THE FORMULATION AND STANDARDISATION OF A BATTERY OF TESTS OF VISUAL-PERCEPTUAL-MOTOR FUNCTION FOR THE ADULT CEREBRO - VASCULAR ACCIDENT PATIENT

Marjorie Eileen Concha

VOLUME I

A Thesis Submitted to the Faculty of Medicine, University of the Witwatersrand, Johannesburg for the Degree of Doctor of Philosophy.

Johannesburg 1998

Declaration

I declare that this thesis is my own unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Sue Concha

Marjorie Eileen Concha

26 day of 9, 1998.

Ethics Number: 18/1/87

Acknowledgements

The candidate wishes to express her appreciation and thanks to the following:

Prof. H. Reef, Head of the Department of Neurology at the University of the Witwatersrand and Chief Neurologist at the Johannesburg Hospital at the time that the work for this Thesis was begun, for his, encouragement and advice.

Prof. V. Fritz, Current Head of the Department of Neurology at the University of the Witwatersrand and the Johannesburg Hospital for taking over the supervision of this thesis, for her encouragement, advice and support.

Miss Angela Richards and Mrs Catherine Collin for the enormous amount of work put into typing the thesis.

Etel Concha, my husband for his unfailing support and assistance.

ABSTRACT

The purpose of this thesis was to construct a comprehensive battery of standardised tests that was robust in terms of its psychometric characteristics to evaluate visual-perceptual-motor (V-P-M) function in the adult who had suffered a Cerebro-Vascular-Accident (C.V.A.). This was to meet the need for the occupational therapist to target specific deficits and thereby to provide an effective treatment to patients. Work done by the candidate with patients, the literature as well as a review of existing tests demonstrated the need for such a test battery. Up to this time, tests available do not meet all the requirements for a reliable and valid assessment tool for adult V-P-M function. This demanded that the tests compiled for this thesis be based on a sound theoretical model, supplied information on age related changes in function and had undergone rigorous psychometric testing. It was also nescessary that normative data based on information collected from a random sample of the population for whom the test was intended and which would include a wide range of age groups, be collected. In order to fulfil these requirements the construction of this test battery followed the eleven steps necessary to produce tests ready for the final stage (step twelve) i.e. the collection of the normative data.

For a standardised test to be acceptable it must be based on a guiding theory which will provide definitions of the theoretical concepts and their relationship to each other. The theory serves as the basis of a model which is then tested through research aimed at making the model operational. The V-P-M test battery described in this thesis was constructed on the basis of a theoretical model devised by the candidate and focused specifically on the visual-perceptual process. The construction of the model was the first essential step in producing the standardised battery of tests. The second step required that the uses for the tests be clearly defined. In this instance it was necessary that the test battery clearly distinguished between the normal decrease in function with advancing age and the problems that might occur subsequent to a C.V.A. as well as being able to diagnose conventional V-P-M syndromes so that treatment could be focused.

Once the model of visual-perceptual-motor function was formulated, the fourth and fifth steps were to construct the test blue-print and the testing material taking into account constraints and variables in test administration which were identified in step three. A battery of fifteen tests was devised. Fourteen of the tests were constructed by the candidate to evaluate form perception, colour perception, part-whole, perceptual constancy, and figure-ground perception as well as the various components of the perception of space. The construction of the V-P-M Battery also took into account the differences that occurred when testing was done through motor free, motor reduced and motor enhanced activities. One published test, i.e. the Test of Visual - Motor *Integration* (Beery and Buktenica 1967 and 1982), used for testing perceptual function in children, was included in the battery as it was thought to be suitable for testing adults.

In step six, the fifteen tests were subjected to content validation by fifteen occupational therapists. At that stage all the tests were found to be better than average at assessing the skills for which they were constructed. At the same time nine C.V.A. patients (step seven) were also evaluated so that difficulties with administration, scoring and execution of the tests could be eliminated. On the basis of the information obtained at this stage, the preliminary battery was finalised (step eight). Preliminary, in this context, meant that changes might still had to be made to the tests in the battery once their psychometric characteristics had been analysed and the model tested in practice.

The battery of fifteen tests was then administered to 118 normal individuals between the ages of twenty years and seventy-nine years to test the hypotheses that there were no decrease in the ability to execute the tests and no increase in the time taken with advancing years (step nine). The data obtained on the normal individuals also provided the information necessary for the psychometric analyses of the tests and of the tests' items (step ten). Further a group of 33 patients who had suffered a right sided C.V.A. was also evaluated, to help in these investigations.

