3	LOUD	NOISE	CROWD	RULE
	4	HOUSE	HOME	GROUSE	SEED
	5	PRIEST	CHURCH	BEAST	LID
	6	STOVE	FIRE	COVE	BOOK
	7	GREEN	GRASS	KEEN	SMILE
	8	MAN	MALE	BAN	TEA
	9	THIEF	STEAL	GRIEF	RUG
	10	BLUE	SKY	FLEW	DRUM

(c) Tests given to different subjects:

COMPETING RESPONSES TESTS (always given in a separate session)

For each type of material, an appropriate introduction is given in the form "I'm going to read you a list of words/a paragraph/show you some pictures/designs and I'd like you to to concentrate hard on it/them and try to remember it/them because afterwards I'll ask you some questions and get you to pick out the answers from a choice/ask you to pick which ones you heard/saw from a choice."

After presentation of the material, about 30 seconds of chat to prevent rehearsal. Then show each of the appropriate choice arrays with appropriate questions. Repeat till all 12 of the stimulus

'sets' have been tested.

Paragraphs The first is adapted from the book 'The Amnesic Syndrome' by G. A. Talland; the others from old newspaper stories. Paragraphs are read clearly and fairly slowly. An example of an 8-choice card appears in the next section.

Paragraph 1: In a city in India, 5000 school children paraded in the main square to celebrate the 70th birthday of the Prime Minister. While reviewing the parade the Prime Minister released a number of white doves, the symbols of peace. One of them, however, perched on top of the Prime Minister's head while he took the salute.

Q1: Where was the city? (Filler answers from: America, Africa, England, Russia, France, Canada, Spain)

2: How many schoolchildren were there? (1000, 200, 10,000, 50, 500, 90, 2000)

3: Whose birthday was it? (queen, painter, soldier, king, cricketer, poet, businessman)

4: How old was he? (48, 65, 50, 55, 57, 63, 60)

5: What were released? (butterflies, dogs, streamers, prisoners, balloons, sparrows, bats)

6: What did one of them perch on? (a wall, a light, his foot, a roof, a tree, a chimney, the ground)

Paragraph 2: David Smith was last night cleared of the murder of Gayle Stewart, a teenager who died after being shot by police marksmen during a siege in Birmingham. He was, however, found guilty of the attempted murder of two police officers and gaoled for twelve years. The jury took seven hours to reach its decision. The verdict was greeted by loud cheering from the public gallery.

Q1: What was the man called? (John, Peter, Chris, Alec, James, Tom, Roger)

2: What was the girl's name? (Susan, Jane, Sally, Kate, Sheila, Ann, Gloria)

3: Where did the siege take place? (London, Glasgow, Chicago, Leeds, Aberdeen, York, Newcastle)

4: How long was the man gaoled for? (5 years, 6 months, 2 years, 10 years, 15 years, 9 months)

5: How long did the jury take to reach a verdict? (1 hour, 20 minutes, 2 days, 5 days, 3 hours, 5 minutes, 1 week)

6: How did the crowd react to the verdict? (silence, disgust, booing, weeping, laughing, distress, sympathy)

Paragraph 3: John Richards, the Irish rugby player, has resigned

from his job in order to go on Ireland's tour of South Africa this summer. He is employed as a manager with the Guinness company and was refused leave by the company for reasons to do with politics rather than sport. He has toured South Africa once before and feels the chance to do so again is worth giving up work for.

Q1: What was the man's surname? (O'brien, Jones, Gibson, Hunter, Munro, Thomson, Stewart)

2: What sport did he play? (tennis, hockey, bowls, football, cricket, squash, darts)

3: For which country? (Scotland, Austria, England, Australia, New Zealand, Italy, Wales)

4: Who does he work for? (Government, McEwan's, a bank, a publisher, Tate & Lyle, Sainsbury's, Ford)

5: What does he do? (storeman, clerk, mechanic, driver, foreman, joiner, plumber)

6: How many times has he been on tour to South Africa before? (3, 5, 0, 2, 6, 4, 7)

Word Lists Words are read at the rate of one word per 2 seconds. In every case the question is "which one of these was in the list I read?" An example of a 4-choice card appears in the next section.

List 1: take, will, dress, sat, act, line.
Fillers: side, day, want, name, wife, let, world, long, green, small, car, hand, age, start, town, felt, plant, learn, down, note, help, them.

List 2: cut, will, just, sport, fine, state.
Fillers: stop, once, hard, lost, book, left, talk, thing, wall, kind, out, room, end, took, three, work, door, home, month, give, quite, place.

List 3: try, head, bring, lie, sound, well.
Fillers: stand, good, near, reach, tell, late, child, more, year, court, when, show, true, done, hope, man, wish, must, yet, big, stood, call.

Photos All photographs were cut from multiple copies of colour magazines. Presentation rate 1 per 2 seconds. In every case the question is "which one of these did I show you?" A photocopy of a 4-choice card appears in the next section (the originals all being in colour).

Designs The designs used were mostly made up from scratch, but some came from a paper by Butters et al (1970, Cortex, vol 6, pp 440-459), and some from the Benton Visual Retention Test and the Graham-Kendall Memory for Designs test. Presentation and questions as in photos. An example of a 4-choice card appears in the next section.

APPENDIX 2

Details of the test materials.

Materials.

Appropriate scoring forms, not reproduced here, were used for each S for ease of recording.

All timing was carried out using a small silent digital electronic stopwatch.

All photocopied materials shown below are reproduced at 71% of actual size. Some materials are shown in a photograph at the end of this section.

Materials are described under the titles of the tests as given in the preceding manual (omitting those tests where no physical materials were required).

VISION/VISUAL FIELD/HEARING 3 white cards with arrays of large (1cm high) letters, small (typewritten) letters, and clusters of lines and dots for counting if necessary.

U B R F P M

y o h c n s

6 8 2 9 5 3

c s d a e

U B R F M

6 8 9 5 3

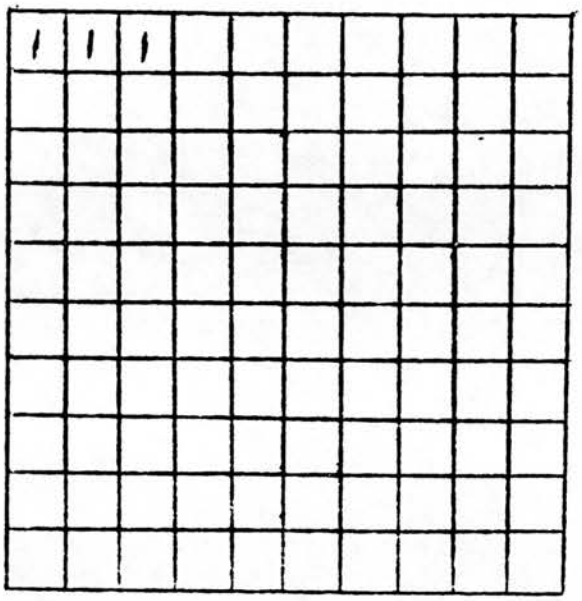
II I III

⁘ ∴ ::

II ⁘

BOX-FILLING 10 X 10 grids of 1cm squares, with a mark like a 1 in the first three, drawn on white A4 paper:

1



TOKEN TEST 20 tokens cut from hardboard 0.3cm thick as follows: 5 large circles (3cm diameter), 5 large squares (3cm side), 5 small circles (2cm diameter), and 5 small squares (2cm side). One of each type of token painted red, one green, one yellow, one white, and one black. Tokens laid out as shown in manual on a 31 X 24cm buff card. (See photograph.)

SUPPLEMENTARY DYSPHASIA TEST Ordinary teaspoon, mortis key, and pencil. (See photo.)

WRITING & READING Sentence printed in lower case lettering 1cm high in two lines on a 20 X 12.5cm white card:

The quick brown fox
jumps over the lazy dog.

ESTIMATED PREMORBID I.Q. N.A.R.T. words printed in capitals 1cm high, with 8 words to a 20 X 12.5cm white card, eg:

DEPOT

AISLE

BOUQUET

PSALM

CAPON

DENY

NAUSEA

DEBT

(6 words from the Burt reading test on the first card, before the start of the N.A.R.T. words.) W.A.I.S. words printed in a similar manner.

ARITHMETIC 8 plain wooden sticks measuring 9 X 0.6 X 0.6cm. (See photo.)

YES/NO PICTURE MEMORY 30 colour photographs, 9 X 7cm, taken from multiple copies of magazines, each mounted on a 14 X 9cm white card: 10 unique photos and 2 copies of another 10. Eg:

Table of d' and B (in parentheses) used in the 20-item Yes-No Picture Recognition Test.

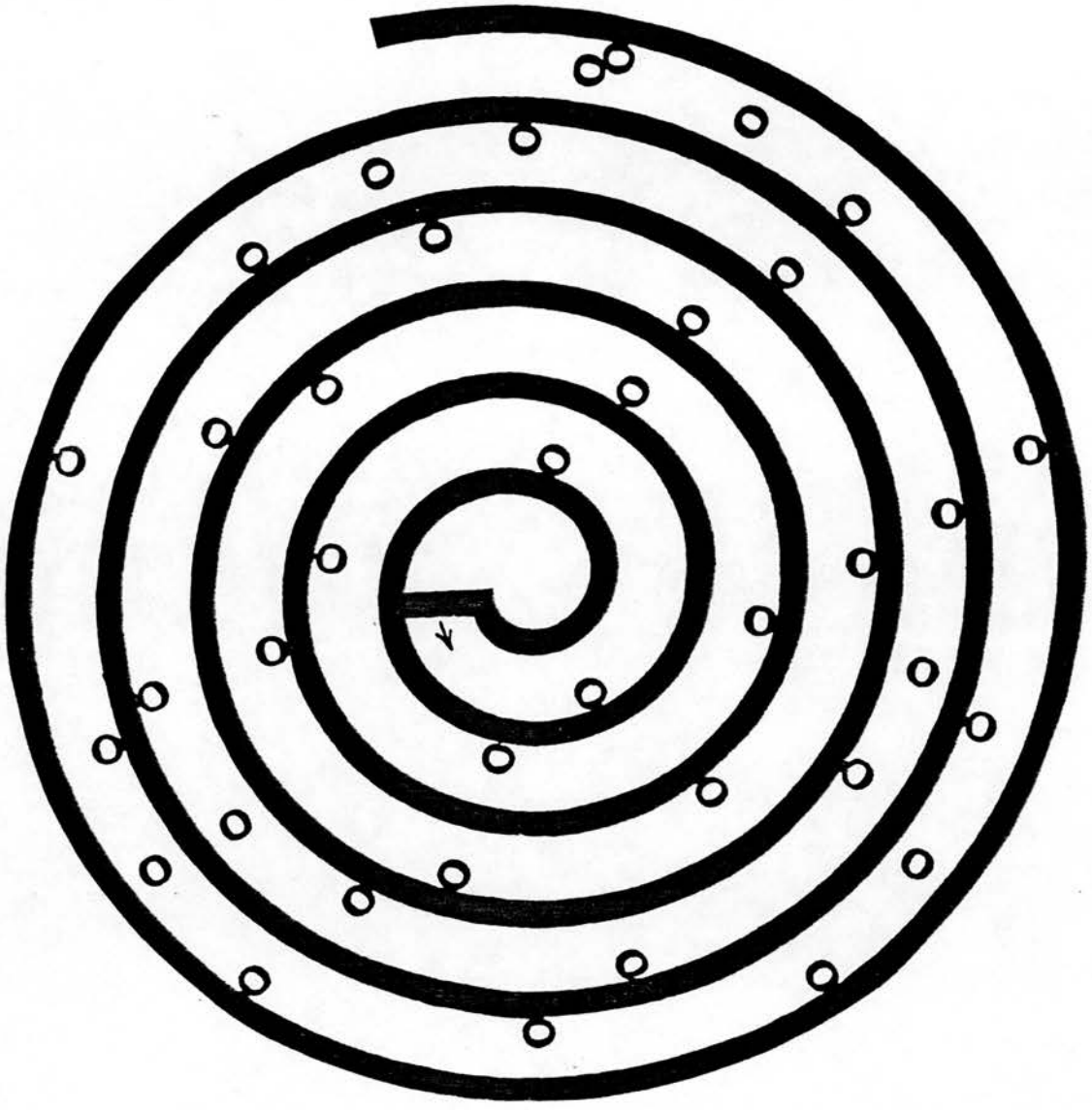
	False Alarms										
	0	1	2	3	4	5	6	7	8	9	10
Hits											
0	0.00	-.47	-.91	etc							
	(1.00)	(0.49)	(0.31)	etc							
1	0.47	0.00	-.44	-.76							
	(2.04)	(1.00)	(0.63)	(0.50)							
2	0.91	0.44	0.00	-.32	-.59						
	(3.25)	(1.60)	(1.00)	(0.81)	(0.72)						
3	1.23	0.76	0.32	0.00	-.27	-.52					
	(4.03)	(1.98)	(1.24)	(1.00)	(0.90)	(0.87)					
4	1.50	1.03	0.59	0.27	0.00	-.25	-.51				
	(4.48)	(2.20)	(1.38)	(1.11)	(1.00)	(0.97)	(1.00)				
5	1.75	1.28	0.84	0.52	0.25	0.00	-.25	-.52			
	(4.63)	(2.27)	(1.42)	(1.15)	(1.03)	(1.00)	(1.03)	(1.15)			
6	2.00	1.53	1.09	0.78	0.51	0.25	0.00	-.27	-.59		
	(4.48)	(2.20)	(1.38)	(1.11)	(1.00)	(0.97)	(1.00)	(1.11)	(1.38)		
7	2.27	1.81	1.37	1.05	0.78	0.52	0.27	0.00	-.32	-.71	
	(4.03)	(1.98)	(1.24)	(1.00)	(0.90)	(0.87)	(0.90)	(1.00)	(1.24)	(1.98)	
8	2.59	2.12	1.68	1.37	1.09	0.84	0.59	0.32	0.00	-.44	-.91
	(3.25)	(1.60)	(1.00)	(0.81)	(0.72)	(0.70)	(0.72)	(0.81)	(1.00)	(1.60)	(3.25)
9	3.03	2.56	2.12	1.81	1.53	1.28	1.03	0.76	0.44	0.00	-.47
	(2.04)	(1.00)	(0.63)	(0.50)	(0.45)	(0.44)	(0.45)	(0.50)	(0.63)	(1.00)	(2.04)
10	3.50	3.03	2.59	2.27	2.00	1.75	1.50	1.23	0.91	0.47	0.00
	(1.00)	(0.49)	(0.31)	(0.25)	(0.22)	(0.22)	(0.22)	(0.25)	(0.31)	(0.49)	(1.00)

This corner remains blank because no S performed worse than chance on this test.

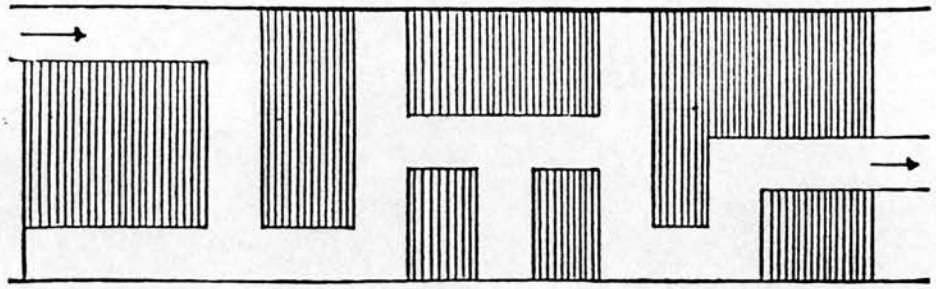
FACE-NAME LEARNING 4 colour photographs of people's faces (2 male, 2 female), 9 X 7cm, each mounted on a 14 X 9cm white card. (See photo.) Also a duplicate set.

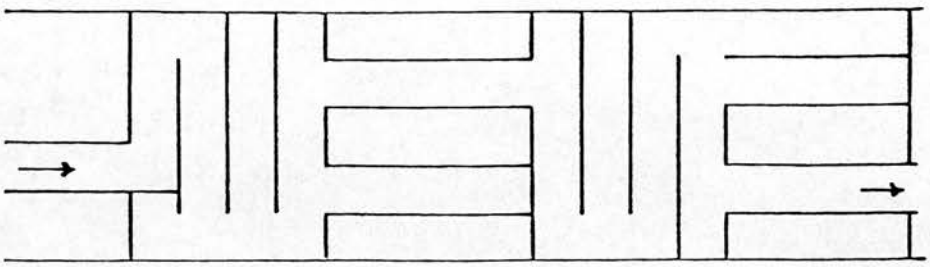
SPIRAL MAZES (1) Standard Gibson Spiral Maze with the outermost circuit removed and all the small circles 'whited out':

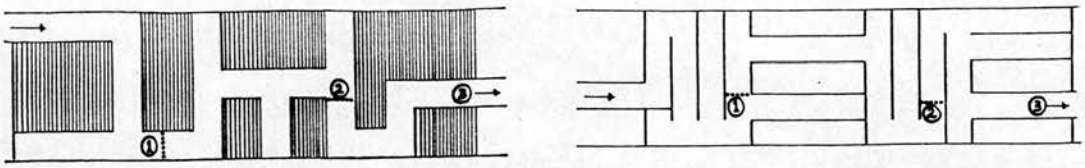
(2) Standard Gibson Spiral Maze with the outermost circuit removed but the small circles left in place:



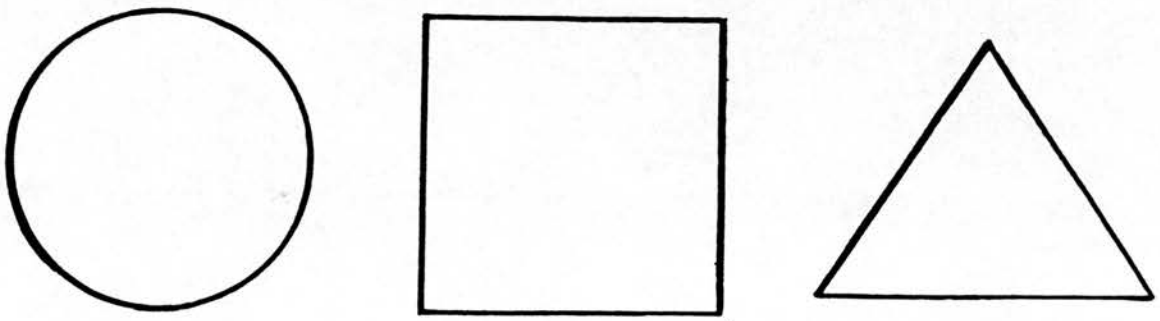
PORTEUS MAZES Standard mazes, years 5 and 6, with an arrow replacing the small rat in the former (Boundaries of 'thirds' of mazes shown below):



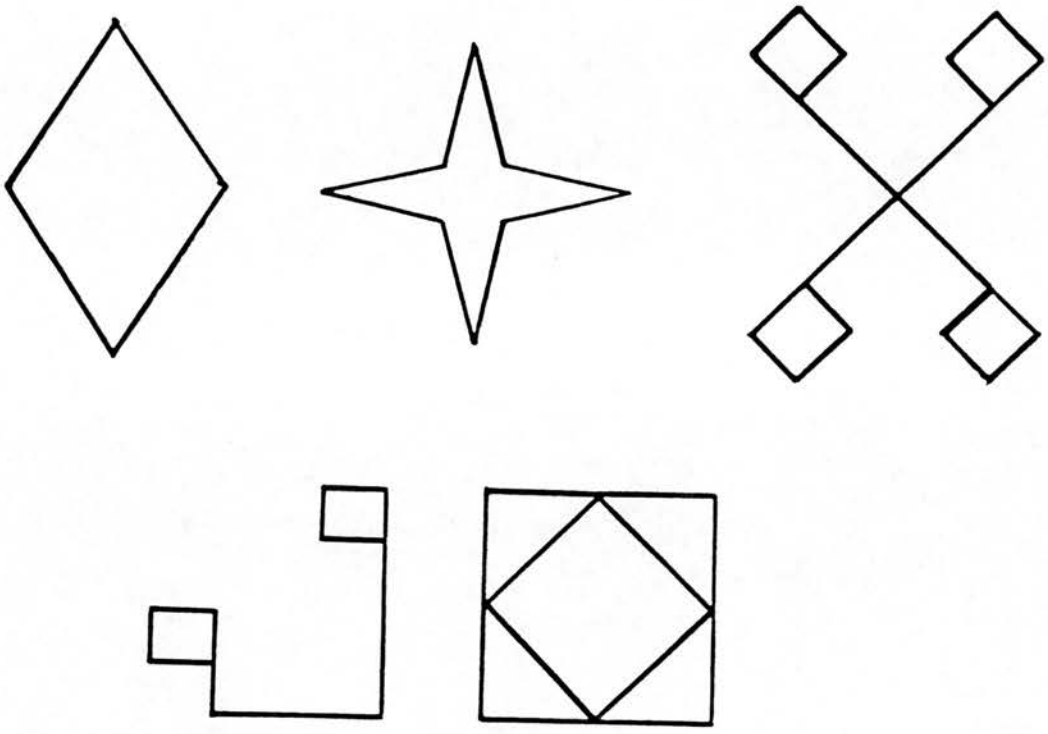




DRAWING SHAPES Circle, square, and triangle, each drawn on a 14 X 9cm white card:

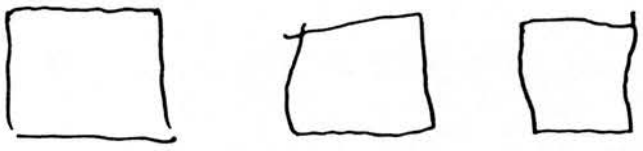


MEMORY FOR DESIGNS and COPYING DESIGNS Diamond, star, 'flags' from the Wechsler Memory Scale form 1, and the two-part design from the W.M.S. form 2, each drawn on a 14 X 9cm white card:



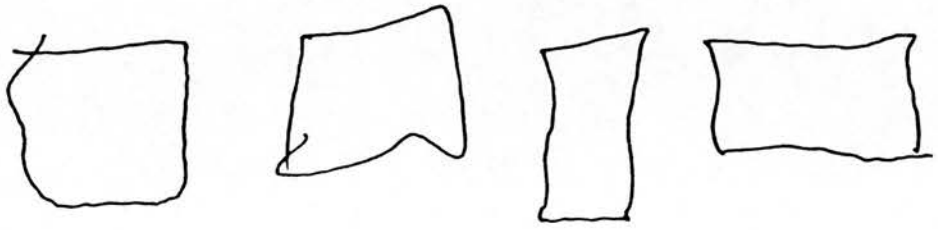
Aids to scoring:

2.