The results of the evaluation of the normal group of individuals (of which 39 were men and 79 were women) demonstrated that there were no differences in the performance between men and women. Nine of the fifteen tests revealed age related changes in test performance. Two of these tests showed changes from fifty years of age, four from sixty and three tests from the age of seventy years. To test for significant differences between the age groups non parametric tests were used. Value counts were performed and 2x2 tables constructed in which the figure equal to and less than the fifty per cent level of accuracy of each test was calculated. When the minimum expected value was equal to or greater than five, the Chi-square Test was used after the application of Yates' Continuity Correction. Fisher's Exact Probabilities were used when the minimum expected value was less than five. The null hypothesis of no decrease in function with advancing age was rejected in nine of the tests and retained in six. It is extremely important to differentiate between changes caused by ageing and those that take place as a result of brain injury so that a correct diagnosis can be made. Existing tests do not make these distinctions clear.

The time taken to complete the tests was longer with advancing age in all but one instance. The null hypothesis of no increase in the time taken was thus rejected in all but one case. The One Way Analysis of Variance was used for the statistical analysis of the time scores as the data were found to be normally distributed. Percentile ranks were calculated for both accuracy and time scores so that the potential of the battery to assess the level of patient performance could be evaluated.

Thirty three C.V.A. patients were evaluated by means of the V-P-M Battery and their scores compared to those of the normal subjects. The One Way Analysis of Variance was used for this purpose. Two tests i.e. the Form and the Colour Matching tests did not reveal any differences between the performance of the normal individuals and that of the patients and therefore their value as testing tools for C.V.A. patients was in doubt.

In step ten rigorous psychometric analyses of the tests in the battery were carried out. This included the use of the following statistical techniques: item efficiency analysis, reliability and validity studies, and factor analysis. An analytical investigation of each test to determine various visual-perceptual-motor syndromes such as apraxia, agnosia, spatial problems, unilateral neglect and achromatopsia was also conducted.

As part of the item efficiency analysis, the item difficulty and the item discrimination indices were calculated. This provided information that helped in the selection of the most appropriate items for the final version of each test in the battery.

The reliability studies included Split-Half reliability, Kuder-Richardson, and inter-rater reliability all of which proved to be within acceptable limits.

The test validity studies included those related to the content of the tests as well as their capacity to measure the skills for which they were constructed. The content was evaluated by experienced occupational therapists. The construct validity was evaluated through the proficiency of the tests to measure age related changes and visual-perceptual-motor problems.

Finally the communality and specificity of each test was determined as a first step in the factor analysis. To complete the factor analysis a correlation matrix and sorted rotated factor loadings were calculated for the tests in the battery. This provided the means of testing the V-P-M model and on the basis of the findings, the model was reformulated. The tests in the battery were also evaluated and adjusted against this model (step eleven). The new model formulated by the candidate provides a well researched theoretical basis of V-P-M function. Through this process it became apparent that certain visual-perceptual skills that were thought to be distinct entities , do not exist as such. This made it possible to reduce the battery from fifteen tests to a more useful twelve tests covering three main visual-perceptual functions i.e. visual discrimination, spatial perception and colour perception, to be tested through the use of motor free, motor reduced and motor enhanced tests.

The completed V-P-M Test Battery, is ready for the final step (step twelve) when it will be administered to a random selection of the adult population for the provision of normative data.

CONTENTS

VOLUME II

VOLUME III

Appendices

TABLES

Table No.		Page Number
11.2	Mean and SD of Accuracy Scores - SR	254
11.3	Mean and SD of Accuracy Scores for Each Age Group	255
11.4	Comparison of Accuracy Scores - SR	255
11.5	Comparison of Accuracy Scores for Ages 20-69 and 70-79 Years - SR	256
11.6	Mean and SD of Accuracy Scores - PIS	256
11.7	Mean and SD of Accuracy Scores for Each Age Group - PIS	257
11.8	Comparison of Accuracy Scores - PIS	257
11.9	Comparison of Accuracy Scores Ages 20-59 and 60-79 Years - PIS	258
11.10	Mean and SD of Time Scores - SR	258
11.11	Mean and SD of Time Scores For Each Age Group - SR	259
11.12	Differences in Time Scores - SR	260
11.13	Differences in Re-Grouped Time Scores - SR	260
11.14	Mean and SD for Accuracy Scores for Patients - SR	262
11.15	Mean and SD for Accuracy Scores for Patients - PIS	262
11.16	Mean and SD of Time Scores for Patients - SR	263
11.17	Differences in Accuracy Scores Obtained by Normal Subjects and Patients - SR	264
11.18	Differences in Accuracy Scores Obtained by Normal Subjects and Patients - PIS	264
11.19	Differences in Time Scores Obtained by Normal Subjects and Patients - SR	265
11.20	Difficulty Index - SR	269
11.21	Difficulty Index - Patient Group - SR	271
11.22	Reliability Co-Efficients - SR	272
11.23	Content Validity - SR and PIS	273
11.24	Correlation with Other Tests in the Battery - SR and PIS	275
11.25	Sorted Rotated Factor Loadings - SR and PIS	276

Table No.	VOLUME II	Page Number
	Chapter 12 The Block Construction Tests	
12.1	Mean and SD of Accuracy Scores - BC I	286
12.2	Mean and SD of Accuracy Scores for Each Age Group - BC I	287
12.3	Comparison of Accuracy Scores - BC I	287
12.4	Comparison of Accuracy Scores for Ages 20-69 and 70-79 Years - BCI	288
12.5	Mean and SD of Accuracy Scores - BC II	288
12.6	Mean and SD of Accuracy Scores for Each Age Group - BC II	288
12.7	Comparison of Accuracy Scores - BC II	289
12.8	Mean and SD of Time Scores - BC I	290
12.9	Mean and SD of Time Scores for Each Age Group - BC I	290
12.10	Differences in Time Scores - BC I	291
12.11	Differences in Re-Grouped Time Scores - BC I	291
12.12	Mean and SD for Time Scores - BC II	292
12.13	Mean and SD of Time Scores for Each Age Group - BC II	292
12.14	Differences in Time Scores - BC II	293
12.15	Differences in Re-Grouped Time Scores - BC II	294
12.16	Mean and SD for Accuracy Scores for Patients - BC I	295
12.17	Mean and SD for Accuracy Scores for Patients - BC II	295
12.18	Mean and SD For Time Scores for Patients - BC I	296
12.19	Mean and SD for Time Scores for Patients - BC II	296
12.20	Differences in Accuracy Scores Obtained by Normal Subjects and Patients - BC I	297
12.21	Differences in Accuracy Scores Obtained by Normal Subjects and Patients - BC II	297
12.22	Differences in Time Scores Obtained by Normal Subjects and Patients - BC I	298
12.23	Differences in Time Scores Obtained by Normal Subjects and Patients - BC II	298
12.24	Difficulty Index - BC I	306
12.25	Difficulty Index - BC II	307
12.26	Difficulty Index - Patient Group - BC I	307
12.27	Difficulty Index - Patient Group - BC II	308
12.28	Reliability Co-Efficients - BC I and BC II	310
12.29	Content Validity - BC I and BC II	311

FIGURES

Fig. No.		Page Number
	Chapter 10 Tests of Directionality - The Arrows and the Dot-Line Tests	

Fig. No.	VOLUME II	Page Number

CHAPTER 1

INTRODUCTION

An occlusion or a rupture of a diseased blood vessel causes damage to the brain and is known as a cerebral vascular accident or stroke. Cerebral Vascular Accident (C.V.A.) is a widespread and debilitating disease. It is estimated that there are 500 000 new adult C.V.A. victims annually in the U.S.A. and at any given time there are two million living C.V.A. patients (Siev and Freishtat 1976). Benaron (1976) stated that there were approximately 7 750 stroke cases in the London area yearly, many of them under the age of thirty five years. The extent of the problem was illustrated by Rose and Capildeo (1981) who found that the incidence of stroke was two per 1 000 of the population in the sixty-five years and older groups. The incidence of stroke in an urban black South African population was found to be approximately one per 1 000 per year (Rosman 1986). An analysis of the patient load in the Occupational Therapy department of the Johannesburg Hospital showed that half of the number of patients treated were diagnosed as hemiplegic as a result of a C.V.A. (Concha 1978). It is clear that the number of C.V.A. patients requiring rehabilitation is very high.

Occupational therapy is concerned with this rehabilitation and in helping patients to lead as normal a life as possible. The patient is assisted to attain independence in activities of daily living (ADL) such as eating and dressing as well as at work and in the use of free time. Persistent Visual-Perceptual-Motor (V-P-M) problems experienced by C.V.A. victims have been associated with poor function in all aspects of activities of daily living and it is important that the cause of dysfunction be correctly diagnosed so that therapy can be correctly applied.

Patients who have suffered a C.V.A. may have many sequelae which result in an inability to lead a normal life. Problems with visual-perceptual-motor function (i.e. the processing and interpretation of visual stimuli and their association with the motor pattern required to react to the stimuli) are only one of a group of symptoms that can result from a C.V.A. Visual-perceptual-motor problems and the other symptoms such as paralysis or paresis of one side of the body, sensory loss and aphasia, make rehabilitation a very complex process.