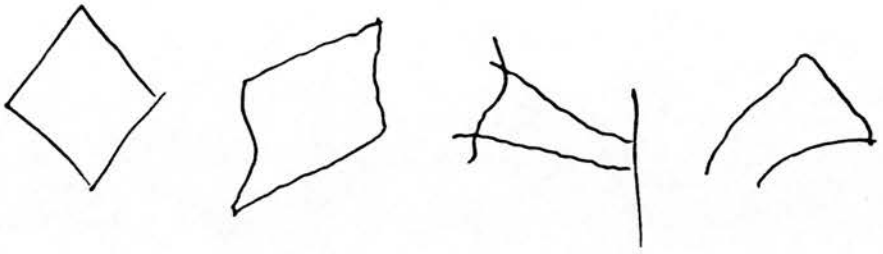


SQUARE

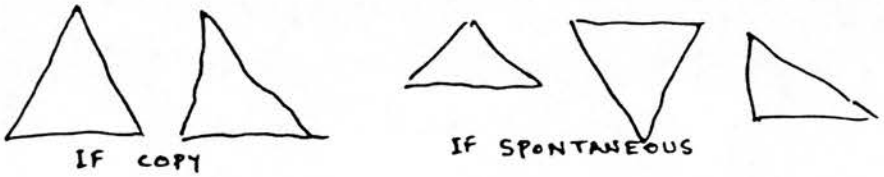
1.



0.



2.



1.

0.

TRIANGLE

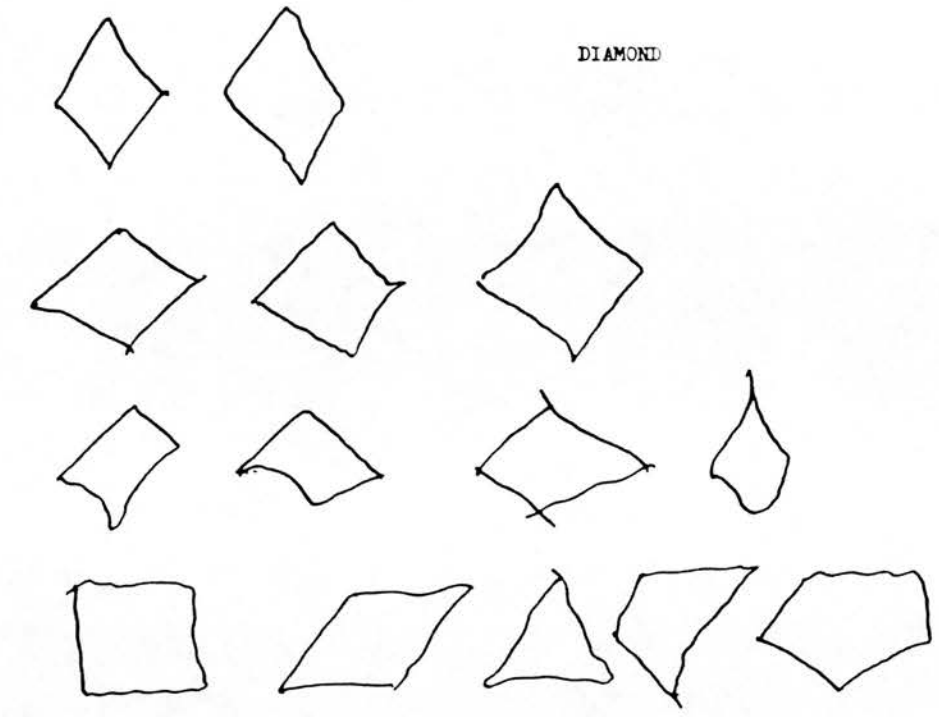
DIAMOND

3.

2.

1.

0.



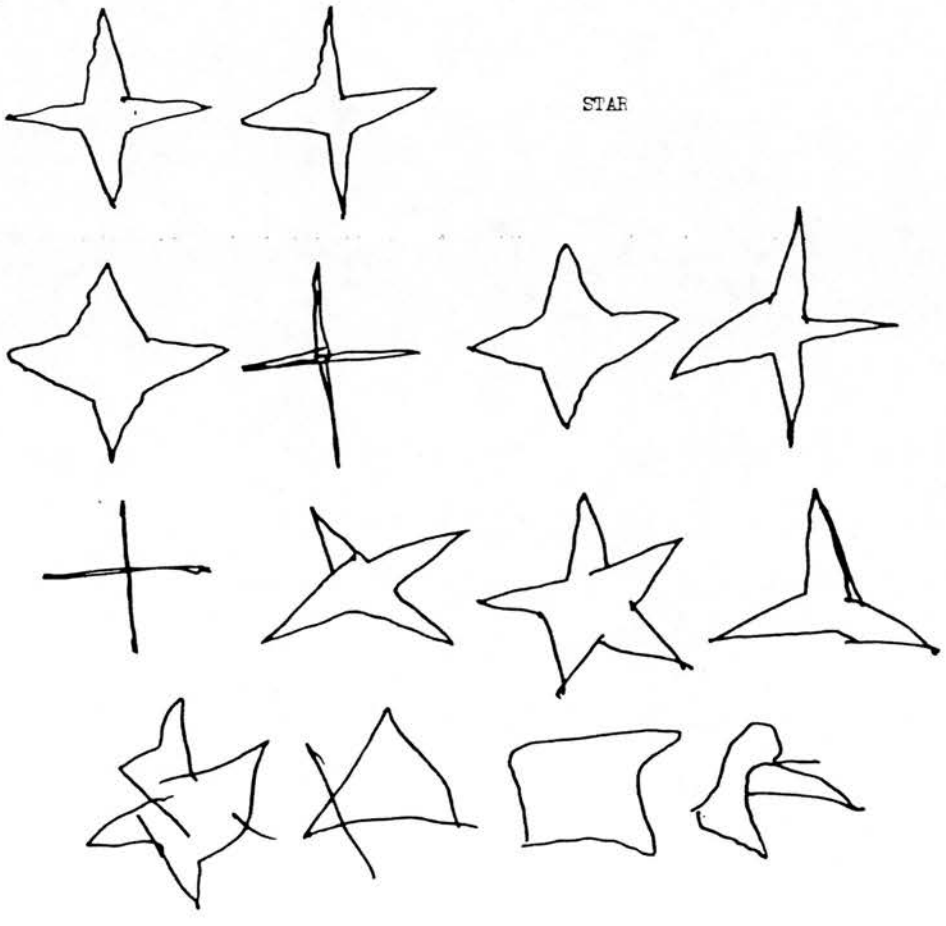
STAR

3.

2.

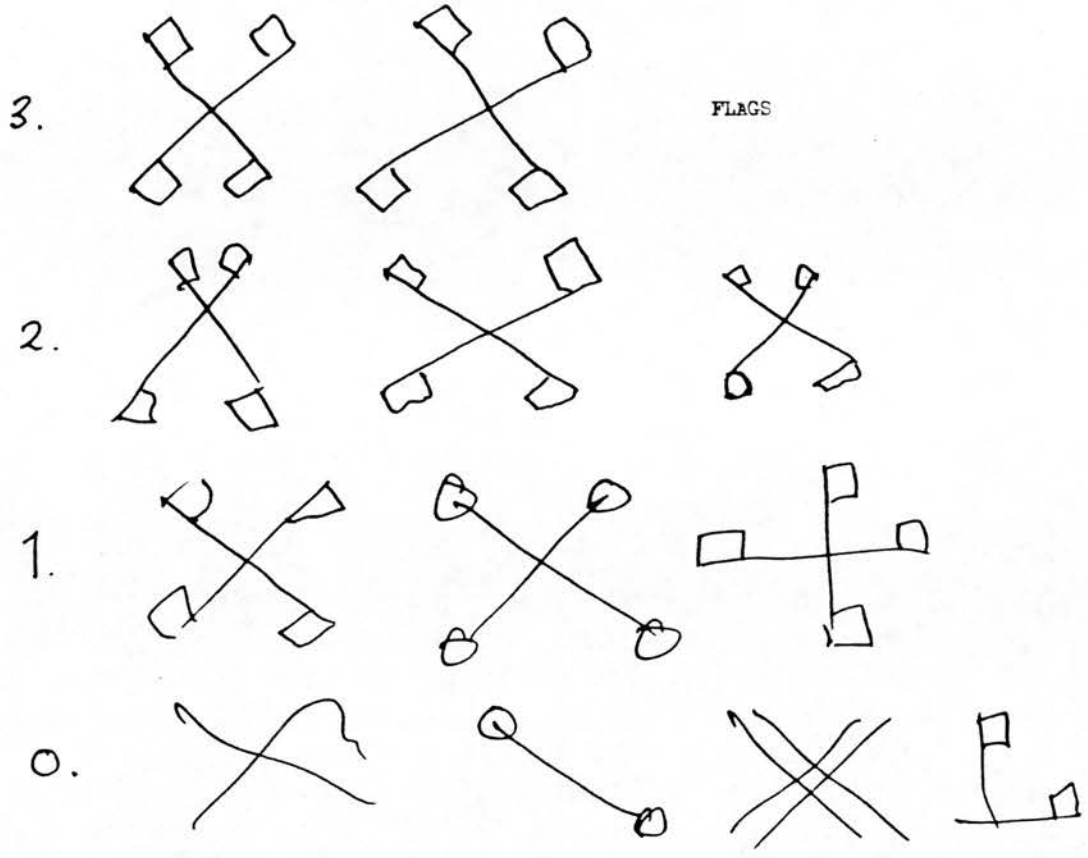
1.

0.



Scoring:

(*a*) 1. Two lines crossed, four flags 1
2. Correctly facing one another 1
3. Accuracy (lines nearly equal, nearly bisected, nearly at right angles; flags nearly square 1
Maximum score 3



FLAGS

Two Designs from the Binet, Form M

Scoring left-hand figure:

1 point for two vertical uprights connected by a base line, with the left upright clearly shorter than the right.

1 point for a small rectangle at the top of each vertical line, the size proportionate to model.

1 point if small rectangles face to the left at distal ends of the vertical lines, as in model.

Maximal score: 3 points for left-hand figure.

Scoring right-hand figure:

1 point for an outer, square-like quadrilateral.

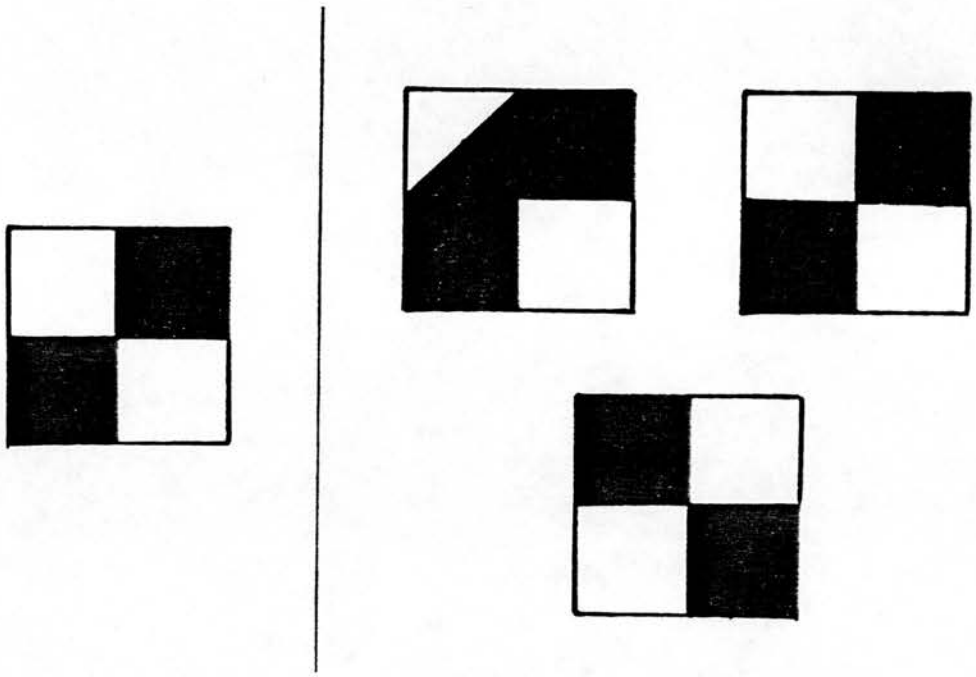
1 point for square-like quadrilateral inside the outer one, rotated 45 degrees, and with at least one corner touching the midpoint of a side of the outer figure.

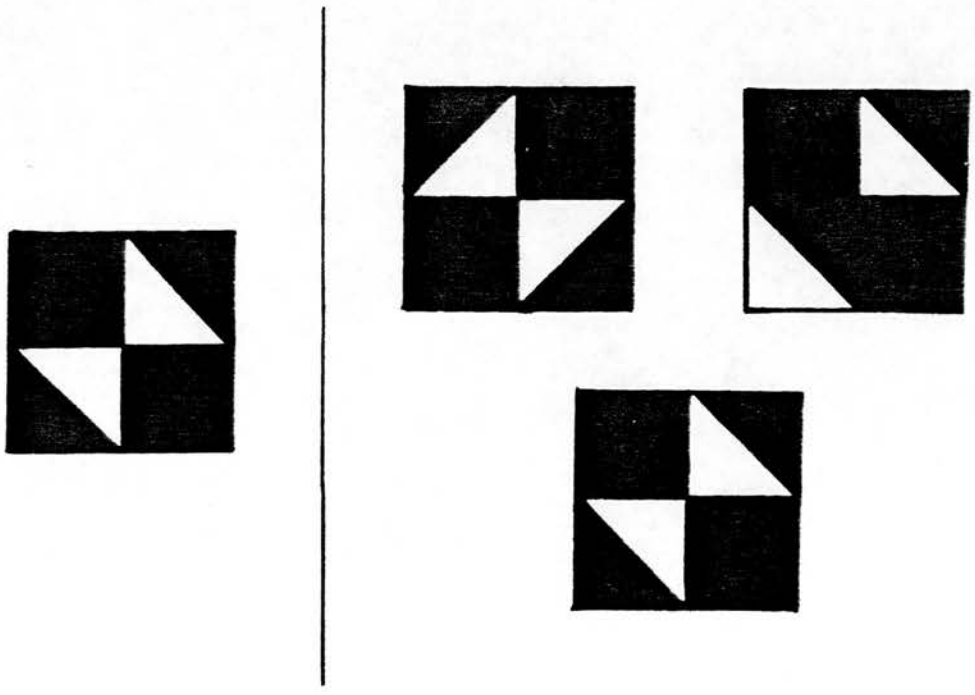
1 point if four corners of the inner figure touch approximately midpoints of sides of outer quadrilateral, and the angles of both quadrilaterals are approximately right angles.

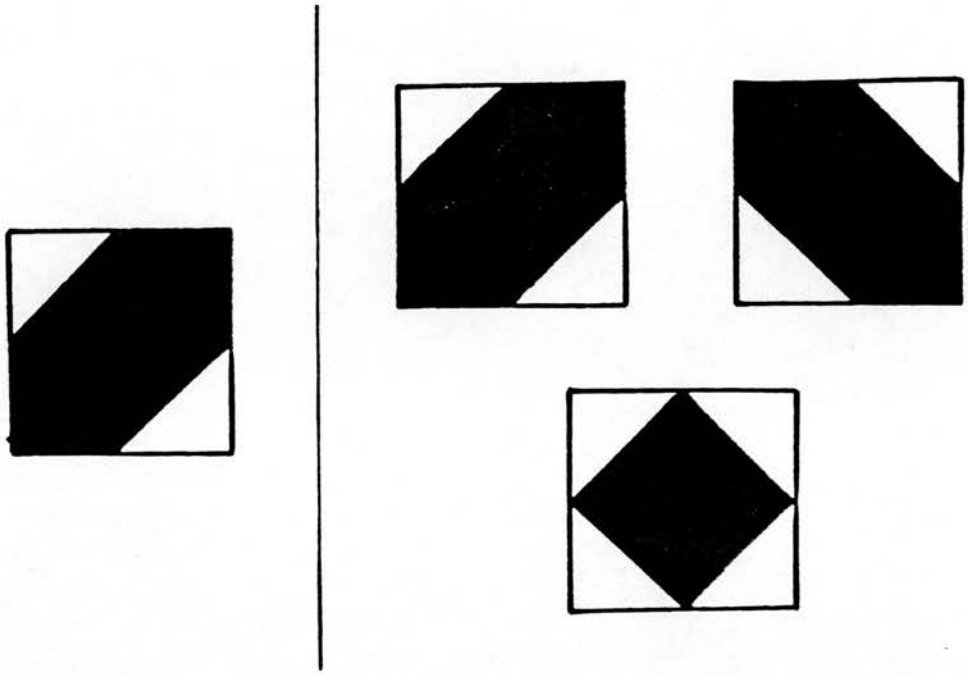
Maximal score: 3 points for right-hand figure.

SQUARE WITH STICKS 8 sticks as in ARITHMETIC.

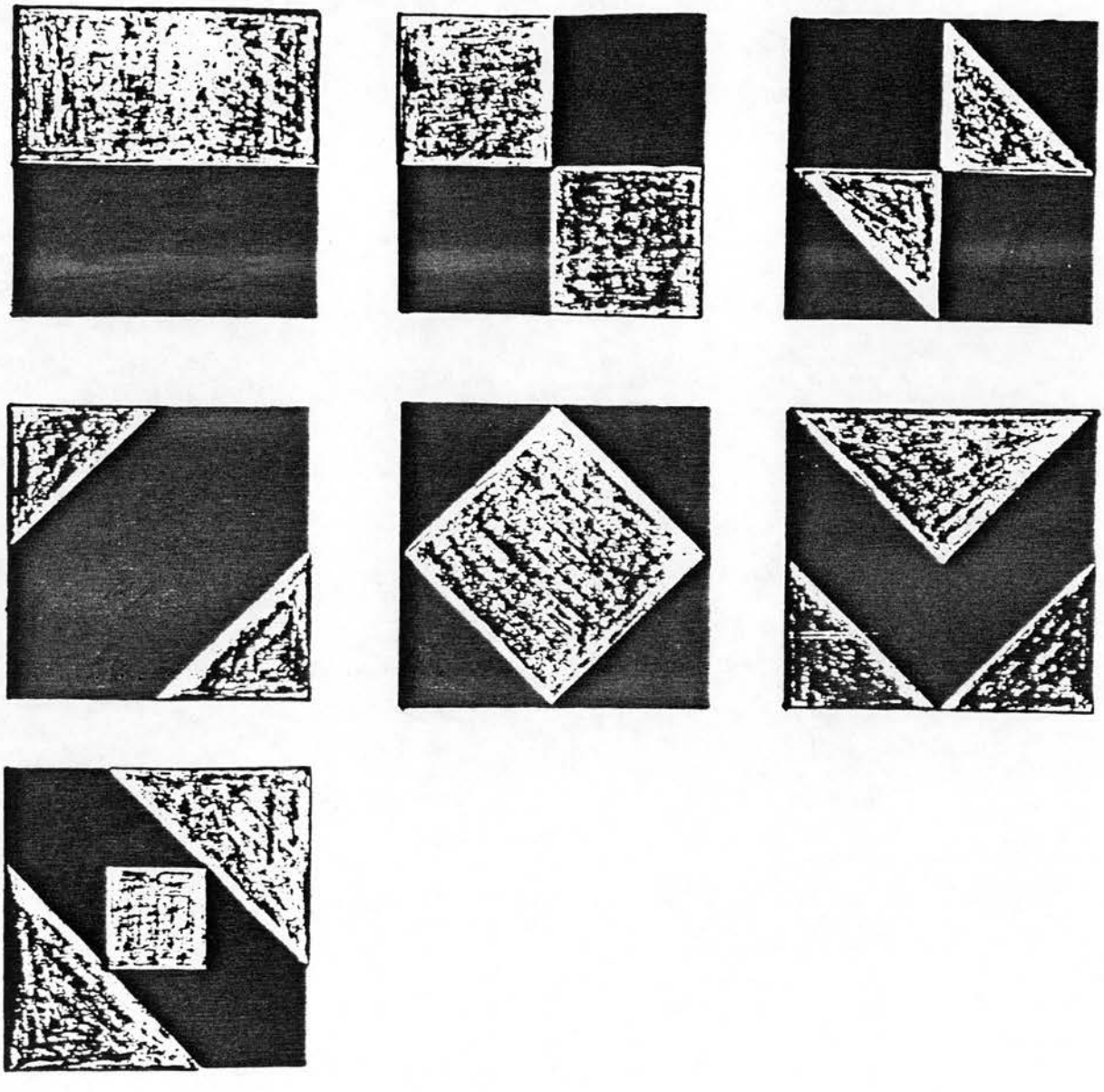
MULTIPLE-CHOICE BLOCK DESIGN 3 white cards, 20 X 12.5cm, each with 3 choice-designs and 1 target design (on the left, separated by a line); all designs 4 X 4cm:



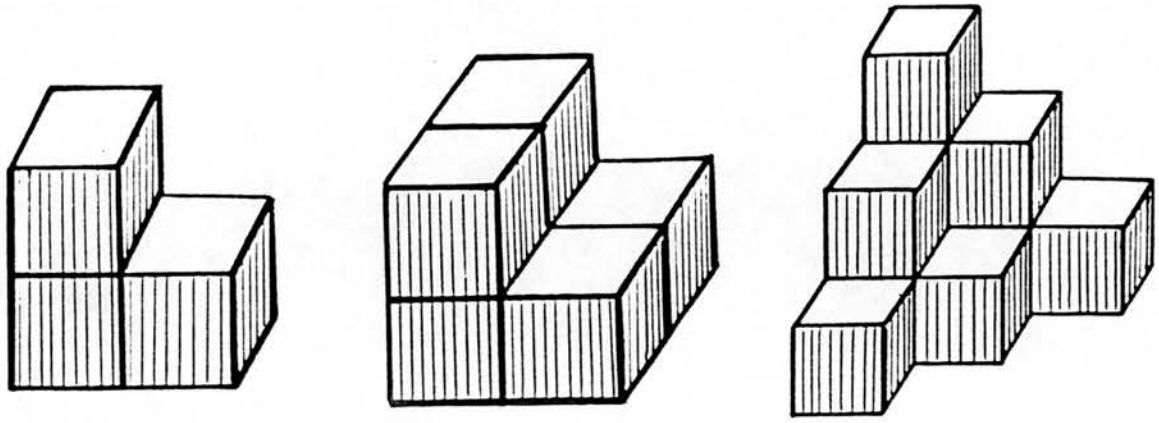




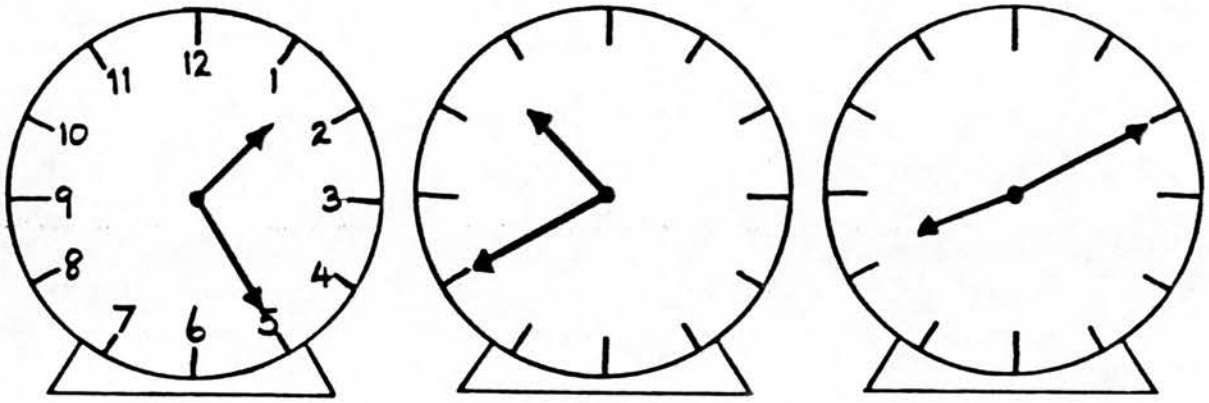
BLOCK DESIGN 9 blocks, 3.5 X 3.5 X 1cm, painted on the top surface only: 2 black, 3 yellow, and 4 half black and half yellow (divided along the diagonal). (See photo.) 7 model designs (W.I.S.C. 'A', W.I.S.C. 'B', W.A.I.S. no. 5, W.A.I.S. no.2, W.I.S.C. no.4, W.A.I.S. no. 6, and one made-up design), 6 X 6cm, each shown on a 14 X 9cm white card:



YERKES TEST Drawings of 3, 6, and 10-block piles, each on a white 14 X 9cm white card:

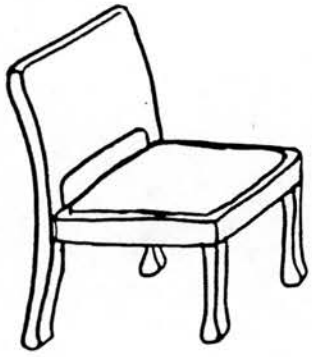


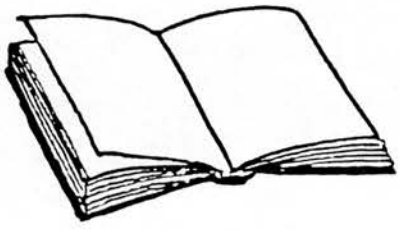
TIME-TELLING Clock faces showing 1:25, 10:40, and 8:10, each drawn on a 14 X 9cm white card:

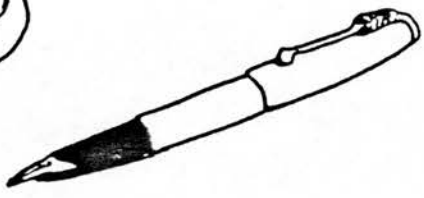


WEIGL TEST 12 pieces cut from hardboard 0.3cm thick as follows: 4 circles (3cm diameter), 4 squares (3cm side), and 4 triangles (3.5cm side). One of each shape painted red, one green, one yellow, and one blue. Pieces initially scattered randomly. (See photo.)

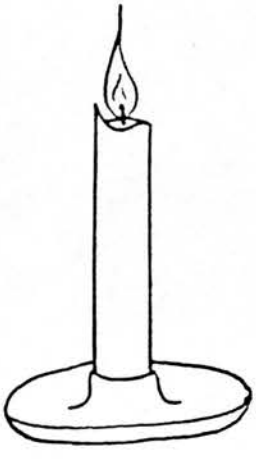
OBJECT RECOGNITION AND NAMING 23 line drawings of objects, each presented on a 14 X 9cm white card: Chair, table, book, key, pen, clock, candle, pencil, drum, hammer, spoon, fork, kettle, leaf, anchor, screw, comb, rake, bicycle, dice, whisk/mixer/egg-beater/switch, teapot, corkscrew/bottle-opener/awl/brad-awl:

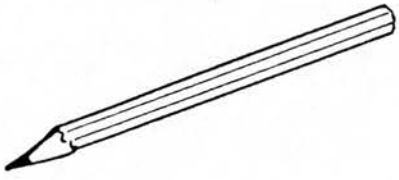


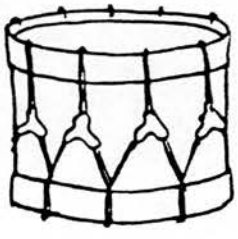


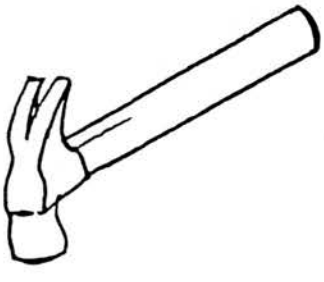



12
1
3
4
5
6
7
8
9
10
11

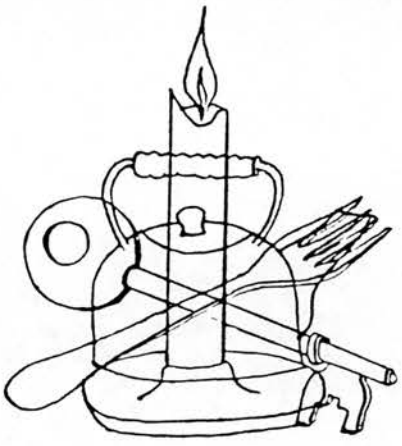


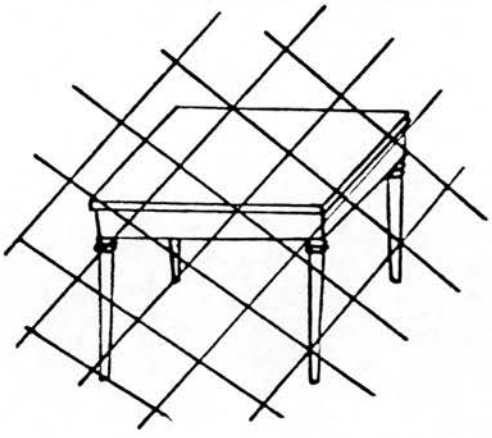






VISUAL PERCEPTION WITH INTERFERENCE (a) 4 of the line drawings used above, superimposed on each other, on a white 14 X 9cm white card: Candle, kettle, key, fork;
(b) Drawing of a table as used above, covered by criss-cross lines, on a 14 X 9cm white card, and 3 of the Burt words used as prelude to the N.A.R.T., covered by criss- cross lines, on a 14 X 9cm white card (big, just, nurse):

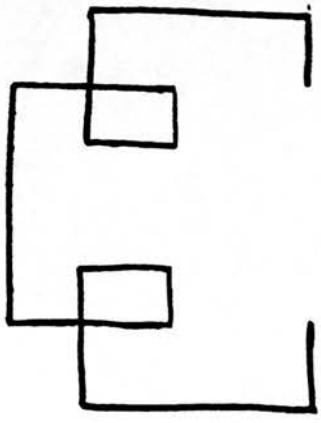




Photograph of test materials

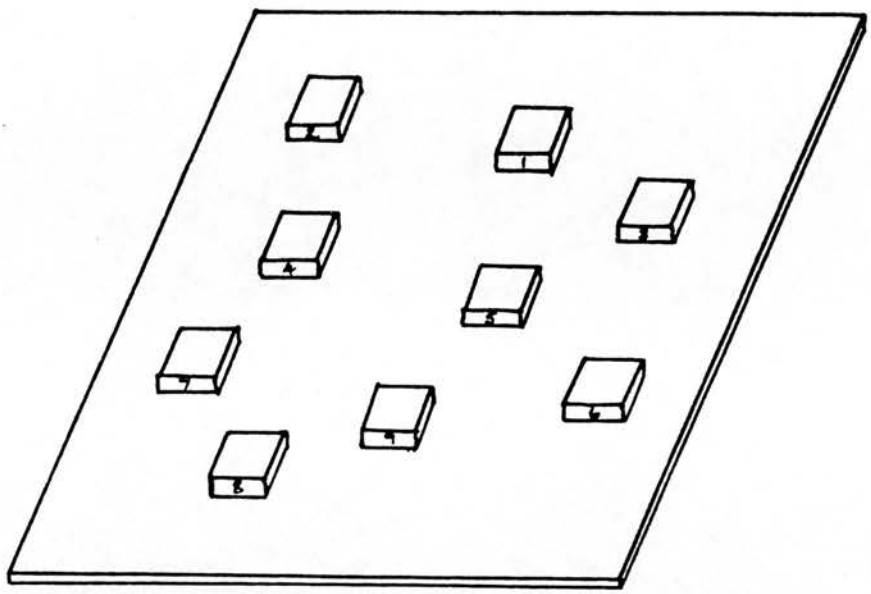
Materials used in additional tests.

EXTENDED YES-NO PICTURE MEMORY (a) 100 colour photographs, 9 X 7cm, cut from multiple copies of magazines, each mounted on a 13 X 10cm white card: 60 unique photos and 5 copies of another 8. (b) 100 designs, of a similar size to the photos, some taken from the Graham-Kendall M.F.D. test and from a paper by Butters et al (1970, Cortex, 6, 440-459) but mostly made up from scratch, each drawn on a 13 X 10cm white card: 60 unique designs and 5 copies of another 8. Eg:



KERTESZ AND HOOPER APRAXIA TEST Common objects as listed in the manual.

SPATIAL DIGIT SPAN 28 X 28cm hardboard square with small white wooden blocks, 2.5 X 2.5 X 0.5cm, glued on top as shown (E's view):



ENCODING MEMORY TEST 20 words, each printed in capitals 1cm high on 13 X 10cm white cards; 20 arrays of 4 words (target, semantic distractor, acoustic distractor, unrelated distractor), printed in capitals 1cm high, each 2 by 2 array on a 20 X 13cm white card, eg:

SHARE SEAT

CHAIR HAND

COMPETING RESPONSES TESTS (3 paragraphs and 3 word lists as given in the previous section.) 3 sets of 6 colour photographs, 6 X 5cm, taken from multiple copies of magazines. 3 sets of 6 designs, of a similar size to the photos and of a similar type to those used in Extended Yes/No Recognition. 4 pads of white A4 paper (one for each type of material), each containing 54 choice arrays (ie 3 times 18, to allow each target item to appear in a 2-choice, a 4-choice, and an 8-choice array). In the two verbal tests, choices printed in capitals 1cm high. In the two non-verbal tests, all choices of the same size and nature as the target items. Eg:

AMERICA INDIA

AFRICA FRANCE

ENGLAND CANADA

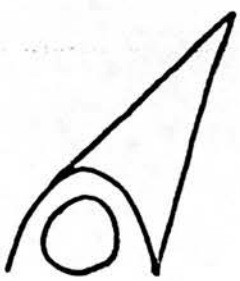
RUSSIA SPAIN

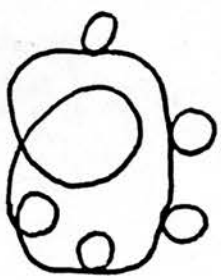
TOOK DOOR

THREE HOME

WORK MONTH

SPORT ROOM





BEHAVIOUR RATING SCALE Rater_____ Date_____

Name_____ DOB_____ S No._____

For each area of functioning listed below, please tick description that applies. (The word "assistance" is used here to mean supervision, direction, or personal assistance.)

BATHING (-either sponge bath, tub bath, or shower)

( ) Receives no assistance (gets in and out of tub by self if tub is usual means of bathing)
( ) Receives assistance in bathing only one part of the body (such as back or a leg)
( ) Receives assistance in bathing more than one part of the body (or not bathed)

DRESSING

( ) Gets all clothes and gets completely dressed without assistance

( ) Gets clothes and gets dressed without assistance except for assistance in tying shoes
( ) Receives assistance in getting clothes or in getting dressed, or stays partly or completely undressed. [If receives assistance, please tick one or both of the following:
  ( ) Because of physical difficulties
  ( ) Because gets muddled or confused over how to dress properly]

TOILETING

( ) Goes to 'toilet room', cleans self, and arranges clothes without assistance (May use object for support such as stick, walker, or wheelchair; and may manage night bedpan or commode, emptying same in morning)
( ) Receives assistance in going to 'toilet room' or in cleansing self or in arranging clothes after elimination or in use of night bedpan or commode
( ) Doesn't go to room termed 'toilet' for elimination process

TRANSFER

( ) Moves in and out of bed as well as in and out of chair without assistance (May be using object for support such as stick or walker)
( ) Moves in or out of bed or chair with assistance

( ) Bedridden or chairbound

CONTINENCE

( ) Controls urination and bowel movement completely by self

( ) Has occasional "accidents"

( ) Supervision helps keep urine or bowel control; catheter is used; or is incontinent [If incontinent, please tick one of the following:
( )Urine ( )Faeces ( )Both]

FEEDING

( ) Feeds self without assistance

( ) Feeds self except for getting assistance in cutting meat or buttering bread
( ) Receives assistance in feeding; or is fed partly or completely using tubes or intravenous fluids

ORIENTATION

( ) Can find way to toilet room, own bed, and eating area without assistance
( ) Can find way to one or two of these without assistance

( ) Cannot find way to any of these without assistance

CARE OF SLEEPING AREA

( ) Makes own bed and keeps area immediately around it tidy without assistance
( ) Makes own bed and tidies immediate area with assistance

( ) Doesn't make bed or tidy immediate area

CONFUSION

( ) Not confused (Aware of situation & surroundings; recognises familiar faces; can find way around and rarely gets lost or loses possesions)
( ) Moderately confused

( ) Severely confused; little or no grasp of situation

COPING OUTSIDE
If allowed outside, would need supervision:

( ) Never ( ) Sometimes ( ) Always

OCCUPATION
Keeps self occupied in a constructive or useful way (works, reads, plays games, has hobbies, etc):

( ) Much of the time ( ) Sometimes ( ) Almost never

COMPREHENSION OF SPEECH

( ) Understands almost everything you communicate

( ) Understands some of what you communicate

( ) Understands almost nothing of what you communicate

COMMUNICATION BY SPEECH

( ) Communicates well enough to be easily understood at all times

( ) Can be understood sometimes or with some difficulty

( ) Can rarely or never be understood for whatever reason

VARIABILITY
From day to day, functioning in the kind of areas listed above tends to vary or fluctuate:

( ) Very little or not at all

( ) Moderately

( ) Very noticeably

Over the course of a single day, functioning in these areas tends to vary or fluctuate:

( ) Very little or not at all

( ) Moderately

( ) Very noticeably

SOCIALISATION

( ) Establishes good relationships with others

( ) Has some difficulty establishing good relationships

( ) Has a great deal of difficulty establishing good relationships

Co-operation: Is willing to go along with requests:

( ) Usually ( ) Sometimes ( ) Almost never

Is objectionable to others (loud or constant talking, pilfering, soiling furniture, interfering with the affairs of others etc) during the day:

( ) Rarely or never

( ) Sometimes

( ) Frequently

And during the night:

( ) Rarely or never

( ) Sometimes

( ) Frequently

Accuses others of doing him/her bodily harm or stealing his/her personal posessions (if accusations are all definitely true, tick "never"):

( ) Never ( ) Sometimes ( ) Frequently

Hoards apparently meaningless items (wads of paper, bits of string, scraps of food, etc):

( ) Never ( ) Sometimes ( ) Frequently

Engages in exaggerated repetitive movements (eg pacing about, rocking, wringing hands, fiddling with clothing):

( ) Never ( ) Sometimes ( ) Frequently

Makes sounds which are not particularly directed towards others (shouting, muttering, moaning):

( ) Never ( ) Sometimes ( ) Frequently

Sleep pattern at night is:

( ) Almost never awake

( ) Sometimes awake

( ) Often awake

Night-time Sedation:

( ) Almost never has sedation

( ) Occasionally has sedation

( ) Regularly has sedation

Emotional lability.

( ) Mood is generally stable, with normal and understandable variation

( ) Moderate fluctuation in mood with some spells of laughing or crying for no apparent reason

( ) Marked emotional lability with frequent spells of innappropriate crying or laughing, the beginning and end of such spells often being sudden and abrupt

APPENDIX 3

Inter-test correlations.

Appendix 3. Inter-test Correlations.

Pearson product-moment correlations at test 1 between individual tests, overall scores, factor scores, and age in each diagnostic group and in the three combined.

The first (i.e. topmost) correlation in each group of four refers to all subject groups combined, the second to the DAT group, the third to the MID group, and the last to the Other group. For Full testing, the numbers of subjects on which these correlations are based are 174, 58, 58, and 58 respectively. For Short testing the numbers are 215, 74, 74, and 67. The only exceptions are with correlations involving I.Q.: here the numbers for Full testing are 151, 49, 49, and 53 respectively, and 153, 49, 50, and 54 for Short testing.

Significance levels for the correlations are not shown in the tables. Critical levels (for a two-tailed test) are as follows:

		Sig. level	
	n	.05	.01
Full testing (DAT,MID,Other)	58	.34	.26
Short testing (DAT,MID)	74	.30	.23
(Other)	67	.31	.24

The sizes of correlation required to reach significance in the combined subject groups (with numbers of 174 and 215 in Full and Short test groups respectively) are of course even smaller.

Abbreviations:

Ori	ORIENTATION	Full	Overall Full score
Par	PARAGRAPH RECALL	Sht	Overall Short score
MfD	MEMORY FOR DESIGNS		
YNP	YES-NO PICTURE MEMORY	I.Q.	Estimated premormid I.Q.
F-N	FACE-NAME LEARNING		
PM	PORTEUS MAZES	Fac1	Full factor 1 score
BD	BLOCK DESIGN	Fac2	Full factor 2 score
Wei	WEIGL TEST	Fac3	Full factor 3 score
Dig	DIGIT SPAN	Fac4	Full factor 4 score
SeR	SENTENCE REPETITION		
Tok	TOKEN TEST	FSh1	Short factor 1 score
SeP	SENTENCE PRODUCTION	FSh2	Short factor 2 score
W&R	WRITING AND READING		
ASp	AUTOMATIC SPEECH		
CD	COPYING DESIGNS		
MVT	MISC. VISUO. TASKS		
Ari	ARITHMETIC		
SM1	SPIRAL MAZE 1		
SM2	SPIRAL MAZE 2		
B-F	BOX-FILLING	SKP	Identifying SKP
ORe	OBJECT RECOGNITION	W/R	Writing name/Reading sentence
NA	NOMINAL ABILITY	SS	Square with Sticks

Full Testing

	Par	MfD	YNP	F-N	PM	BD	Wei	Dig	SeR	Tok	SeP	W&R
Ori	.75	.64	.61	.77	.62	.60	.62	.39	.60	.65	.51	.43
	.70	.64	.56	.70	.61	.59	.52	.43	.63	.67	.56	.51
	.77	.44	.49	.78	.42	.45	.55	.25	.46	.62	.45	.41
	.62	.65	.62	.63	.63	.54	.55	.48	.48	.50	.48	.45
Par		.64	.58	.75	.60	.55	.64	.31	.54	.56	.48	.25
		.51	.45	.64	.48	.35	.39	.27	.53	.47	.40	.28
		.57	.48	.73	.54	.49	.59	.17	.49	.58	.46	.30
		.58	.64	.65	.39	.50	.62	.33	.43	.46	.52	.18
MfD			.59	.62	.71	.77	.72	.46	.51	.57	.47	.30
			.55	.43	.61	.74	.71	.38	.51	.60	.46	.43
			.54	.47	.69	.78	.65	.31	.31	.44	.42	.27
			.62	.60	.66	.65	.58	.54	.49	.52	.47	.21
YNP				.56	.55	.56	.49	.29	.47	.48	.43	.33
				.59	.60	.58	.41	.31	.48	.51	.42	.35
				.40	.40	.52	.39	.11	.22	.27	.19	.34
				.54	.47	.51	.55	.41	.59	.63	.65	.24
F-N					.56	.53	.53	.30	.46	.48	.37	.28
					.44	.46	.35	.19	.40	.30	.27	.27
					.46	.41	.44	.13	.29	.49	.33	.31
					.47	.41	.37	.37	.46	.36	.38	.26
PM						.68	.61	.42	.49	.55	.42	.37
						.52	.49	.33	.54	.58	.42	.49
						.66	.52	.23	.24	.34	.35	.19
						.70	.61	.58	.51	.60	.39	.35
BD							.72	.45	.52	.60	.48	.38
							.64	.42	.56	.64	.55	.52
							.62	.24	.30	.40	.42	.30
							.72	.53	.52	.64	.37	.33
Wei								.50	.57	.66	.48	.33
								.44	.58	.66	.38	.45
								.33	.42	.56	.51	.33
								.61	.53	.66	.51	.26
Dig									.63	.58	.48	.32
									.59	.68	.46	.52
									.53	.42	.44	.20
									.75	.60	.52	.24
SeR										.78	.65	.47
										.80	.57	.55
										.74	.71	.51
										.71	.64	.40

	ASp	CD	MVT	Ari	SM1	SM2	B-F	ORe	NA	Age	Full	Sht
Ori	.48	.58	.68	.64	.34	.37	.37	.56	.50	-.58	.80	.72
	.51	.50	.69	.61	.34	.37	.40	.49	.44	-.24	.81	.75
	.29	.44	.50	.61	.36	.39	.27	.59	.51	-.62	.73	.64
	.35	.59	.58	.55	.32	.28	.42	.46	.34	-.41	.73	.63
Par	.36	.44	.56	.60	.39	.46	.42	.44	.39	-.54	.75	.64
	.24	.24	.53	.46	.18	.27	.28	.38	.36	-.29	.60	.51
	.32	.40	.50	.67	.51	.53	.44	.59	.37	-.57	.78	.68
	.31	.39	.41	.47	.39	.44	.46	.39	.43	-.24	.70	.54
MfD	.44	.76	.73	.68	.46	.49	.52	.41	.29	-.57	.83	.73
	.36	.70	.67	.63	.48	.47	.49	.34	.28	-.33	.79	.70
	.42	.80	.72	.67	.52	.57	.55	.48	.10	-.47	.80	.73
	.38	.70	.65	.63	.35	.35	.46	.44	.42	-.40	.77	.62
YNP	.31	.53	.56	.54	.31	.33	.36	.43	.25	-.39	.68	.57
	.29	.48	.51	.48	.22	.21	.40	.28	.20	-.20	.66	.53
	.05	.50	.51	.48	.40	.43	.33	.65	.08	-.38	.59	.51
	.36	.53	.57	.53	.32	.32	.38	.59	.36	-.34	.73	.55
F-N	.30	.46	.52	.49	.33	.38	.41	.38	.30	-.60	.70	.60
	.20	.31	.39	.35	.04	.02	.29	.26	.19	-.23	.53	.38
	.14	.35	.37	.53	.43	.44	.37	.51	.27	-.66	.66	.57
	.28	.42	.43	.33	.39	.44	.46	.22	.31	-.35	.65	.58
PM	.42	.67	.71	.58	.56	.58	.58	.42	.21	-.56	.79	.81
	.30	.60	.63	.45	.51	.53	.50	.33	.13	-.49	.74	.74
	.33	.65	.67	.51	.65	.68	.71	.51	.02	-.40	.74	.80
	.48	.68	.72	.63	.52	.50	.52	.53	.33	-.51	.79	.79
BD	.49	.76	.77	.63	.56	.60	.58	.41	.25	-.48	.83	.77
	.44	.73	.68	.61	.44	.44	.49	.34	.21	-.17	.79	.68
	.43	.77	.76	.56	.61	.69	.67	.46	.04	-.43	.79	.78
	.46	.73	.76	.58	.60	.59	.53	.63	.53	-.33	.82	.76
Wei	.50	.66	.71	.65	.51	.51	.51	.33	.33	-.55	.81	.73
	.44	.60	.63	.66	.47	.39	.39	.20	.19	-.35	.72	.63
	.43	.67	.67	.59	.49	.52	.49	.31	.38	-.50	.78	.70
	.50	.56	.68	.58	.54	.50	.56	.58	.46	-.29	.81	.72
Dig	.52	.44	.53	.56	.31	.26	.35	.28	.37	-.23	.59	.59
	.56	.42	.56	.60	.35	.33	.41	.33	.41	-.02	.63	.65
	.44	.30	.35	.33	.10	.09	.18	.13	.46	-.09	.41	.42
	.51	.60	.63	.68	.34	.24	.31	.53	.30	-.24	.69	.63
SeR	.61	.54	.67	.72	.30	.28	.33	.49	.55	-.30	.75	.69
	.59	.50	.79	.75	.33	.28	.33	.55	.55	-.04	.79	.74
	.53	.38	.41	.64	.14	.20	.19	.34	.63	-.18	.61	.53
	.62	.62	.60	.65	.34	.25	.36	.51	.29	-.21	.73	.65

	I.Q.	Fac1	Fac2	Fac3	Fac4	FSh1	FSh2
Ori	.39	.72	.29	.13	.44	.36	.68
	.39	.51	.40	.17	.44	.33	.69
	.42	.70	.14	.10	.51	.34	.55
	.41	.67	.33	.20	.01	.38	.57
Par	.50	.75	.12	.22	.38	.43	.47
	.29	.55	.13	.07	.41	.22	.46
	.59	.74	.07	.29	.43	.49	.40
	.58	.73	.06	.29	.28	.44	.29
MfD	.48	.75	.42	.28	-.02	.54	.48
	.43	.55	.48	.30	.07	.54	.42
	.48	.72	.37	.37	-.28	.60	.32
	.54	.76	.38	.21	-.35	.41	.48
YNP	.37	.70	.22	.12	.16	.36	.45
	.39	.64	.30	.06	.12	.31	.41
	.34	.70	.04	.22	.01	.43	.21
	.38	.70	.31	.15	.07	.34	.48
F-N	.35	.79	.05	.17	.30	.38	.46
	.20	.73	.06	-.08	.23	.09	.43
	.47	.75	-.08	.24	.38	.44	.29
	.32	.67	.02	.31	.22	.44	.35
PM	.48	.62	.35	.45	.03	.67	.44
	.44	.48	.37	.42	.09	.64	.38
	.57	.61	.17	.61	-.17	.78	.18
	.44	.49	.48	.42	-.35	.58	.50
BD	.40	.61	.48	.43	-.04	.62	.46
	.48	.45	.59	.30	.03	.46	.48
	.33	.62	.33	.55	-.28	.71	.25
	.40	.49	.46	.49	-.30	.62	.39
Wei	.45	.60	.50	.32	.02	.53	.50
	.44	.39	.61	.25	-.03	.44	.43
	.36	.57	.42	.31	-.03	.51	.42
	.56	.48	.48	.41	-.21	.55	.44
Dig	.39	.17	.73	.11	.07	.25	.67
	.53	-.07	.74	.21	.24	.27	.66
	.26	.01	.67	-.07	.05	.04	.70
	.38	.32	.78	.11	-.40	.29	.75
SeR	.46	.34	.70	.05	.38	.23	.81
	.57	.20	.70	.09	.46	.23	.79
	.40	.14	.70	-.09	.43	.05	.85
	.40	.33	.74	.13	-.10	.30	.78

	SeP	W&R	ASp	CD	MVT	Ari	SM1	SM2	B-F	ORe	NA	Age
Tok	.68	.51	.59	.65	.78	.77	.39	.37	.38	.49	.57	-.34
	.70	.68	.68	.68	.83	.80	.48	.44	.43	.52	.54	-.14
	.70	.64	.54	.54	.67	.75	.34	.36	.24	.41	.67	-.22
	.57	.31	.42	.62	.79	.67	.29	.21	.34	.72	.61	-.24
SeP		.46	.57	.53	.58	.60	.32	.32	.34	.49	.43	-.24
		.51	.56	.54	.63	.56	.33	.40	.39	.62	.46	.01
		.56	.54	.47	.55	.64	.29	.32	.27	.34	.46	-.22
		.34	.57	.51	.47	.51	.30	.20	.31	.36	.20	-.27
W&R			.50	.47	.49	.45	.39	.33	.36	.24	.30	-.11
			.62	.61	.60	.63	.46	.45	.35	.25	.32	.02
			.39	.43	.47	.48	.32	.33	.25	.28	.50	-.14
			.43	.37	.45	.25	.38	.26	.47	.19	.07	-.11
ASp				.50	.59	.58	.35	.35	.42	.43	.46	-.17
				.45	.59	.64	.33	.29	.48	.46	.53	.09
				.44	.47	.49	.15	.31	.35	.15	.33	-.01
				.47	.54	.46	.49	.40	.38	.46	.11	-.17
CD					.80	.69	.48	.43	.41	.43	.30	-.39
					.72	.65	.52	.39	.35	.35	.25	-.14
					.80	.63	.47	.51	.42	.41	.16	-.32
					.80	.76	.43	.33	.40	.53	.38	-.23
MVT						.75	.53	.52	.51	.60	.51	-.42
						.78	.49	.44	.45	.61	.56	-.11
						.68	.65	.69	.56	.51	.32	-.31
						.70	.49	.40	.47	.73	.52	-.29
Ari							.40	.39	.39	.47	.46	-.33
							.41	.37	.35	.41	.50	.07
							.43	.48	.37	.58	.42	-.36
							.31	.26	.36	.61	.30	-.24
SM1								.87	.69	.25	.16	-.34
								.79	.64	.22	.16	-.36
								.91	.69	.43	.13	-.33
								.89	.72	.35	.23	-.25
SM2									.73	.22	.11	-.37
									.62	.13	.01	-.35
									.79	.47	.14	-.37
									.75	.27	.23	-.24
B-F										.31	.16	-.39
										.37	.19	-.45
										.38	.02	-.33
										.26	.26	-.26
ORe											.62	-.29
											.68	-.04
											.15	-.47
											.55	-.21

	Full	Sht	I.Q.	Fac1	Fac2	Fac3	Fac4	FSh1	FSh2
Tok	.81	.73	.41	.39	.72	.14	.33	.34	.74
	.88	.83	.64	.14	.83	.25	.38	.37	.77
	.75	.64	.28	.31	.65	.05	.43	.25	.73
	.74	.58	.31	.47	.67	.09	-.10	.27	.66
SeP	.69	.62	.41	.25	.62	.14	.37	.27	.65
	.73	.69	.58	.06	.60	.24	.49	.25	.69
	.68	.57	.29	.18	.69	.08	.29	.24	.63
	.66	.53	.39	.35	.56	.11	.05	.28	.57
W&R	.55	.57	.15	.04	.56	.32	.22	.37	.48
	.70	.69	.48	-.03	.74	.36	.19	.41	.56
	.55	.48	-.07	.09	.50	.17	.36	.23	.49
	.47	.56	.04	-.10	.44	.41	.16	.44	.39
ASp	.66	.69	.41	.09	.68	.26	.33	.30	.75
	.67	.72	.53	-.20	.74	.28	.47	.21	.79
	.55	.56	.28	.03	.69	.15	.08	.20	.68
	.64	.70	.44	.03	.62	.38	.00	.42	.63
CD	.80	.72	.41	.55	.63	.25	-.06	.49	.54
	.77	.69	.56	.33	.71	.27	.00	.46	.48
	.78	.71	.35	.59	.55	.28	-.33	.53	.41
	.78	.67	.37	.50	.65	.23	-.44	.43	.57
MVT	.88	.84	.42	.52	.62	.33	.24	.53	.67
	.89	.84	.60	.26	.71	.27	.42	.41	.73
	.85	.84	.34	.51	.52	.47	-.07	.68	.40
	.83	.76	.34	.44	.67	.33	-.26	.49	.64
Ari	.82	.72	.59	.50	.67	.14	.20	.38	.68
	.82	.72	.65	.21	.79	.15	.30	.30	.70
	.83	.71	.53	.56	.53	.16	.21	.41	.58
	.74	.60	.60	.50	.68	.11	-.37	.32	.62
SM1	.64	.77	.35	.19	.22	.88	.06	.91	.09
	.61	.73	.46	-.05	.39	.86	.06	.88	.14
	.69	.78	.31	.43	.01	.86	.03	.92	-.07
	.70	.84	.27	.06	.15	.91	.09	.94	.01
SM2	.65	.76	.38	.26	.14	.89	.05	.93	.06
	.57	.69	.45	.00	.33	.86	-.01	.90	.08
	.74	.84	.35	.44	.06	.90	.03	.94	.00
	.65	.79	.35	.15	-.03	.93	.12	.94	-.11
B-F	.65	.76	.42	.26	.20	.80	.10	.82	.21
	.63	.72	.49	.07	.28	.77	.22	.69	.33
	.65	.78	.40	.37	.09	.84	-.08	.86	.05
	.70	.79	.36	.19	.13	.82	.11	.87	.07
ORe	.58	.55	.40	.32	.25	.12	.70	.18	.63
	.57	.57	.53	.01	.30	.14	.83	.06	.70
	.63	.58	.51	.61	.04	.27	.30	.46	.28
	.65	.55	.36	.40	.52	.16	-.05	.28	.57

	Age	Full	Sht	I.Q.	Fac1	Fac2	Fac3	Fac4	FSh1	FSh2
NA	-.13	.49	.45	.22	.12	.37	-.02	.78	.02	.68
	.20	.50	.48	.44	-.16	.41	-.02	.86	-.10	.74
	-.15	.42	.36	.00	.00	.47	-.11	.68	-.04	.68
	-.07	.47	.35	.08	.42	.21	.13	.14	.25	.24
Age		-.55	-.50	-.24	-.63	-.04	-.24	-.07	-.41	-.26
		-.25	-.28	-.17	-.37	.13	-.41	.13	-.50	.11
		-.52	-.44	-.28	-.68	.11	-.21	-.16	-.37	-.17
		-.40	-.39	-.18	-.38	-.11	-.19	.13	-.29	-.23
Full			.95	.58	.64	.57	.43	.27	.66	.69
			.95	.70	.35	.72	.41	.37	.56	.75
			.95	.56	.68	.47	.47	.15	.70	.56
			.93	.58	.58	.54	.54	-.10	.72	.58
Sht				.55	.46	.54	.63	.27	.78	.62
				.69	.16	.69	.60	.40	.68	.70
				.54	.56	.40	.65	.11	.82	.46
				.47	.32	.48	.75	-.04	.86	.46
I.Q.					.38	.30	.26	.11	.40	.39
					.08	.57	.36	.30	.44	.55
					.54	.06	.22	.10	.41	.31
					.47	.26	.21	-.17	.32	.33
Fac1						-.00	.00	-.00	.31	.32
						-.11	-.20	-.20	.14	.05
						-.12	.23	-.04	.51	.15
						.06	-.08	-.14	.13	.36
Fac2							.00	.00	.13	.72
							.17	.14	.27	.70
							-.16	-.05	-.04	.78
							-.07	-.47	.08	.83
Fac3								-.00	.92	-.14
								.02	.88	-.03
								-.12	.93	-.30
								.19	.96	-.21
Fac4									-.07	.51
									-.15	.68
									-.13	.38
									.09	-.22
FSh1										.00
										-.04
										-.12
										-.05

Short Testing

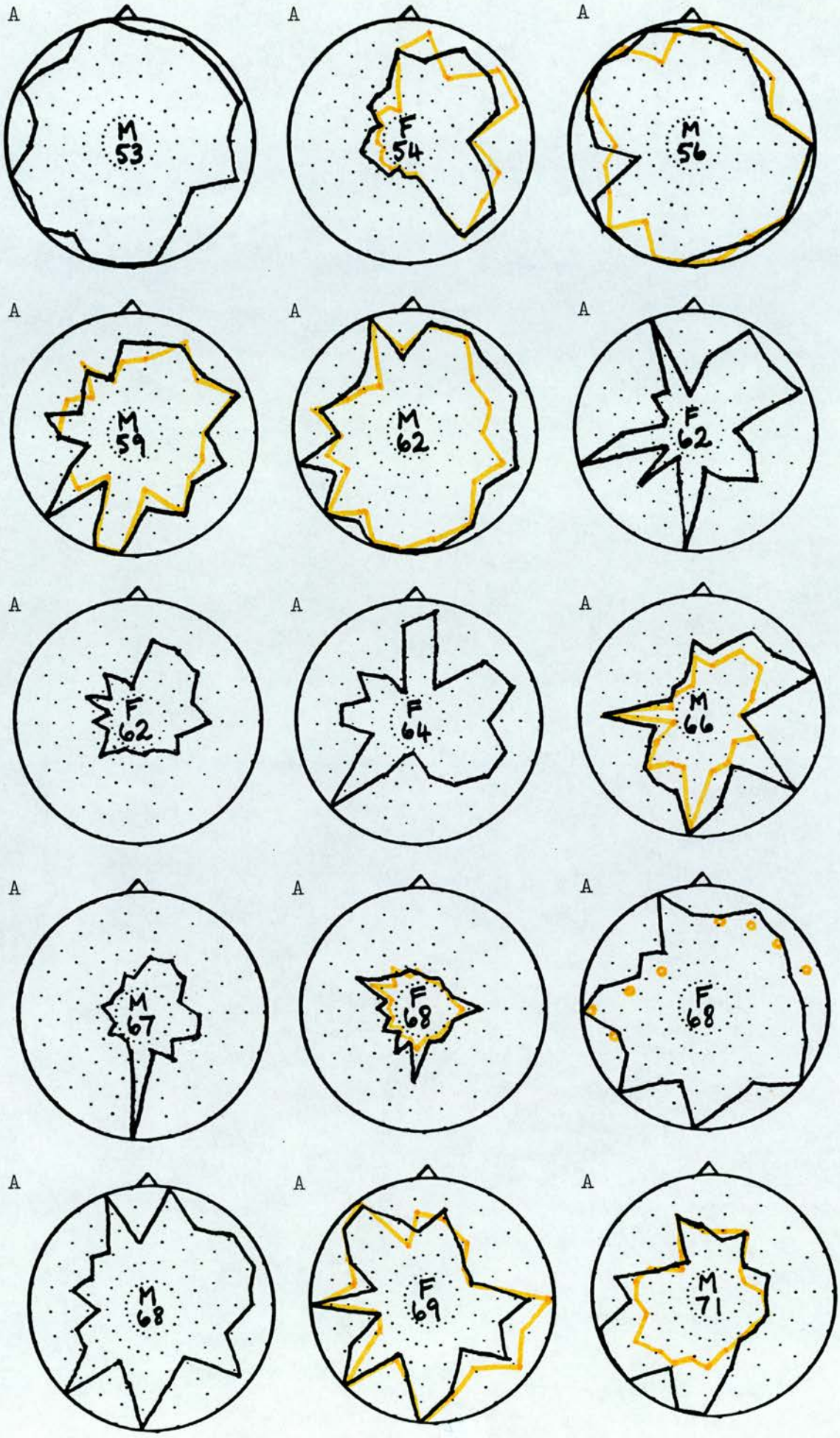
	Dig	SeR	ASp	SKP	W/R	SM1	SM2	B-F	PM	SS	Age	Sht
Ori	.57	.73	.64	.53	.57	.54	.56	.59	.73	.61	-.58	.81
	.53	.68	.59	.51	.52	.51	.55	.58	.69	.52	-.26	.76
	.48	.66	.57	.52	.60	.57	.57	.54	.64	.57	-.53	.79
	.80	.83	.77	.56	.76	.57	.54	.68	.76	.79	-.60	.88
Dig		.80	.75	.54	.67	.52	.48	.57	.56	.63	-.26	.79
		.74	.71	.62	.70	.53	.52	.56	.48	.63	-.11	.79
		.77	.75	.55	.61	.42	.37	.47	.47	.58	-.10	.74
		.91	.83	.53	.75	.58	.52	.64	.73	.72	-.46	.88
SeR			.80	.63	.72	.54	.51	.60	.65	.70	-.35	.86
			.74	.68	.75	.56	.53	.58	.62	.75	-.15	.87
			.77	.61	.67	.44	.44	.51	.55	.58	-.20	.81
			.88	.56	.79	.59	.53	.66	.72	.76	-.48	.90
ASp				.64	.72	.55	.53	.62	.59	.69	-.25	.85
				.69	.70	.51	.49	.60	.44	.66	-.01	.81
				.64	.68	.45	.47	.57	.58	.63	-.08	.82
				.54	.80	.65	.60	.67	.70	.79	-.45	.90
SKP					.68	.35	.32	.43	.39	.57	-.15	.67
					.71	.40	.35	.54	.35	.63	-.04	.74
					.75	.31	.28	.38	.38	.52	.02	.67
					.58	.31	.31	.38	.42	.51	-.29	.60
W/R						.54	.50	.56	.50	.70	-.16	.80
						.56	.57	.55	.45	.73	-.10	.83
						.52	.47	.53	.52	.62	-.07	.80
						.57	.50	.66	.60	.78	-.33	.84
SM1							.90	.78	.67	.57	-.39	.80
							.84	.74	.63	.62	-.32	.79
							.93	.78	.74	.56	-.35	.79
							.91	.80	.66	.53	-.42	.81
SM2								.80	.67	.55	-.42	.78
								.73	.65	.59	-.34	.78
								.85	.74	.55	-.39	.78
								.81	.62	.49	-.40	.77
B-F									.70	.61	-.45	.82
									.62	.57	-.44	.80
									.78	.60	-.36	.81
									.70	.63	-.45	.85
PM										.61	-.56	.81
										.56	-.42	.74
										.62	-.39	.82
										.63	-.64	.84