Many C.V.A. patients are elderly (Rose and Capildeo 1981) and therefore cognisance must be taken of the normal decrease in perceptual function with advancing age. Lezak (1983) has stated that age graded scores were important aids to differential diagnosis and were essential to the evaluation of test performance in patients over sixty-five years. Current tests used by occupational therapists do not provide adequate information on the effects of ageing. For example, Cockburn et al (1982) found no decrease in function with advancing age when using the Rivermead Perceptual Assessment Battery (RPAB) nor did Itzkovich et al (1990) when using the Loewenstein Occupational Therapy Cognitive Assessment (LOTCA), in spite of the fact that adult scores on some tests varied with age. For instance the median scores of the Wechsler Adult Intelligence Scale (WAIS) Block Design and Object Assembly Tests reached a peak at eighteen to nineteen years of age then gradually declined (Lezak 1983). Taylor (1968), Kaplan and Hier (1982) and Farver and Farver (1982) found decreased function with age when using visuo-spatial tests but no data on normal subjects were made available. This made it difficult to use the tests mentioned by these authors for diagnostic purposes.

Additionally there are no tests with South African norms for the occupational therapist to use. Two tests published by the National Institute for Personnel Research i.e. the Fret Repetition Test (Grant 1969) and the Symco Test (Crawford-Nutt 1977) were made available to the candidate. Neither test was found to be suitable for the evaluation of V-P-M function as defined in this thesis.

Up and until the time that the study for this thesis was begun, there was no comprehensive test battery available for the occupational therapist to use in the delimitation of visual-perceptual-motor function in the adult C.V.A. patient. Thus research carried out on perceptual function was often unsatisfactory.

Anderson and Choy (1970) have suggested that the emphasis in treatment should be on perceptual training and not on activities of daily living (ADL) and if rehabilitation strategies for visual perceptual problems were to be improved as suggested by Kaplan and Hier (1982), the diagnosis of these problems had to be refined. According to Ottenbacher and Goar (1986:151) *"there appears to be a professional consensus within occupational therapy that without valid and reliable assessment devices the discipline's ability to substantiate treatment effectiveness will be reduced or compromised."* Gillette (1982) likewise observed that the lack of standardised tests on which to base the outcome of research was a strong deterrent to a data base showing the results of intervention through occupational therapy.

1.1. PURPOSE OF THE STUDY

It was therefore the purpose of this study to:

a) Formulate a model of Visual-Perceptual-Motor function of relevance to occupational therapy. This is an important first step to demarcating the perceptual areas to be tested and is a pre-requisite to the formulation of any test material.

b) Design and construct suitable test material to evaluate the level of an individual's V-P-M function in the areas demarcated in the model as well as in the various age groups.

c) Follow the steps necessary for the ultimate standardisation of the test battery which included obtaining data on a normal group of individuals in the age groups for which the battery was intended.

d) Try out the test battery thus formulated, on a group of patients in order to determine its value in the assessment of V-P-M function in the C.V.A. patient and to determine the way in which the battery could be used to diagnose certain V-P-M syndromes.

e) To test the model of V-P-M function through the analysis of the factors evaluated through the battery of tests and to draw up and adjust the model if necessary.

1.2. LIMITATIONS OF THE STUDY

a) The data obtained provided the information to construct the final format of the test battery. This final format will not be included due to the limitations of space although the guidelines for it will be presented.

b) No attempt was made in this thesis to collect information on the types of performance observed in left brain damaged subjects as compared to right brain damaged subjects. This could only be done once a well constructed test battery was made available.

c) As the specific aim of this thesis was to develop a meaningful battery of tests to evaluate V-P-M function in the C.V.A. patient, the focus of the research was on the steps needed to develop this battery and the consequent testing of the Model of V-P-M function. A brief investigation of patient performance was part of this process. A detailed analysis of patient's performance is beyond the scope of this thesis and will only be possible once a well researched test battery is available. Detailed medical information was not collected on each patient as the purpose of testing patients was only to determine whether V-P-M dysfunction could be demonstrated by means of the tests.

d) Where it was possible to illustrate the mistakes made by the normal group of individuals and the patient group, this was done by the inclusion of copies of the tests or the pictures of the constructions.

1.3. METHOD OF PRESENTING THE THESIS

On account of the complexity of the standardisation process this thesis has been written up in the following way:

The review of the literature (Chapter 2) is a general review of the work done in the field. The discussion of the model of visual-perceptual-motor function used as the basis for construction of the tests in the battery (Chapter 4) follows the chapter on research methods (Chapter 3).

The results for each test or pair of tests are discussed in separate chapters. (Chapters 5-15). Each of these chapters contain a review of the literature related to the test, the results of the statistical analysis and a discussion of the results.

A review of the types of dysfunction found in stroke patients and the way each test demonstrated this dysfunction is to be found in chapters 16-20. The final chapter (Chapter 21) presents the overall results and conclusions.

Appendix 1 contains copies of the tests that make up the V-P-M battery together with the administration and scoring instructions. The percentile scores for each age group for both time and accuracy scores for each test are to be found in Appendix 2 and a brief description of each patients' performance in Appendix 3. Appendix 4 contains a summary of the different types of function found in left versus right cerebral lesions and in Appendix 5 are presented the journal articles published by the candidate in support of this thesis.