	I.Q.	FSh1	FSh2
Ori	.37	.53	.74
	.39	.49	.69
	.38	.50	.69
	.41	.59	.85
Dig	.39	.43	.82
	.53	.42	.81
	.26	.28	.82
	.38	.55	.90
SeR	.46	.45	.90
	.57	.44	.89
	.41	.31	.88
	.40	.55	.92
ASp	.41	.46	.86
	.53	.38	.86
	.25	.36	.86
	.44	.62	.87
SKP	.10	.15	.85
	.27	.16	.89
	-.18	.10	.86
	.02	.16	.77
W/R	.08	.48	.81
	.37	.49	.83
	.01	.41	.81
	.00	.58	.84
SM1	.34	.91	.44
	.46	.89	.47
	.30	.91	.38
	.27	.93	.45
SM2	.38	.93	.41
	.45	.91	.45
	.34	.94	.35
	.35	.93	.40
B-F	.41	.84	.54
	.49	.76	.58
	.39	.86	.47
	.36	.89	.56
PM	.47	.73	.59
	.44	.71	.51
	.55	.78	.53
	.43	.70	.68

	Age	Sht	I.Q.	FSh1	FSh2
SS	-.32	.82	.31	.59	.74
	-.15	.84	.51	.58	.75
	-.17	.80	.31	.58	.66
	-.51	.83	.02	.58	.79
Age		-.45	-.23	-.45	-.28
		-.26	-.17	-.43	-.06
		-.31	-.27	-.40	-.10
		-.56	-.18	-.45	-.46
Sht			.54	.75	.86
			.69	.71	.88
			.51	.71	.84
			.47	.80	.86
I.Q.				.39	.39
				.44	.55
				.40	.29
				.32	.33
FSh1					.32
					.30
					.21
					.39

APPENDIX 4

Details of factor analyses.


A
M 53
A
F 54
A
M 56
A
M 59
A
M 62
A
F 62
A
F 62
A
F 64
A
M 66
A
M 67
A
F 68
A
F 68
A
M 68
A
F 69
A
M 71

Full Testing

	All (n=174)				DAT (n=58)			
	Fac1	Fac2	Fac3	Fac4	Fac1	Fac2	Fac3	Fac4
Ori	.717	.293	.132	.444	.347	.400	.688	.197
Par	.754	.119	.217	.382	.074	.332	.755	.120
MfD	.747	.415	.278	-.023	.568	.077	.525	.352
YNP	.704	.224	.121	.160	.366	.080	.681	.125
F-N	.792	.054	.174	.304	.135	.097	.871	-.062
PM	.618	.349	.447	.027	.363	.080	.550	.495
BD	.607	.475	.429	-.039	.663	.091	.430	.290
Wei	.599	.499	.320	.019	.731	.036	.316	.225
Dig	.172	.732	.108	.065	.550	.462	.013	.188
SeR	.338	.702	.052	.383	.558	.558	.350	.071
Tok	.391	.719	.136	.329	.690	.529	.241	.234
SeP	.250	.622	.141	.366	.363	.588	.245	.269
W&R	.037	.556	.324	.222	.668	.297	.083	.264
ASp	.090	.684	.261	.333	.509	.618	-.001	.180
CD	.549	.627	.247	-.058	.762	.111	.251	.268
MVT	.517	.621	.325	.238	.600	.510	.373	.261
Ari	.503	.666	.139	.204	.738	.426	.251	.110
SM1	.188	.216	.879	.060	.343	.100	-.000	.837
SM2	.265	.139	.891	.046	.262	.041	.061	.878
B-F	.257	.198	.801	.103	.129	.244	.247	.767
ORe	.318	.246	.123	.700	.027	.807	.271	.159
NA	.116	.373	-.017	.783	.145	.858	.113	-.069
% Var	24.3	23.4	15.0	10.2	24.3	17.5	17.2	14.0
Total % Var				73.0				73.0

	MID (n=58)				Other (n=58)			
	Fac1	Fac2	Fac3	Fac4	Fac1	Fac2	Fac3	Fac4
Ori	.393	.147	.789	.132	.326	.672	.315	.166
Par	.320	.344	.732	.153	.261	.805	.045	.250
MfD	.206	.360	.283	.760	.484	.605	.210	.201
YNP	-.041	.106	.614	.537	.402	.680	.273	.068
F-N	.182	.291	.778	.054	.088	.792	.134	.301
PM	.098	.627	.252	.506	.543	.264	.351	.414
BD	.178	.532	.205	.677	.693	.236	.177	.491
Wei	.436	.383	.271	.443	.600	.332	.246	.394
Dig	.627	-.015	-.096	.274	.545	.193	.569	.070
SeR	.828	-.009	.232	.140	.426	.327	.687	.048
Tok	.800	.145	.383	.170	.775	.287	.336	.025
SeP	.756	.159	.175	.234	.211	.498	.630	-.001
W&R	.614	.195	.247	.048	-.003	.096	.602	.363
ASp	.673	.221	-.151	.319	.240	.065	.715	.321
CD	.323	.283	.179	.784	.615	.275	.450	.206
MVT	.420	.490	.235	.581	.748	.207	.392	.292
Ari	.550	.197	.476	.440	.622	.282	.482	.058
SM1	.077	.857	.288	.159	.204	.114	.216	.875
SM2	.122	.887	.260	.217	.149	.191	.046	.911
B-F	.097	.841	.096	.264	.146	.301	.203	.781
ORe	.077	.183	.701	.360	.810	.129	.198	.107
NA	.789	-.002	.275	-.231	.730	.278	-.266	.120
% Var	22.4	17.9	17.1	16.2	25.0	16.9	15.5	15.4
Total % Var				73.7				72.8

Short Testing (in subjects completing Full testing)

All (n=174)

	F1	F2
Ori	.362	.682
Dig	.247	.669
SeR	.233	.814
ASp	.297	.750
SKP	-.124	.636
W/R	.325	.317
SM1	.912	.088
SM2	.928	.057
B-F	.816	.213
PM	.665	.440
SS	.508	.361
% Var	31.9	27.5
Total % Var		59.4

DAT (n=58)

	F1	F2	F3
Ori	.400	.483	.456
Dig	.261	.676	.311
SeR	.318	.524	.552
ASp	.219	.656	.500
SKP	-.040	.037	.876
W/R	.092	.910	-.184
SM1	.865	.214	-.061
SM2	.877	.210	-.152
B-F	.754	.089	.301
PM	.700	.248	.202
SS	.618	.093	.426
% Var	30.4	21.8	18.2
Total % Var			70.4

MID (n=58)

	F1	F2	F3
Ori	.375	.440	.309
Dig	.075	.845	-.148
SeR	.053	.782	.383
ASp	.238	.702	.141
SKP	-.105	.461	.710
W/R	.285	-.021	.839
SM1	.875	-.019	.225
SM2	.903	.060	.227
B-F	.857	.145	.078
PM	.829	.221	.008
SS	.647	.151	-.058
%Var	33.8	21.1	14.6
Total % Var			69.5

Other (n=58)

	F1	F2	F3	F4
Ori	.588	.206	.289	.336
Dig	.897	.136	-.023	-.042
SeR	.897	.117	.118	-.060
ASp	.615	.287	.317	.126
SKP	.043	-.034	-.026	.954
W/R	.348	.183	.718	.046
SM1	.232	.896	.173	-.030
SM2	.116	.955	.066	.023
B-F	.240	.805	.256	.011
PM	.609	.471	.090	.349
SS	-.006	.148	.851	-.037
%Var	26.8	25.4	14.1	10.6
Total % Var				77.0

Short Testing (in subjects completing at least Short testing)

	All (n=215)		DAT (n=74)	
	F1	F2	F1	F2
Ori	.615	.514	.492	.589
Dig	.769	.346	.765	.353
SeR	.828	.378	.797	.421
ASp	.806	.376	.816	.315
SKP	.829	.090	.864	.147
W/R	.803	.315	.805	.346
SM1	.263	.894	.292	.851
SM2	.216	.922	.250	.881
B-F	.377	.815	.406	.752
PM	.427	.729	.271	.797
SS	.700	.448	.702	.461
%Var	41.5	34.7	40.0	34.6
Total % Var		76.2		74.6

	MID (n=74)		Other (n=67)	
	F1	F2	F1	F2
Ori	.580	.518	.813	.405
Dig	.807	.226	.819	.392
SeR	.834	.294	.852	.387
ASp	.819	.326	.793	.463
SKP	.836	.101	.740	.045
W/R	.788	.334	.810	.357
SM1	.220	.912	.277	.915
SM2	.182	.943	.216	.931
B-F	.325	.852	.430	.808
PM	.374	.800	.587	.595
SS	.608	.506	.796	.349
%Var	39.7	36.1	47.1	33.1
Total % Var		75.9		80.2

APPENDIX 5

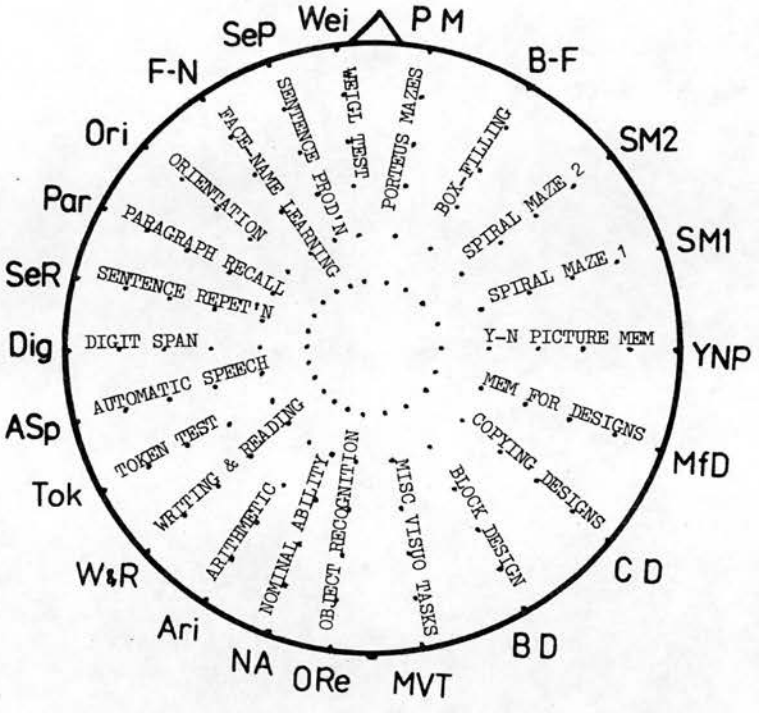
Individual subject graphs.

Abbreviated diagnostic labels:

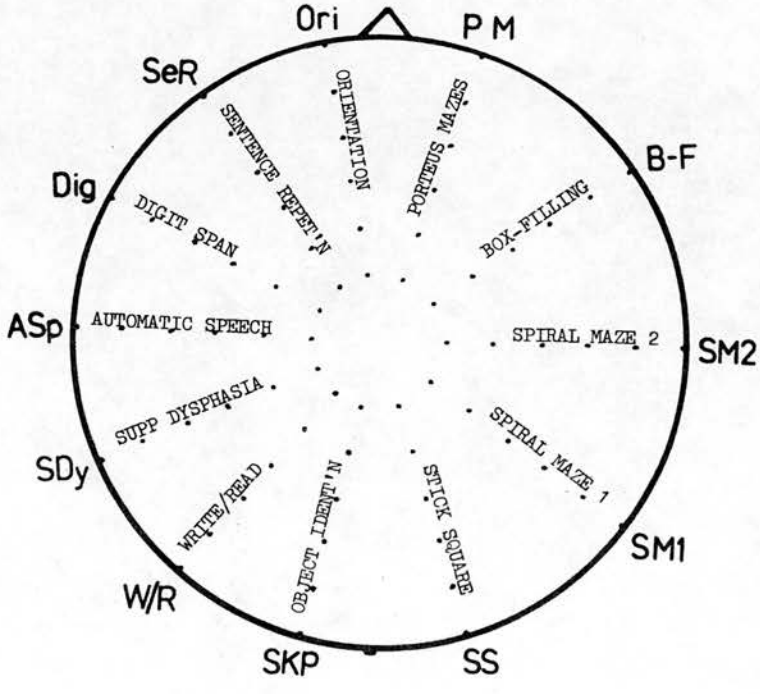
A	Dementia of the Alzheimer Type
MI	Multi-Infarct Dementia
A/MI	Suspected mixed Alzheimer/Multi-Infarct
Pick	Suspected Pick's Disease
HuCh	Huntington's Chorea
SeCh	Senile Chorea
Park	Parkinson's Disease
S-R	Steele-Richardson Syndrome
MS	Multiple Sclerosis
LOD	Late Onset Demyelination
Hydr	Hydrocephalus (* successfully treated by shunt surgery)
BIH	Benign Intracranial Hypertension
AlcD	Alcoholic Dementia
K/Alc	Mixed features of Korsakov's & Alcoholic Dementia
Pug	Dementia Pugilistica ('Punchy' syndrome)
GPI	General Paresis of the Insane (infection ?20 yrs previously)
SID	Steroid-Induced Dementia
BEnc	Binswanger Encephalopathy
EncL	Encephalitis Lethargica
CeAt	Cerebellar Ataxia of unknown origin
Tum	Brain Tumour (Respectively, Craniopharyngioma; Left frontal glioma; Right frontal glioma; Left temporal glioma)
Met	Metastatic spread to brain from extra-cranial malignancy
Stro	Single left fronto-parietal stroke
Hinj	Left parietal head injury
Hinj*	Frontal head injury with evidence of added dementing process
TLob	Left temporal lobectomy for chronic epilepsy
Depr	Depression
Fun	'Functional' condition (Hysterical or pseudo-dementia)
Schi	Chronic schizophrenia
MH	Mental Handicap
N	Normal with peripheral hand tremor
ndr	No diagnosis reached

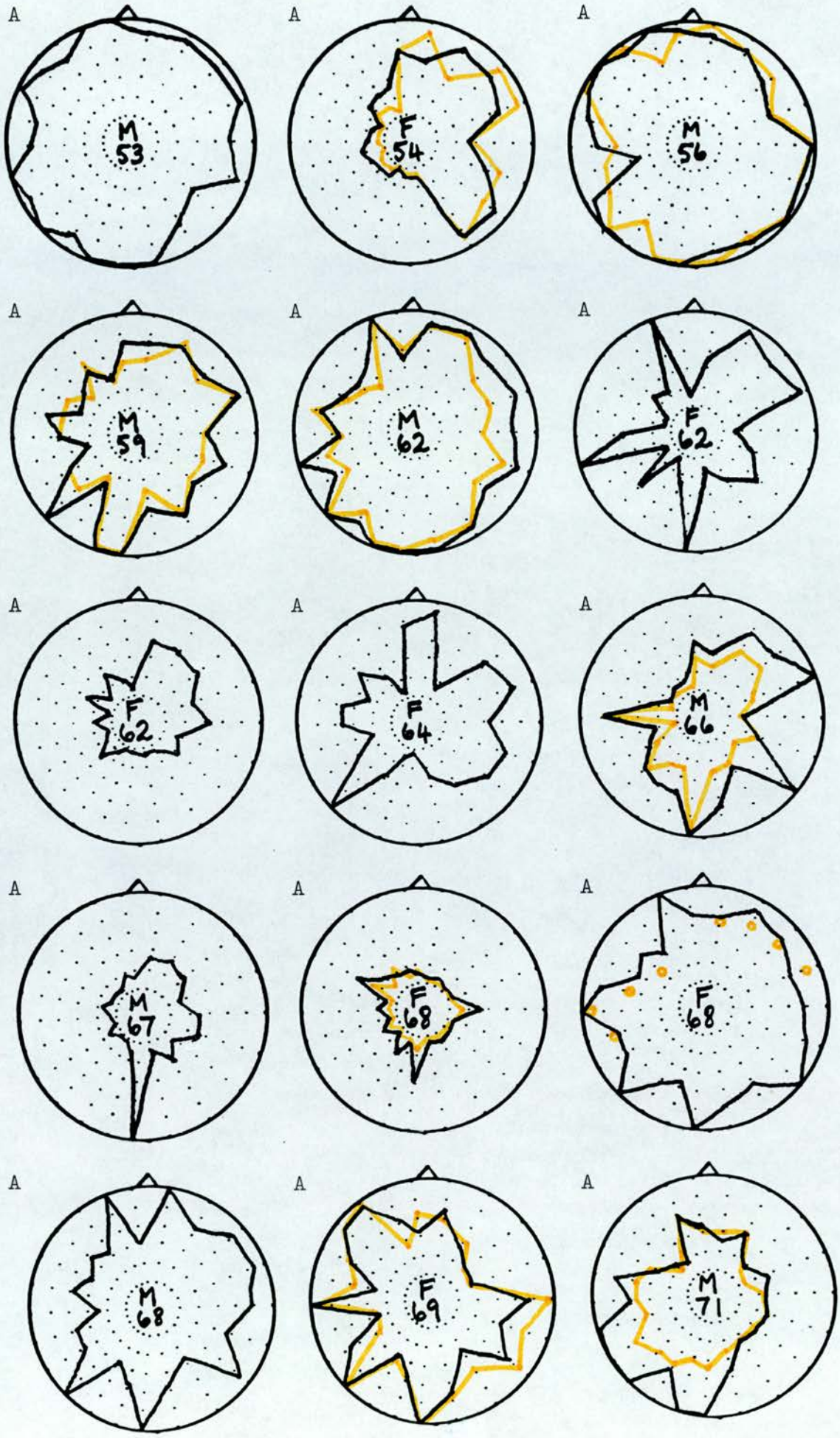
Graphs as described in Chapter 2.
Age shown at centre of each graph.
Test 1 scores in black; Test 2 (where available) in yellow or grey.

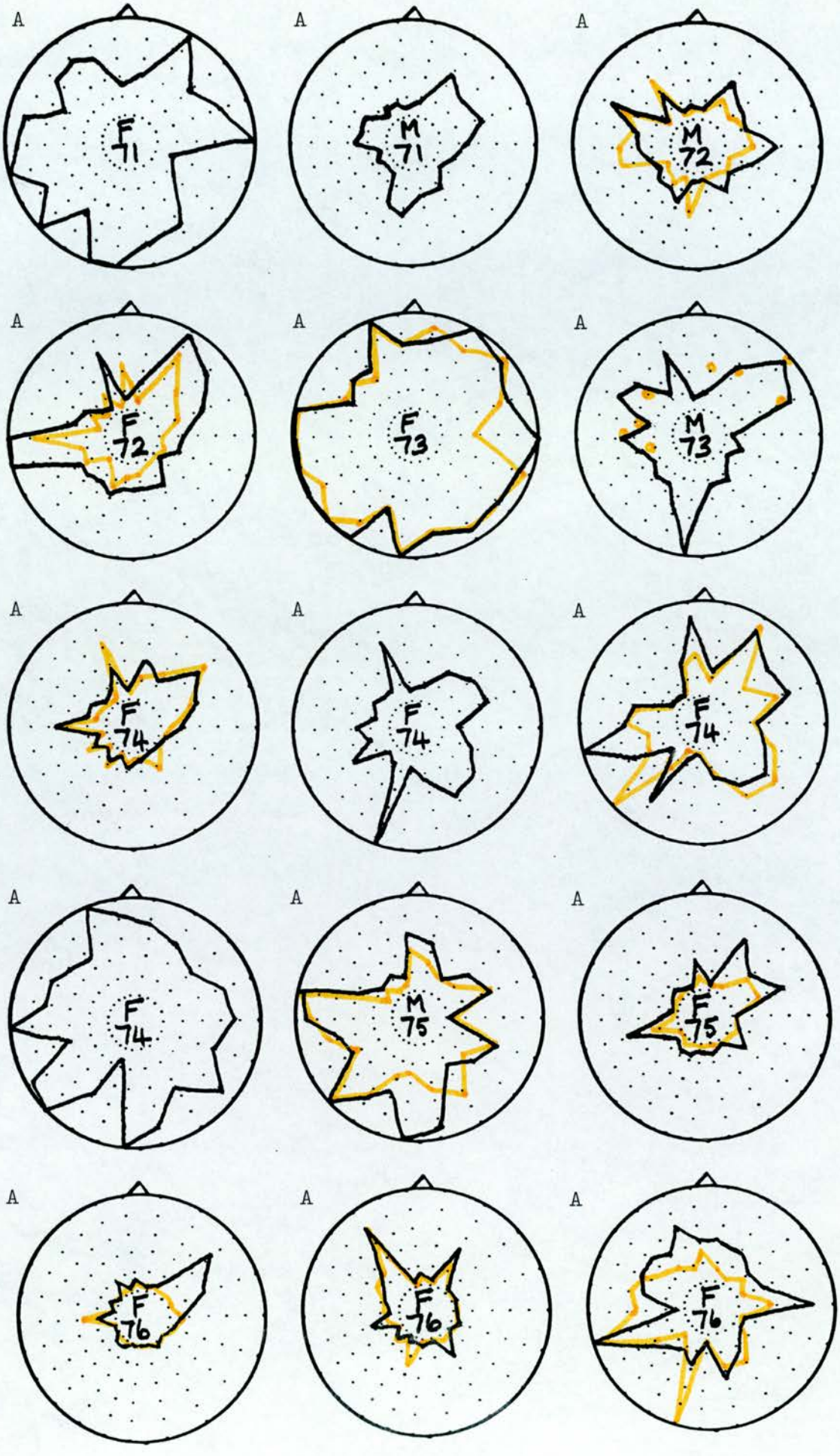
Key (Full testing):

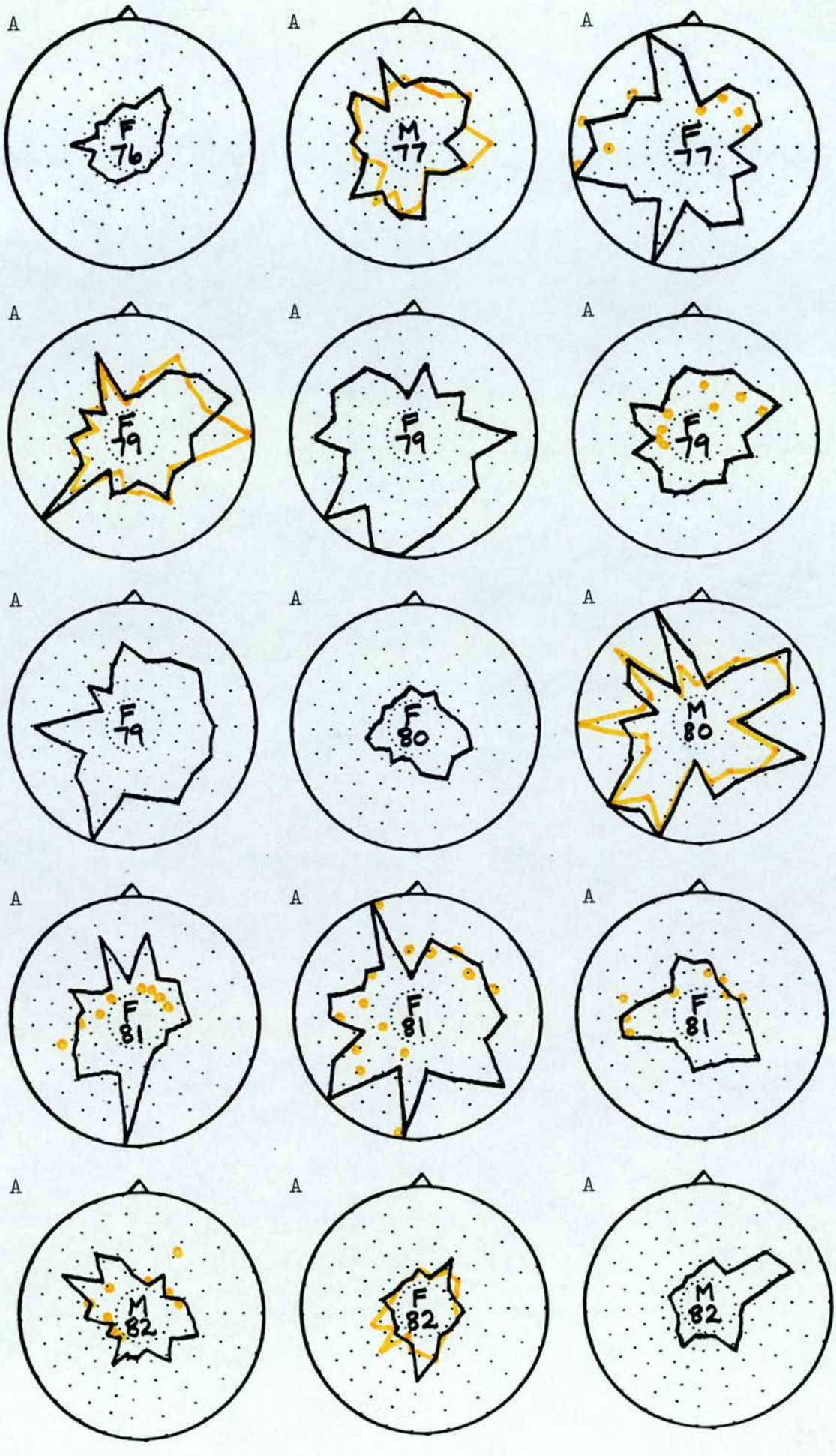


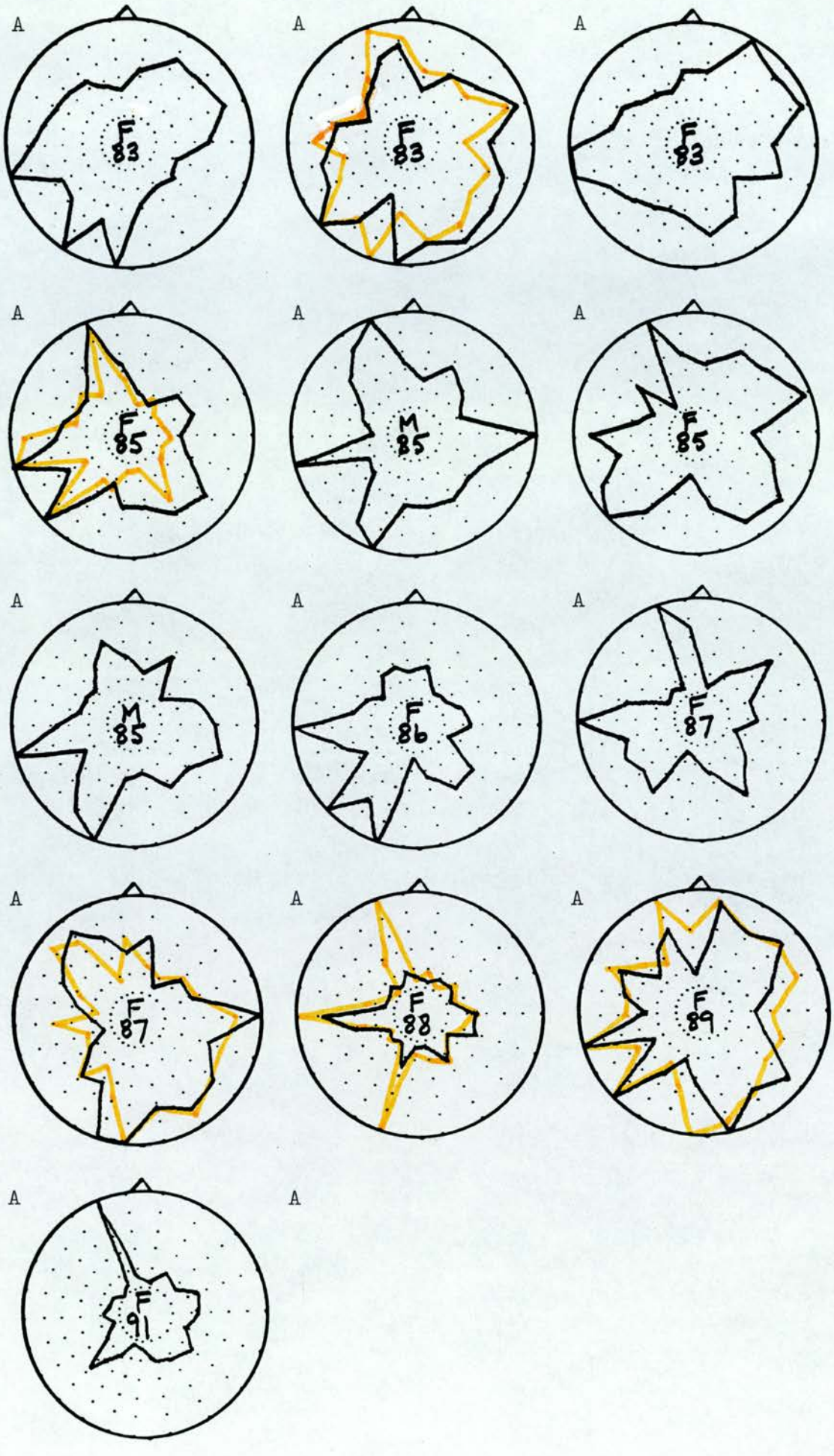
Key (Short testing):