CHAPTER 2

REVIEW OF THE LITERATURE

2.1. REVIEW OF PREVIOUS RESEARCH IN THE SAME FIELD

This review is concerned only with the general field of testing V-P-M function. More detailed reviews of individual tests are presented in the chapters dealing with the tests that make up the V-P-M battery. The literature concerned with the various V-P-M syndromes is reviewed in separate chapters.

Eminent workers such as Luria and Critchley have described the range of perceptual and cognitive problems which might occur as a result of brain injury. They have also described ways in which these problems might be evaluated (Critchley 1953 and Luria 1966, 1980).

However, these evaluations have not been standardised for purposes of administration and scoring. Christensen (1975) attempted to draw up a standard administration procedure for Luria's Neuropsychological Investigation and Golden et al (1980) prepared a standardised version for some parts of the tests described by Christensen. Golden et al (1980) obtained normative data on these tests and these were published as the Luria Nebraska Neuropsychological Battery (Golden et al 1985). Unfortunately, Luria's tests were primarily designed for the hemispheric localisation of brain disorders which made them unsuitable for occupational therapists although both Burnell (1985) and Barrett (1986) have suggested that they might be used by them as an assessment tool. In addition Golden et al's version, according to Lezak (1983), had some serious limitations. Certainly the Screening Test for Adults (Golden 1987) contained no visual perceptual components.

Benton et al (1983:vii) put together a collection of twelve tests, each of which *"was developed to answer a specific question or cover a specific need"* and which were utilised by Benton and his colleagues during the years preceding publication. Of the twelve tests described by Benton et al

(1983), four could be useful for testing aspects of visual-perceptual-motor function, for example, the judgement of line direction and the three dimensional block construction. As Benton et al's tests were only commercially made available after the V-P-M battery was drawn up for this thesis, their inclusion in the battery was not possible. Each of Benton et al's sub-tests have data on normal subjects, and each sub-test is discussed in the relevant chapters of this thesis.

The most widely used set of clinical procedures designed to assess brain behaviour relationships is the Halstead-Reitan Neuropsychological Battery (Boll 1981). This battery, however, contained no visual perceptual tests and must be augmented by tests such as the Wechsler Adult Intelligence Scale to obtain information on visual perceptual function.

Some occupational therapists have used Critchley's (1953) text on the Parietal Lobes as well as the tests described therein, to draw up evaluation procedures. The tests described by Critchley consisted of tasks such as draw-a-daisy and a bicycle, copy forms of match sticks, and building with blocks (Gregory and Aitken 1971, Little and Sylvester 1971, McMenamin 1976 and Sylvester et al 1974). None of these tasks had standards for administration and scoring or offered data on the performance of normal people, although Lezak (1983) has provided a scoring system for free drawings of a bicycle, house and clock face.

Jean Ayres in the USA has been the most significant and influential of all occupational therapists in the field of perception. Ayres felt the need for standardised testing material for children with learning disorders in the field of sensory-motor function and drew up a test battery, the Southern California Sensory Integration Tests (Ayres, 1962, 1972, 1980a). The tests that make up this battery encompass aspects of visual perception such as figure-ground perception and position in space as well as several tests for the evaluation of kinaesthetic and tactile function. The battery was standardised for children between the ages of four years and eight years eleven months and extensive information is available on the use of the tests.

Several of the Ayres' Tests have been used by occupational therapists for testing adults (Taylor 1968, Gane 1976, Siev and Freishtat 1976, Kaplan and Hier 1982, Fox and Harlowe 1984, Jongbloed et al 1986 and Siev, Freishtat and Zoltan 1986) although no published data on normal adult performance is available.

Marianne Frostig, a remedial teacher, who formulated a test of visual perception (Frostig 1966) for use with children with learning problems, has also influenced the work of occupational therapists in the field of perception testing. The Developmental Test of Visual Perception (DTVP) (Frostig 1966) has well defined sub-tests for the evaluation of figure-ground perception, form constancy, position in space and space relations which have been standardised for children between the ages of four years and seven years eleven months. The sub-tests have also been adapted by occupational therapists for use in testing adults (Taylor, 1968, Gane 1976, Siev and Freishtat 1976, Harlowe and Van Deusen 1984, Siev et al 1986). The test was updated and revised by Hammill, Pearson and Voress in 1993, although normative data were still provided only to age ten years eleven months, and no adult studies have been carried out using the revised DTVP. As with the Ayres' tests normative data was not available on the DTVP for adults and the scoring criteria used to determine dysfunction in adult studies were arbitrarily allocated.

Siev and Freishtat (1976) and Siev, Freishtat and Zolton (1986), made a significant contribution to the field of adult perceptual evaluation by defining terms and listing tests that could be used to evaluate function in each defined area. However, most of the tests listed were similar to the type of tests suggested by Critchley (1953) or belonged to the Developmental Test of Visual Perception (Frostig 1966) or to the Sensory Integration Tests (Ayres 1972). As has been stated, none of these tests have been standardised for adults. Similarly, the Ontario Study Group on the Brain Damaged Adult (1972) drew up a manual containing unstandardised tests of a simple nature. To test the *"validity of the OSOT Perceptual Evaluation for differentiating neurologically normal persons from those ... diagnosed as neurologically impaired"* (pg. 93), Boys et al (1988) studied the results of eighty patients and seventy normal control subjects. Cut off scores to determine impaired function were established from these results.

Normative data for perceptual tests of relevance to the occupational therapist have been obtained in other studies (Taylor 1968, Farver and Farver 1982, Cockburn et al 1982, Tyerman et al 1986, Itzkovich et al 1990 and Warrington and James 1991). Only one study, that of Taylor (1968) was carried out prior to the formulation of the test material for this thesis which was first described in 1982 (Concha 1982). Taylor's (1968) study was one of the most comprehensive studies which was carried out on a range of tests obtained from different sources. The study provided ranges, means and standard deviations of the scores obtained for each test used. Split-half reliability and validity studies (by comparing test scores to ratings obtained in activities of daily living) were also included. However, test instructions and administration procedures were not given and the tests have not been classified according to the various aspects of perceptual function. The scientific value of the study could have been enhanced if information had been more comprehensively published.

Farver and Farver (1982) tested sixty-eight subjects between the ages of forty years and eighty-eight years in order to gather preliminary data for the Parietal Lobe Tests of the Boston Diagnostic Aphasia Examination and its two supplementary visuo-spatial tests. The candidate analysed these tests in order to determine the perceptual skills evaluated. The analysis revealed that they covered position in space, space relations, figure-ground and part-whole concepts. Farver and Farver's study, as they stated, was a preliminary study but results did provide information on the effects of age on the test results, there being a statistically significant age effect. This study could have been of greater usefulness had the data on the normal subjects been provided in spite of the fact that it was collected from only sixty-eight subjects.

Cockburn et al (1982) and Bhavnani et al (1983) recognising the need for organised, standardised testing material for adults suffering from brain damage, drew up a battery of tests at the Rivermead Rehabilitation Centre at Oxford in the United Kingdom. These tests published later by Whiting et al (1985) and called the Rivermead Perceptual Assessment Battery (RPAB) have administration and scoring instructions and criteria for determining perceptual deficits which were obtained by testing only sixty-nine normal subjects between the ages of sixteen and sixty-nine. Lincoln and Clarke (1986) obtained additional data on the tests based on a sample of 100 subjects between the ages of sixty-five and ninety-five years. The RPAB was not constructed to cover perceptual tasks with and without motor output. When the candidate analysed the perceptual skills needed to execute the various tasks, it became apparent that most aspects of visual-perception had been included, but had not been classified by the authors and no model of perceptual function was provided nor were there distinctions made between the different types of motor activity required in the tests.

The Rivermead Perceptual Assessment Battery, although a useful tool, has some very simple tests (the mean for the normal group is the maximum score in thirteen of the sixteen tests) and no decrease of function with age was found in either the Cockburn et al's (1982) study or the Lincoln and Clarke (1986) study. This might show a ceiling effect. Lincoln and Clarke, in fact, found that the elderly subjects performed better on the Body Image, Three Dimensional Copying and Figure-Ground tests than the younger subjects. These findings were contrary to the findings in other studies (Taylor 1968, Farver and Farver 1982 and Concha 1982, 1986) and probably indicated the simplistic nature of the tests. It did, however, mean that the tests could be used with the elderly population as age did not have a deleterious effect on test performance.

Jongbloed et al (1986) published some preliminary findings on the reliability and validity of a Sensorimotor Test Battery. This battery emphasised the sensory system and used ten measures of kinaesthetic and tactile sensation from the Southern California Sensory Integration Tests which they adapted for adults. The only visual perceptual test included in the battery was the Hooper Visual Organisation Test (Hooper 1957). There were no data on normal subjects for the tests in this battery as was the case in the St Mary's C.V.A. Evaluation (Fox and Harlowe 1984, Harlowe and Van Deusen 1984) which also contained measures of perception adapted from the Southern

California Sensory Integration Tests. Patient performance on the tests in this instance was rated as intact, impaired or absent according to the standards established by St Mary's occupational therapy staff. None of these ratings were compared to normal performance.