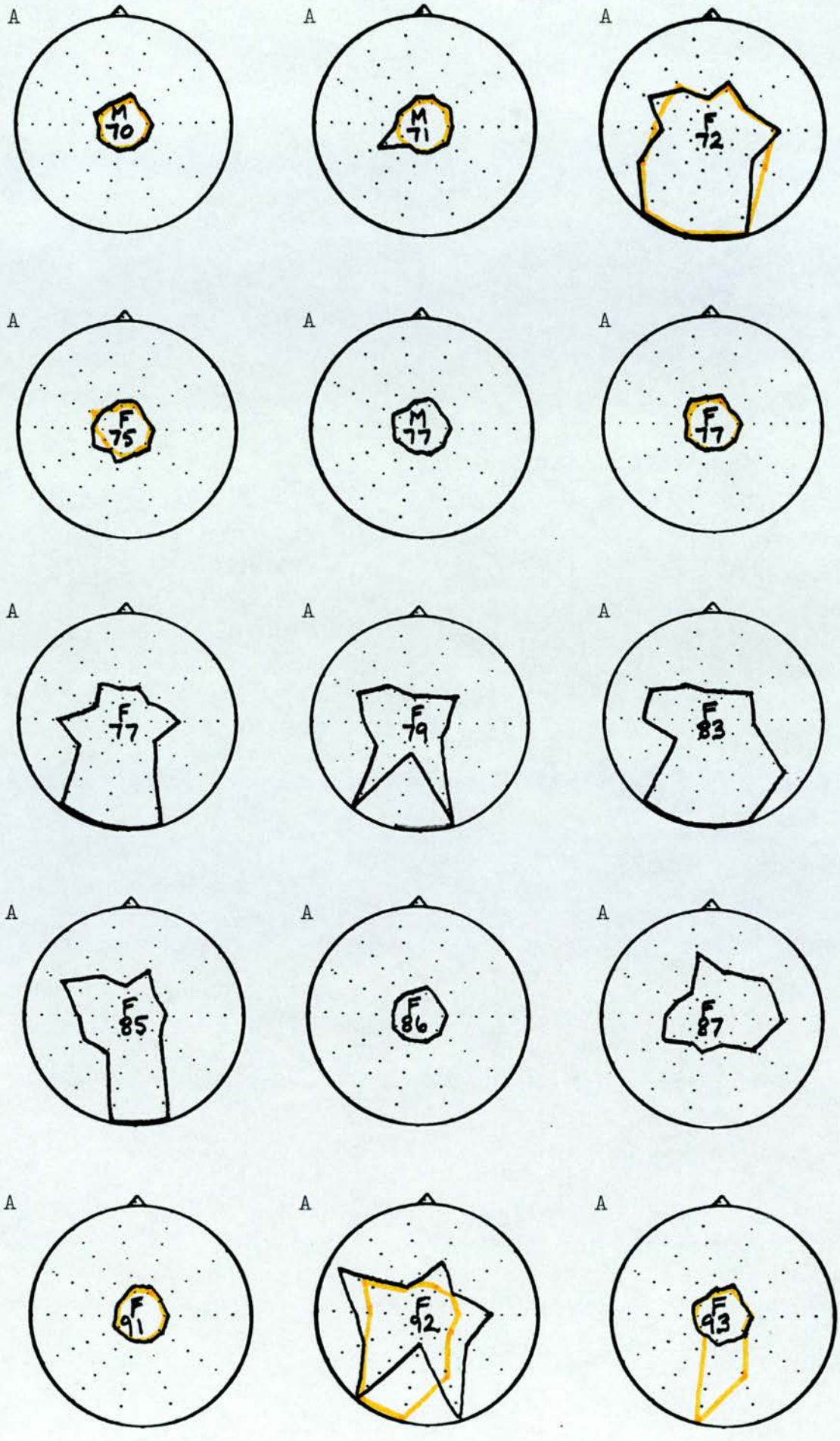



A
M 53
A
F 54
A
M 56
A
M 59
A
M 62
A
F 62
A
F 62
A
F 64
A
M 66
A
M 67
A
F 68
A
F 68
A
M 68
A
F 69
A
M 71

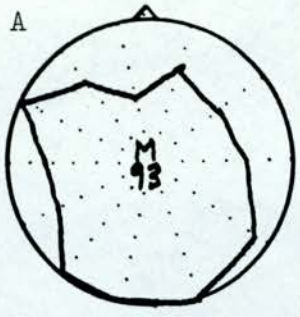

A
F 71
A
M 71
A
M 72
A
F 72
A
F 73
A
M 73
A
F 74
A
F 74
A
F 74
A
F 74
A
M 75
A
F 75
A
F 76
A
F 76
A
F 76

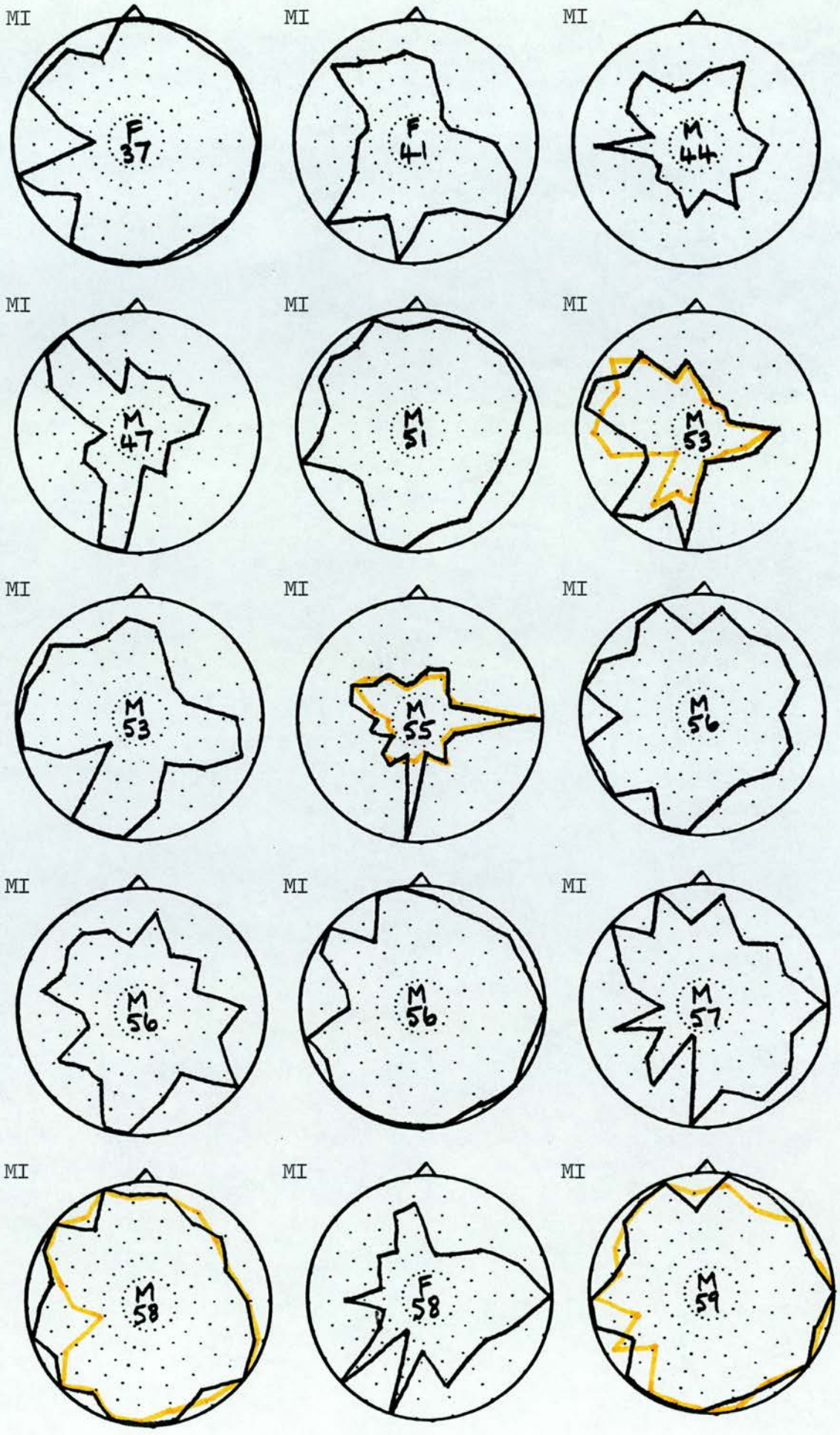

A
F 76
A
M 77
A
F 77
A
F 79
A
F 79
A
F 79
A
F 79
A
F 80
A
M 80
A
F 81
A
F 81
A
F 81
A
M 82
A
F 82
A
M 82

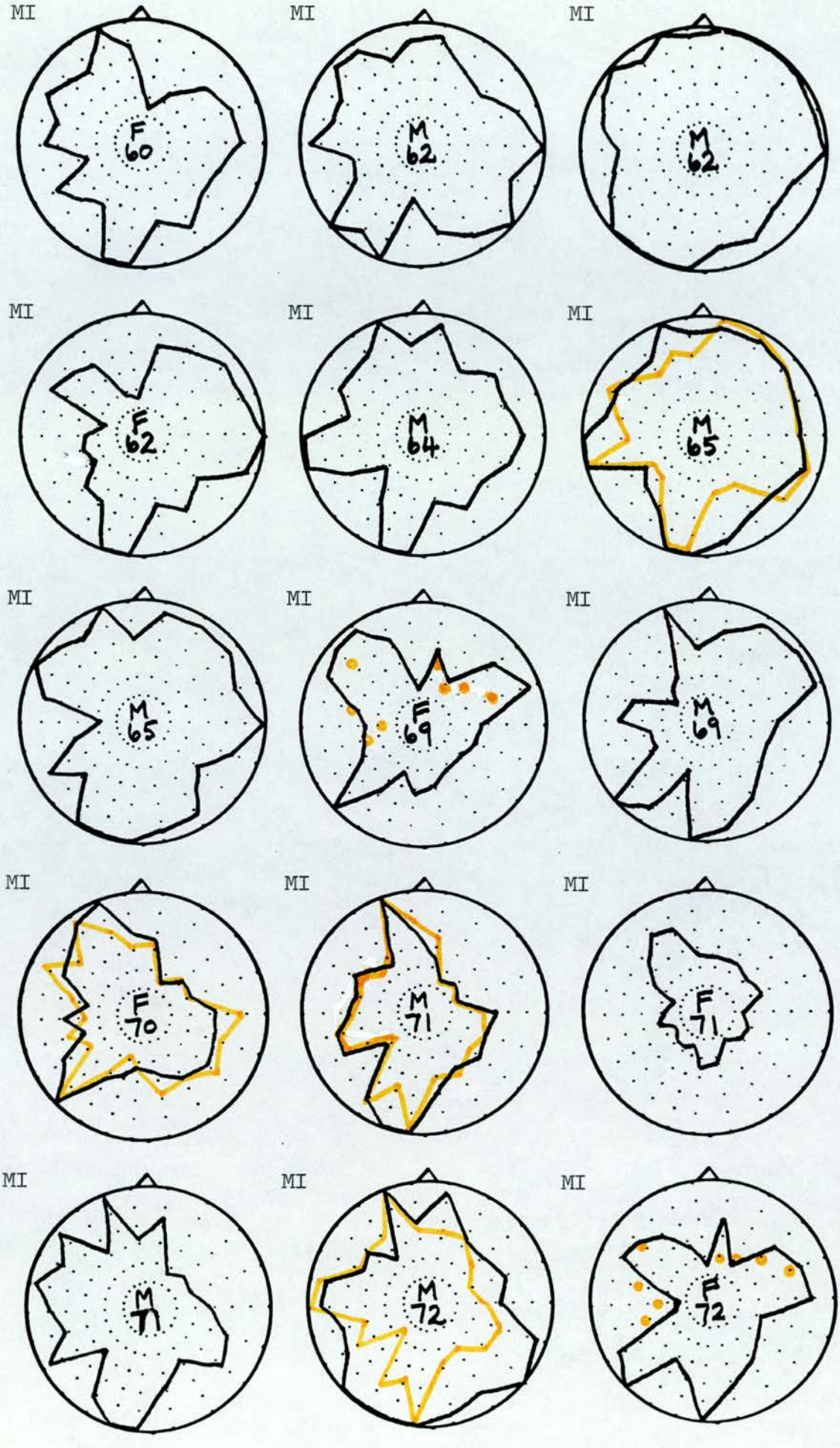

A
F 83
A
F 83
A
F 83
A
F 85
A
M 85
A
F 85
A
M 85
A
F 86
A
F 87
A
F 87
A
F 88
A
F 89
A
F 91
A

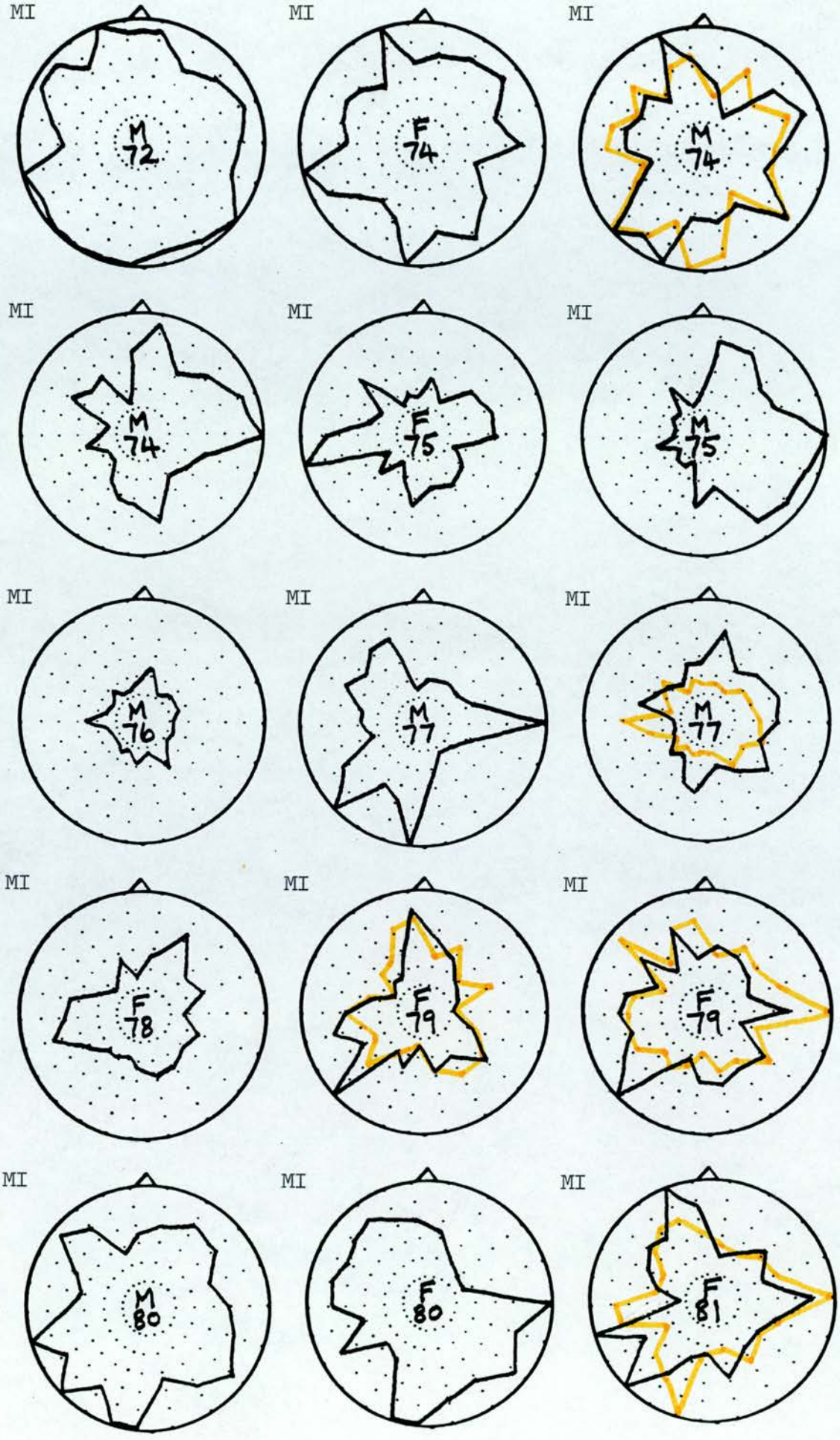

A
M 70
A
M 71
A
F 72
A
F 75
A
M 77
A
F 77
A
F 77
A
F 79
A
F 83
A
F 85
A
F 86
A
F 87
A
F 91
A
F 92
A
F 93

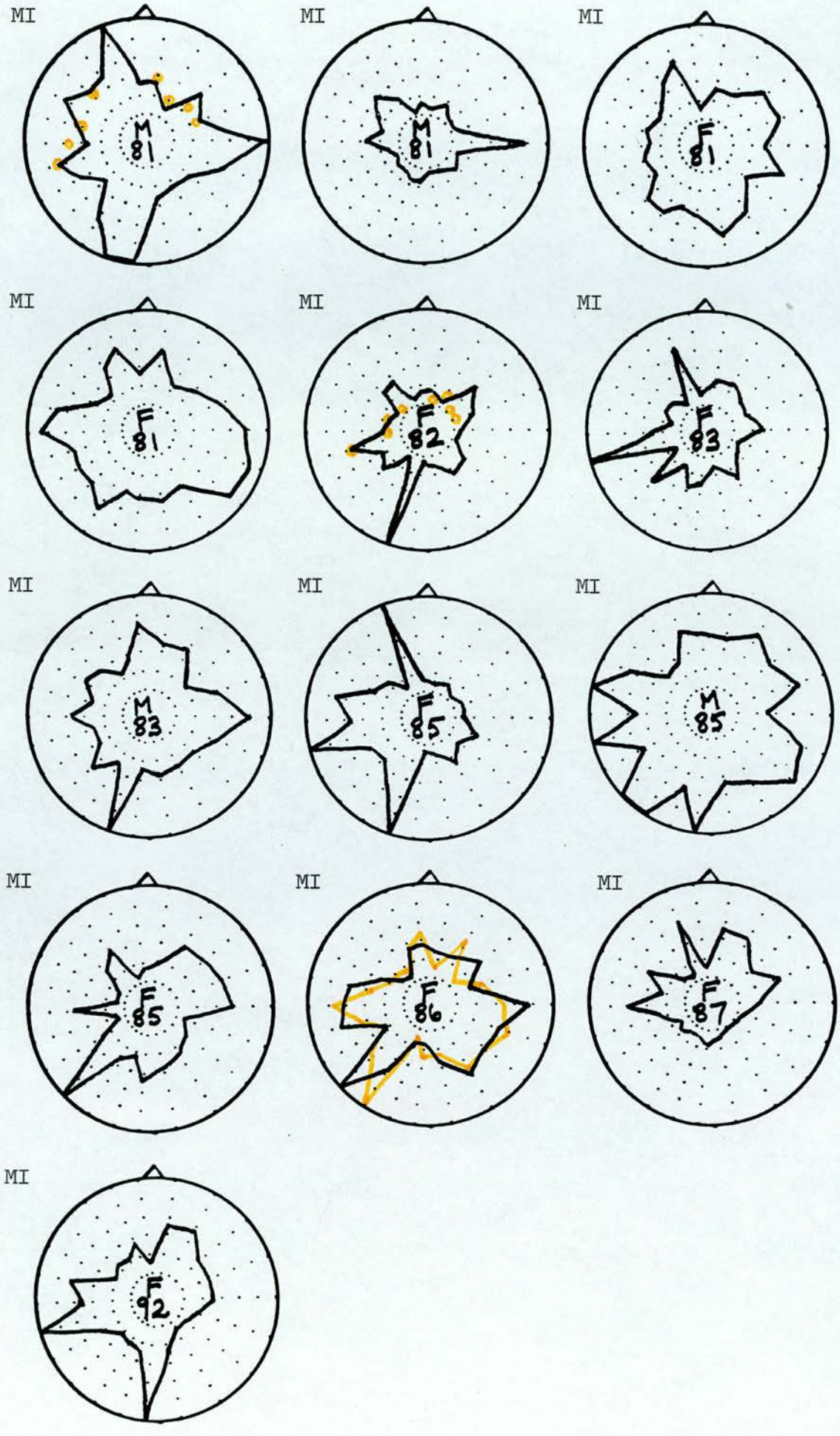
A

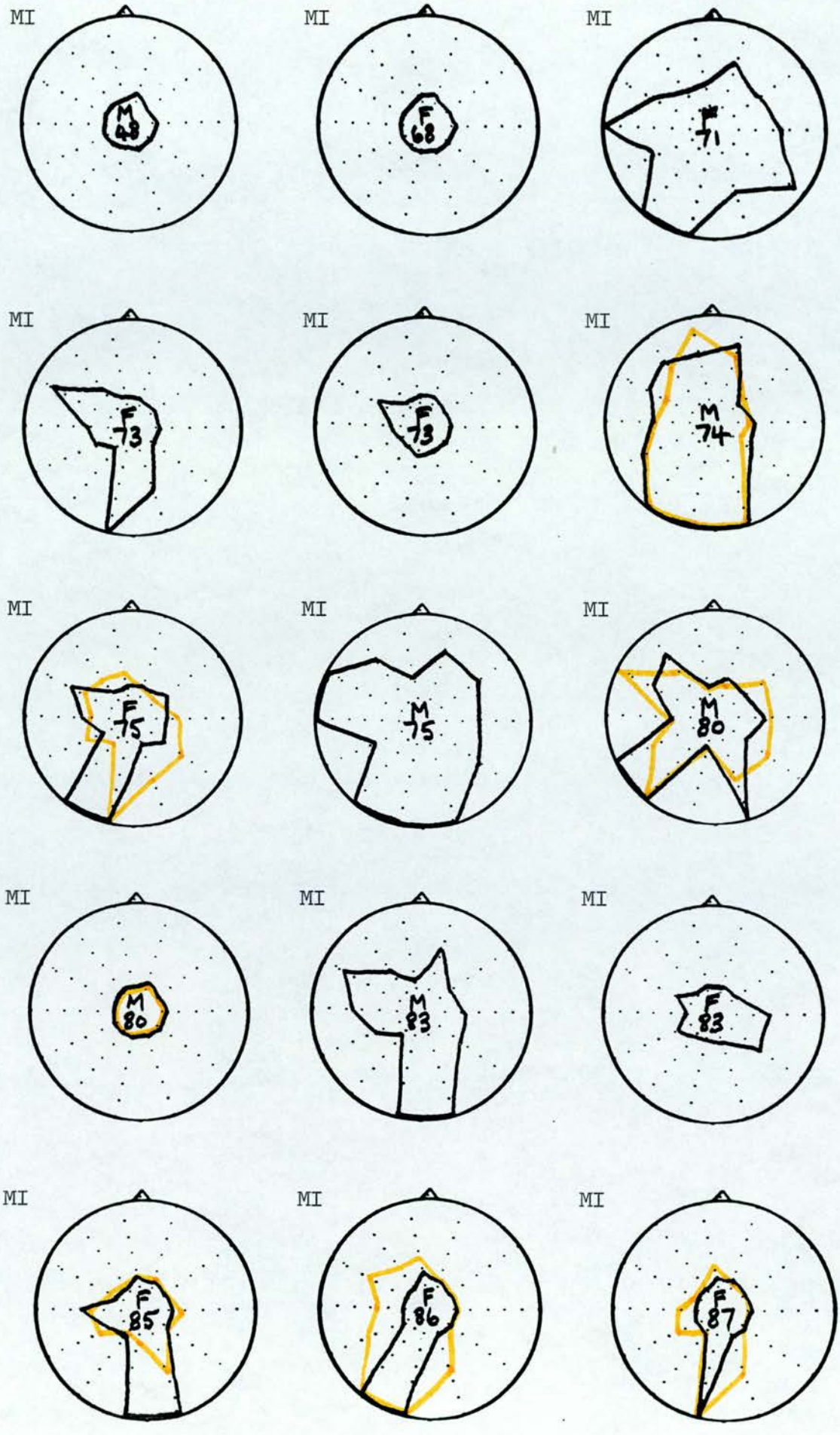



MI
F 37
MI
F 41
MI
M 44
MI
M 47
MI
M 51
MI
M 53
MI
M 53
MI
M 55
MI
M 56
MI
M 56
MI
M 56
MI
M 57
MI
M 58
MI
F 58
MI
M 59

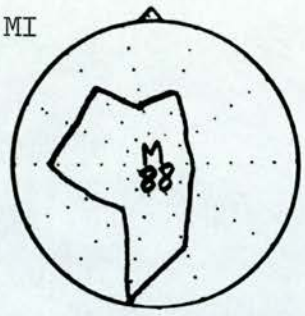

MI
F 60
MI
M 62
MI
M 62
MI
F 62
MI
M 64
MI
M 65
MI
M 65
MI
F 69
MI
M 69
MI
F 70
MI
M 71
MI
F 71
MI
M 71
MI
M 72
MI
F 72

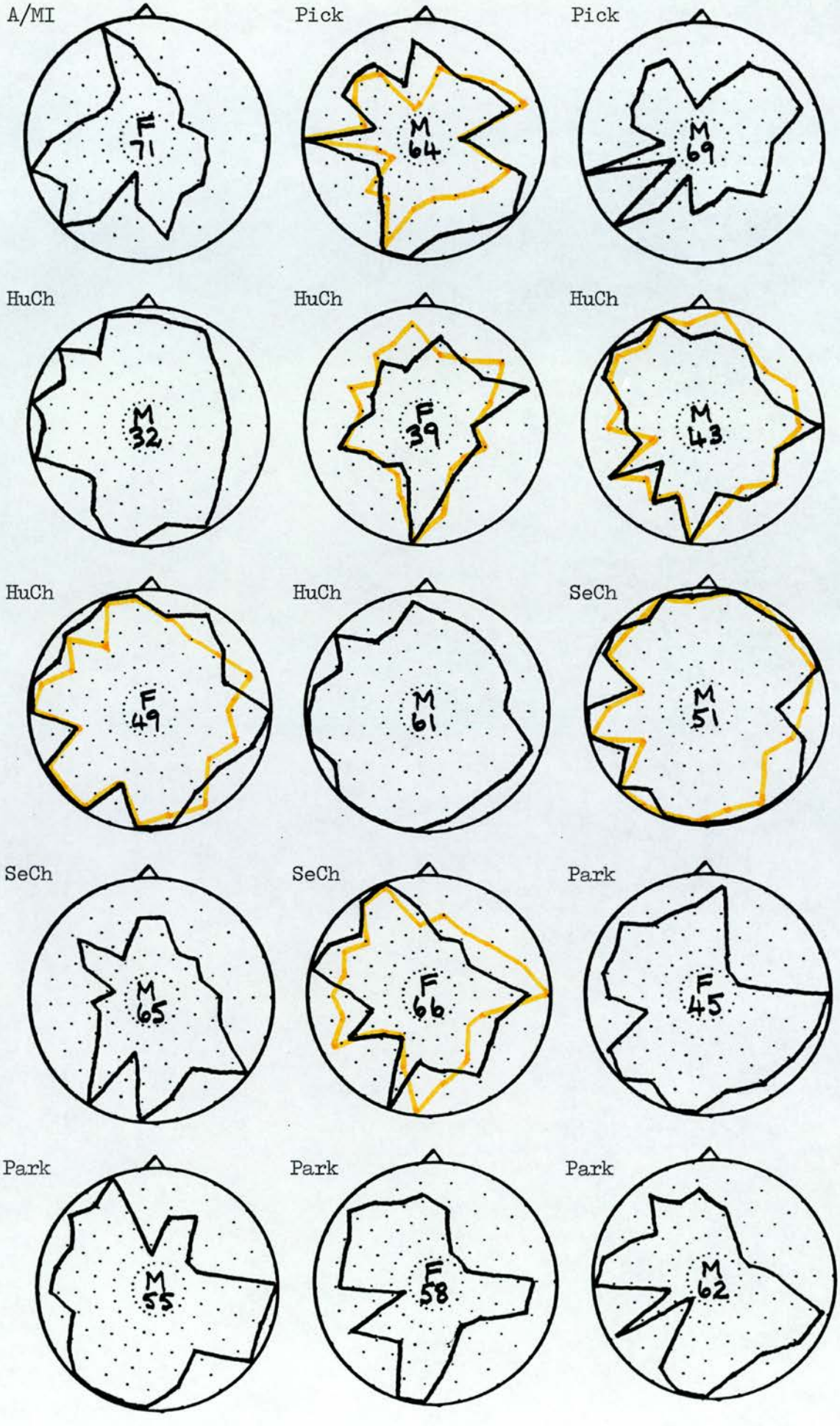

MI
M 72
MI
F 74
MI
M 74
MI
M 74
MI
F 75
MI
M 75
MI
M 76
MI
M 77
MI
M 77
MI
F 78
MI
F 79
MI
F 79
MI
M 80
MI
F 80
MI
F 81

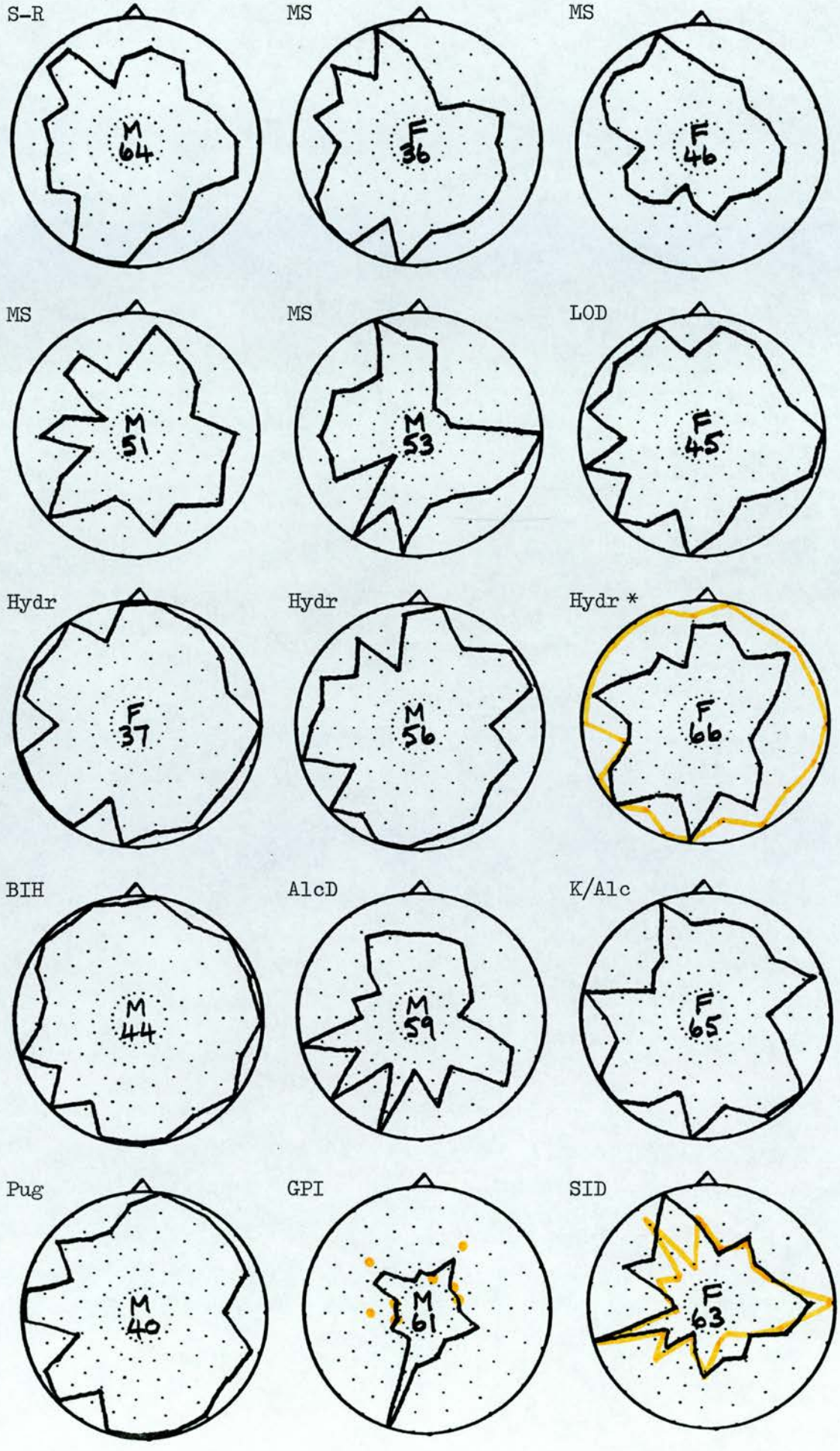

MI
M 81
MI
M 81
MI
F 81
MI
F 81
MI
F 82
MI
F 83
MI
M 83
MI
F 85
MI
M 85
MI
F 85
MI
F 86
MI
F 87
MI
F 92

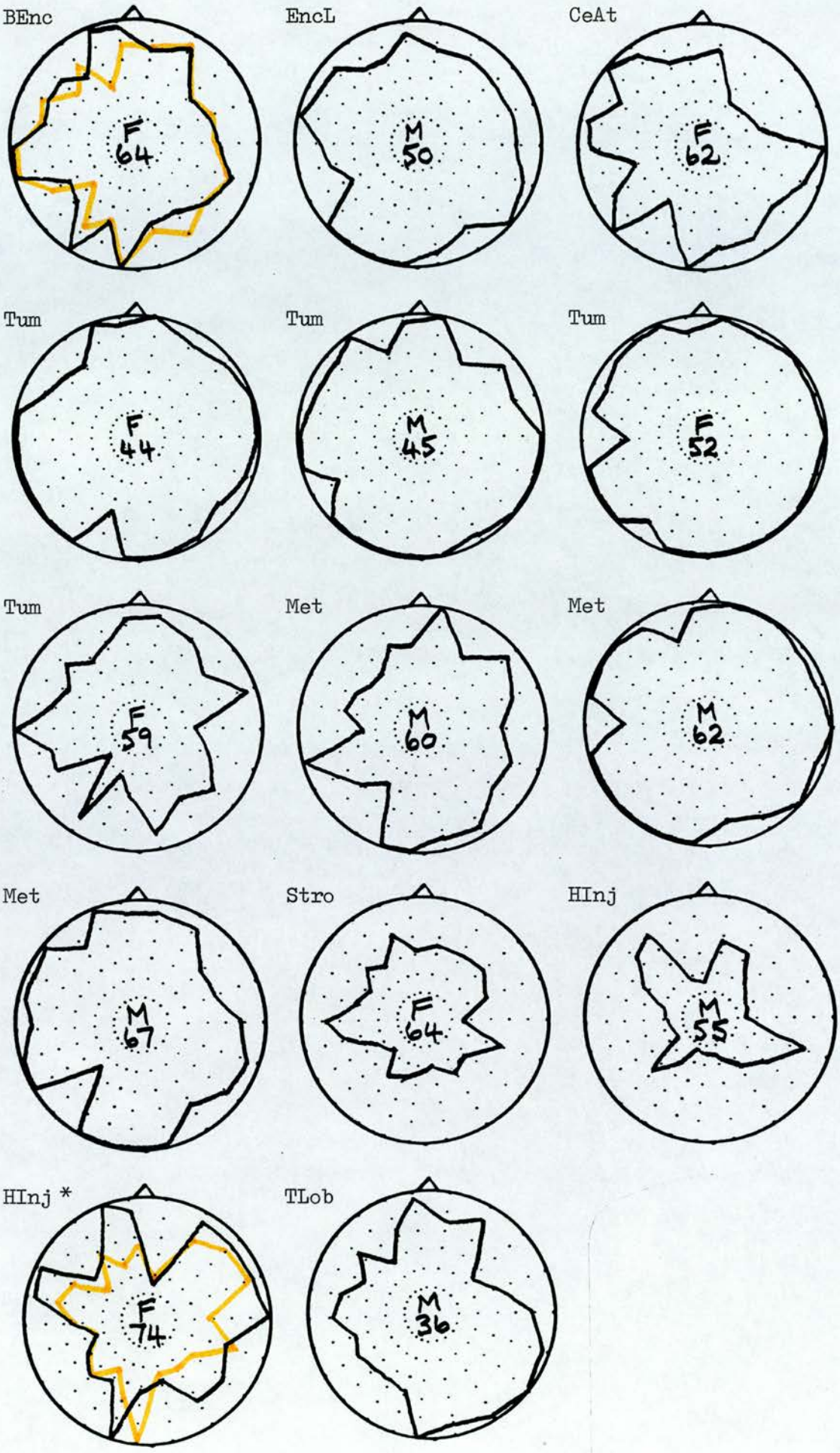

MI
M 48
MI
F 68
MI
F 71
MI
F 73
MI
F 73
MI
M 74
MI
F 75
MI
M 75
MI
M 80
MI
M 80
MI
M 83
MI
F 83
MI
F 85
MI
F 86
MI
F 87

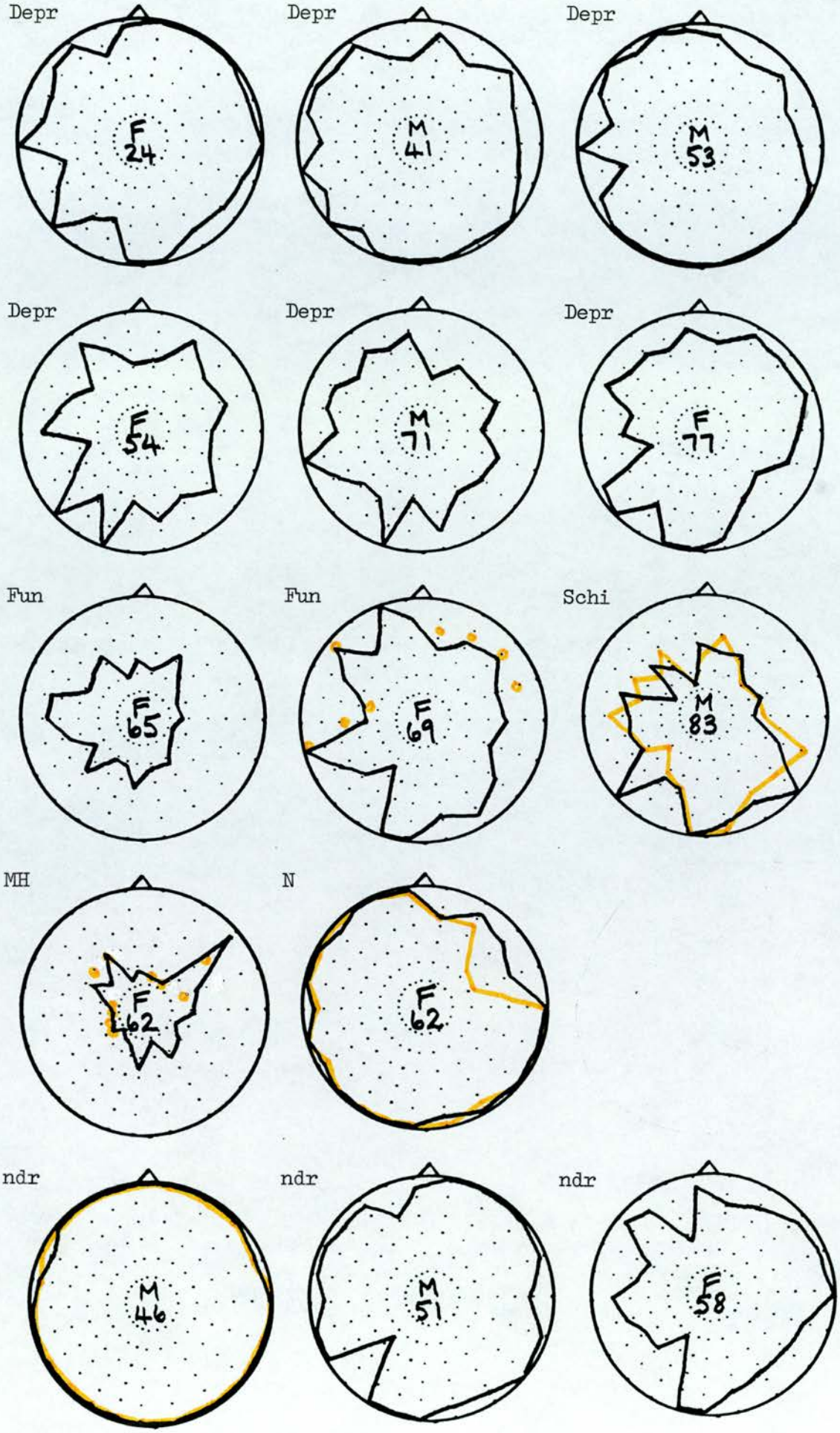
MI




A/MI
F
71
Pick
M
64
Pick
M
69
HuCh
M
32
HuCh
F
39
HuCh
M
43
HuCh
F
49
HuCh
M
61
SeCh
M
51
SeCh
M
65
SeCh
F
66
Park
F
45
Park
M
55
Park
F
58
Park
M
62

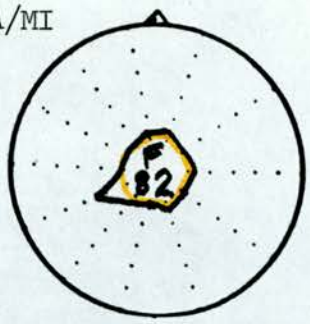

S-R
M 64
MS
F 36
MS
F 46
MS
M 51
MS
M 53
LOD
F 45
Hydr
F 37
Hydr
M 56
Hydr *
F 66
BIH
M 44
AlcD
M 59
K/Alc
F 65
Pug
M 40
GPI
M 61
SID
F 63


BEnc
F
64
EncL
M
50
CeAt
F
62
Tum
F
44
Tum
M
45
Tum
F
52
Tum
F
59
Met
M
60
Met
M
62
Met
M
67
Stro
F
64
HInj
M
55
HInj *
F
74
TLob
M
36

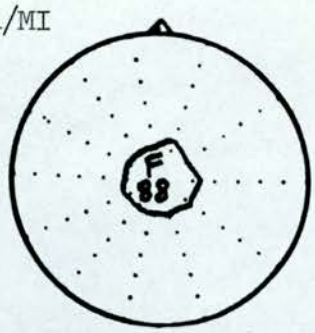

Depr
F 24
Depr
M 41
Depr
M 53
Depr
F 54
Depr
M 71
Depr
F 77
Fun
F 65
Fun
F 69
Schi
M 83
MH
F 62
N
F 62
ndr
M 46
ndr
M 51
ndr
F 58

A/MI

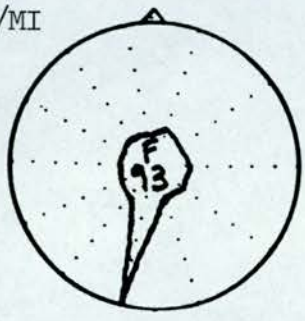
A/MI



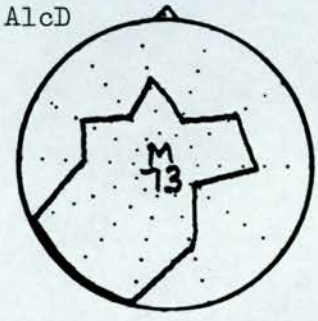
A/MI



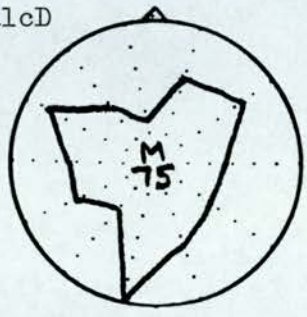
A/MI



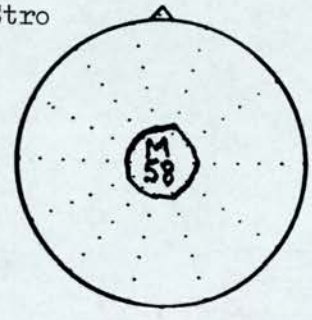
AlcD



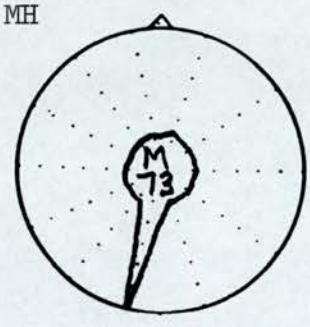
AlcD



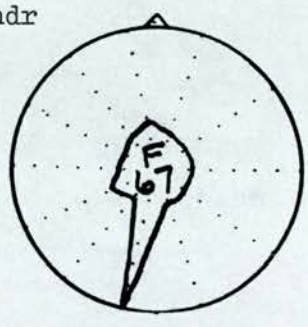
Stro



MH



ndr



APPENDIX 6

A published paper.

British Journal of Clinical Psychology (1985), **24**, 145–147 *Printed in Great Britain*

Response competition in recognition memory in three amnesic groups

Robert Taylor and **Ralph J. McGuire**

Experimental studies have suggested that a feature of the memory defect in some amnesic conditions may be a failure to inhibit competing responses at retrieval. It is not known whether this feature is present equally in different syndromes involving amnesia. Employing a recognition memory paradigm with varying numbers of response alternatives (based on Miller, 1978), three amnesic groups (Korsakov's, senile dementia, and multi-infarct dementia) were compared using four types of test material. The groups were not differentially affected by increasing numbers of response alternatives.

Considerable interest has been shown in the contribution of retrieval, as opposed to encoding or consolidation, defects in amnesic syndromes since Warrington & Weiskrantz' (1970) report that amnesic subjects' memory performance could be greatly improved by appropriate cueing at the time of retrieval. Miller (1975) has confirmed the ability of cueing to improve performance in presenile dementia to normal levels, and Morris *et al.* (1983) have replicated this in senile dementia. One explanation of the benefit of cueing is that amnesic subjects suffer from a relative inability to inhibit competing responses and that the partial information or cue acts by reducing the potential number of competing responses. Miller (1978, Expt 2) developed an elegant paradigm to test this: a recognition memory task using learning of word lists where at retrieval each target word appeared in a choice array of two, four, or eight words. He reasoned that his amnesic subjects should be particularly disadvantaged (compared with normals) by larger numbers of response alternatives, and this was what he found. He considers the possible clinical implications of such findings in Miller (1984).

In view of the diversity of amnesic syndromes (as reviewed by Hirst, 1982), the present study was designed to compare three amnesic conditions – alcoholic Korsakov's (K), senile dementia of the Alzheimer type (SD), and multi-infarct dementia (MID) – using a variant of Miller's design. Differences exist in the locations and types of anatomical damage in the three conditions (relatively characteristic in K and SD and variable in MID) as well as differences in their clinical characteristics. One might therefore expect the nature of the memory impairment to differ between the groups.

Method

Subjects and procedure

The subject groups consisted of 12 chronic hospitalized K subjects (10 M, 2 F; mean age 59·1 years, range 47–69), 12 SD subjects (9 M, 3 F; 75·5 years, 64–82), and 12 MID subjects (9 M, 3 F; 75·1 years, 65–83). Selection criteria involved a number of factors including consultant's diagnosis, score on the Hachinski index, aspects of medical status and history, and history of alcohol use. (No SD or MID subject had a history of alcohol abuse, while all K subjects were alcoholic Korsakov patients with no recent history of drinking.) People with prominent dysphasic or dyslexic difficulties, or with uncorrected impairments of sight or hearing, were excluded. Each subject was tested on 12 brief recognition memory tasks: three each of four types of test as follows.

(1) A prose paragraph was read out to the subject. After 30 s of general conversation s/he was asked a series of six questions about it, in each case a card being shown with the correct answer amongst one, three, or seven plausible but incorrect filler answers (all in lettering 1 cm high). The subject was required to choose the correct answer, and to guess if s/he did not know.

(2) A list of six common words (Thorndike & Lorge category AA) was read to the subject at a rate of one every 2 s. After 30 s of conversation s/he was shown a series of six cards, each with one

target word and one, three, or seven fillers drawn from the same pool (again in 1 cm lettering). The subject was asked to choose which word had been read out, and to guess if s/he did not know.

(3) A series of six meaningful photographs (5 × 6 cm, taken from colour magazines) was shown at a rate of one every 2 s. After 30 s of conversation the subject was required to choose which picture s/he had seen from each of six arrays (the one, three, or seven filler pictures being taken from the same source) and to guess if s/he did not know.

(4) This was as (3), but using meaningless designs.