In 1986 the Chessington Occupational Therapy Neurological Assessment Battery (COTNAB) was published (Tyerman et al 1986). This test battery contained twelve tests, only three of which were classified as visual perceptual and three as constructional in nature. Although the test battery was "standardised" on one hundred and fifty normal individuals between the ages of sixteen years and sixty-five years, no studies on the differences in performance between the different age groups (i.e. sixteen to twenty, twenty-one to forty-nine and fifty to sixty-five) were carried out. Information on an additional forty-seven subjects in the age range sixty-five to eighty-four years was obtained to extend the upper limits of the battery in 1992. Although the scores provided by the normal individuals tested contributed to the way in which the scoring system was developed for each of the age groups, it was not entirely clear how this was done. It seemed that the percentiles were evaluated within broad bands of ability e.g. unable or unwilling, impaired, below average or borderline, within normal limits. A point system was allocated for each band and was used to determine the level of function of the individual. The six visual-perceptual-motor tasks cover figure-ground ability, form constancy, space relations, part-whole perception, and constructional ability. Each of these tests will be discussed and evaluated at the appropriate place in the following chapters. No age related studies were carried out on the COTNAB nor were there any published reliability studies. Validity was restricted to comparisons between the functions of the normal group and that of a group of brain damaged subjects and C.V.A. patients. In spite of the above limitations the COTNAB is a well presented test battery. It is, however, extremely expensive to purchase is not transportable due to its size and the limited visual perceptual tests do not warrant its extensive use for testing V-P-M function.

The Loewenstein Occupational Therapy Cognitive Assessment (LOTCA) (Itzkovich et al 1990) was divided into four sections i.e. Orientation (two tests), Perception (six tests), Visuo-Motor Organization (seven tests), and

Thinking Equations (four tests). Normative data were collected on fifty-five normal individuals (a very small number for standardisation purposes) between the ages of twenty and seventy years. This is a fairly well presented battery of tests containing simple tasks. However, the instructions for administering the battery focused performance more on verbal understanding than on perceptual ability. In addition, it is the opinion of the candidate that the tests are not well formulated for assessing the named skills as they are too simple for adults and the authors found no change in test performance with advancing age. All normal (control group) subjects scored almost perfectly ($\bar{x}$ score = 3,7-3,9, max. score = 4) on the perception and visuo-motor sections. Information on the performance of two hundred and forty normal children aged six years to twelve years was provided which should make the test useful in evaluating brain damaged children.

Studies conducted on the (LOTCA), to determine its reliability and validity have been described (Katz et al 1989). These studies were carried out on forty-eight patients and fifty-five non-brain injured controls.

Shan-Shan, Fisher and Cermak (1987) used the Test of Visual- Perceptual Skills (non motor) (Gardner 1982) to evaluate the performance of learning disabled (mean age 18,8 years) and normal young men (mean age 20,4 years). They concluded that an alternative shorter version excluding the easiest items could be used for testing adult function. This conclusion was based on the fact that *"a ceiling effect was only seen among the normal subjects on three out of seven subtests. Furthermore, none of the subjects obtained a perfect total TVPS score"* (pg. 795). Since the Shan-Shan et al (1987) study was published, Gardner (1988 and 1992) developed a different version of the test, called the Test of Visual-Perceptual Skills (non motor) Upper Level, to evaluate visual perceptual function in young adults between the ages of twelve and eighteen years.

Arnadottir (1990) drew up a test instrument, the Arnadottir OT - ADL Neurobehavioral Evaluation (A - ONE) to test activities of daily living and neurobehavioral components simultaneously. In this test the patient was required to carry out a number of standardised personal management tasks such as dressing, eating and transfers, tasks which were not culture free. A neurobehavioral impairment score is provided once the observation scores have been totalled. The neurobehavioral check list used to provide this impairment score gives an excellent summary of the problems that may be observed. However, the method used to diagnose the problems required an extremely experienced therapist to make the correct observations and there may be "*multiple correct answers to the same question*" (Arnadottir 1990: 265).

It is the candidate's opinion that a less time consuming assessment could be carried out through the use of a properly standardised V-P-M test which would be compared to the relevant cultural method of executing self care and other activities of daily living, particularly within the multi-cultural context of South Africa.

In 1991, Warrington and James published the Visual Object and Space Perception Battery (VOSP). This battery contained nine tests (all in two dimensional form) which evaluated part-whole perception, perceptual constancy, spatial relationships and position in space (space visualisation). The normative data for the battery were obtained from two hundred subjects for some tests and from one hundred and fifty subjects for others. The data were divided into two i.e. for those younger than fifty years and for those older than fifty years as *"there appeared to be a minor difference in performance"* after the age of fifty (Warrington and James 1991:8). For each test, the mean, the standard deviation and the percentage cut off scores were provided for the two age bands. The VOSP offers some useful tests but is limited by the lack of tests including a motor component.