Three 'sets' (*a*, *b*, *c*) of each of these four types of test were used (to collect enough data without making any single test too long or difficult): i.e. three paragraphs, three lists of six words, three sets of six photos, and three sets of six designs. They were presented in a cyclical fashion (1*a*, 2*a*, 3*a*, 4*a*, 1*b*, 2*b*, etc). Any given target item appeared in a two-choice array for a third of each subject group, a four-choice array for a third, and an eight-choice array for a third. The order (from 1 to 6) in which two-, four-, and eight-choice arrays appeared was also appropriately counterbalanced. The physical position of correct choices in arrays was randomized. No subject ever saw the same filler item twice.

Results

The mean numbers of correct choices (possible score 0–6) in each condition for the three groups are shown in Table 1.

Table 1. Mean number of correct choices (standard deviations in parentheses) in each condition

	Paragraphs			Words			Photos			Designs		
	2	4	8	2	4	8	2	4	8	2	4	8
K	4·17 (1·70)	2·75 (1·29)	1·92 (1·78)	4·33 (1·30)	3·08 (1·44)	2·50 (1·24)	5·42 (0·67)	3·83 (1·27)	3·58 (1·44)	4·58 (1·31)	3·67 (1·56)	1·83 (1·03)
SD	3·50 (1·57)	2·33 (1·44)	1·25 (1·14)	3·67 (1·50)	1·92 (1·16)	1·33 (1·07)	4·08 (1·51)	2·75 (1·36)	1·33 (1·23)	4·25 (0·87)	2·08 (1·24)	1·58 (1·00)
MID	3·67 (1·61)	3·25 (1·86)	3·17 (1·64)	3·58 (1·83)	2·67 (1·87)	1·75 (1·14)	5·50 (0·90)	4·67 (1·07)	3·50 (1·51)	4·58 (1·31)	3·58 (0·79)	2·42 (1·73)

Appropriate repeated measures ANOVAs were carried out, first on all the data and then for each type of material separately. To summarize: the SD group happened to perform rather more poorly than the other groups, and photos was an easier test than the others. The effect of number of response alternatives (NR) was, not surprisingly, always highly significant. The interaction effects of interest – those involving subject group by NR – were *not* significant (though there was some trend in paragraphs for the MID group to be less affected than the other groups by increasing NR). The analyses were repeated using chance-corrected scores (= [Po − Pc]/[1 − Pc] where Po is the obtained proportion correct and Pc is that expected by chance) and then using d' scores (as discussed in Meudell & Mayes, 1981). The outcome was largely unchanged, the point of note being that the main effect of NR was obliterated in the two verbal tests but remained in the two non-verbal ones.

Appropriate correlational analyses (in each group separately, and in all three combined) indicated no consistent relationship between overall *levels* of performance and the extent of the *decline* in performance as NR increased. Similarly neither the overall levels nor the extent of decline correlated with age in any group or in the three groups combined. The only apparent ceiling effect was in the MID group in photos two-choice, and this does not distort the results of the analyses in an important way.

Comparison of the three groups with two other subject groups – normal elderly and

depressed elderly – was attempted. Unfortunately the tests were too easy for these intact subjects: those tested scored five or six in all conditions. This ceiling effect would have led to a spurious difference between the pattern of performance of the impaired and the intact groups, and so the comparison was abandoned. Consideration of Miller's (1978, p.146) graph suggests that a ceiling effect in the two-alternative condition might have occurred in his study also, though Meudell & Mayes (1981) argue against this interpretation.

Discussion

Despite the differing neuropathological bases and clinical features of the three conditions, little difference between the groups emerged in this study regarding the importance of failure to inhibit competing responses. The extent to which *all* the groups were suffering from such a failure cannot be judged here in the absence of normal control data. The avoidance of ceiling effects in unimpaired subjects as well as floor effects in amnesic subjects is problematic, but controls must have 'room for improvement' just as amnesics do under the easiest or most helpful experimental manipulation if conclusions are to be drawn about essential defects in amnesia. Squire (1980) discusses whether amnesic subjects are *uniquely* benefited by certain kinds of prompts given at the time of the retention test, and concludes that they are not. Similarly, Meudell & Mayes (1981) found with two- vs. eight-choice word recognition that normals after long retention intervals showed similar patterns of performance to those seen in amnesic subjects after short intervals. Using a signal detection analysis they also found that d' scores were no worse with eight than with two choices: they suggest that a failure of response inhibition may not be present and that the concept of 'weak memory' is just as compatible with the findings. In the present study, a d' analysis removed the effect of increasing NR only in the two verbal tests. The study suggests that a particular recognition memory paradigm which is presumed to involve the inhibition of competing responses shows little difference between three amnesic groups, but that failure of response inhibition may nevertheless be present at least with certain types of material.

References

Hirst, W. (1982). The amnesic syndrome: Descriptions and explanations. *Psychological Bulletin*, **91**, 435–460.

Meudell, P. & Mayes, A. (1981). A similarity between weak normal memory and amnesia with 2 and 8 choice word recognition: A signal detection analysis. *Cortex*, **17**, 19–30.

Miller, E. (1975). Impaired recall and the memory disturbance in presenile dementia. *British Journal of Social and Clinical Psychology*, **14**, 73–79.

Miller, E. (1978). Retrieval from long-term memory in presenile dementia: Two tests of an hypothesis. *British Journal of Social and Clinical Psychology*, **17**, 143–148.

Miller, E. (1984). *Recovery and Management of Neuropsychological Impairments*. Chichester: Wiley.

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

Squire, L. (1980). Specifying the defect in human amnesia: Storage, retrieval and semantics. *Neuropsychologia*, **18**, 368–372.

Warrington, E. & Weiskrantz, L. (1970). Amnesic syndrome: Consolidation or retrieval? *Nature*, **228**, 628–630.

Received 18 April 1984; revised version received 5 November 1984

Requests for reprints should be addressed to Robert Taylor, University Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh EH10 5HF.

Ralph J. McGuire is also at the above address.

Additional information on the samples, materials, and results has been deposited with the British Lending Library, Boston Spa, Yorkshire, UK, as Supplementary Publication No. SUP 90, 098.

Appendix 7

A statistical note.

Indication of significance at the .10 level is used (on Page 133 and subsequently) as a convenient way of drawing attention to potentially interesting trends. It is not intended to indicate results which could be considered acceptable as significant in the conventional statistical sense.