When the research for this thesis began, as previously mentioned, there was only one reference in the occupational therapy literature (Taylor 1968) to a battery of tests for visual-perceptual-motor function in the adult stroke patient and the tests were not commercially available. The candidate, therefore, felt that it would be useful to draw up a standardised test for the evaluation of visual-perceptual-motor function and to provide South African norms against which patients could be assessed. The appearance of test batteries such as those published by Whiting et al (1985) Tyerman et al

(1986), Itzkovich et al (1990) have indicated the need for a comprehensive battery of tests which would cover a clearly defined field of visual-perceptual-motor function. Tests covering all aspects of V-P-M function and which include normative data obtained from a representative sample of the South African population still remains a high priority as correct diagnosis of perceptual problems must lead to more effective means of therapy.

2.2. AGE AS A FACTOR IN PERCEPTUAL DEVELOPMENT

Decreased perceptual function with age has been well documented (Birren and Schaie 1977). These changes in function might be due to a number of factors such as:-

a) Changes in the transmission and accommodative power of the eye that begins to occur at about thirty-five to forty-five years of age (Fozard et al 1977). These changes affect distance vision, sensitivity to glare, binocular depth perception and colour sensitivity. Further deterioration in the eye takes place at about fifty-five to sixty years of age and affects the size of the visual field, sensitivity to low quality of light and sensitivity to flicker.

b) Loss of visual acuity which reduces the value of monocular vision for distances (Fozard et al 1977).

c) Diminished flexibility and adaptability which affect test performance in the elderly causing low scores on tests and lead to incorrect conclusions about the intellectual capacity of older persons (Lezak 1983).

d) Short attention span and increased distractibility which may also lead to poor scores on tests requiring concentration.

e) Structural changes in the nervous system such as decrease in size and flattening of the surface of the brain as well as increasing amounts of intracranial space which form part of the normal ageing process (Lezak 1983). Concomitant microscopic, biochemical and electro-physiological changes take place as well. All this affects the normal visual-perceptual-motor process.

f) Decrease in intellectual ability.

Lezak (1983) stated that the normal intellectual decline associated with old age showed up most strikingly in four areas of intellectual activity i.e.:-

i. the secondary memory;

ii. diminished ability for abstract and complex conceptualisation;

iii. mental inflexibility;

iv. slowing of general behaviour.

Decline of function in abstract and complex conceptualisation, mental inflexibility and slowing of behaviour were revealed in the results of this thesis.

Cristarella (1977) had listed the visual functions that decline with age and which affect the accomplishment of daily living tasks as - acuity, accommodation, visual fields, light sensitivity, dark adaptation, brightness discrimination, colour discrimination, visual spatial ability, visual figure-ground, perception of ambiguous figures, visual memory, visual closure, depth perception, visual serial learning ability, resistance to flicker fusion, and resistance to visual illusions. Visual fields, colour discrimination, spatial ability, figure-ground perception, visual closure and depth perception are perceptual functions tested by the V-P-M battery.

Research has shown that older adults tended to score less well in tests of fluid intelligence (perceptual function) than younger adults, but score much the same in crystallised intelligence (verbal function) until quite late in life (Horn and Cattell 1966, Storck, Looft and Hooper 1972, Hayslip and Sterns 1979 and Botwinick 1981). Many tests that were thought to measure crystallised intelligence were similar to those that measure left hemisphere function (e.g. vocabulary tests). Also those that were thought to measure fluid intelligence were the same as right hemisphere tests (e.g. block designs). It was not surprising, therefore, that several investigators had noted the similarity of the performance defects of normal older people to those of patients with right hemisphere damage (Klisz 1978, Concha 1982, Farver and Farver 1982 and Concha 1986). Schaie and Schaie (1977) reminded us that although tests of old people might look like those of right hemisphere patients, they did not necessarily reflect right hemisphere dysfunction. It seemed that the normal ageing process depressed skills that were integrative and sensory perceptual in nature and that involved speed of response (Botwinick 1981). Tests developed on the basis of young people's performance should not be used with the elderly.

Reitan (1962) and Davies (1968) have equated the effects of brain damage in adults of thirty to thirty-five years of age with that of ageing. This, however, did not explain the functional loss that occurred in the young brain damaged versus the elderly non-brain damaged individual.

It is important to be able to evaluate the effects of a C.V.A. and the effects of the ageing process separately as therapists have often confused the effects of ageing with those resulting from the C.V.A. and patients have been diagnosed as suffering from visual-perceptual dysfunction whereas in reality they were performing normally for their age.

For each of the tests in the battery described in this thesis, the relevant literature regarding performance of the elderly will be reviewed. The development of the skills in childhood is summarised in Chapter Three as it demonstrated clearly the way in which components of V-P -M function are structured. In addition it has been found that recovery from brain damage in the adult often takes place in a sequence similar to the developmental of V-P-M function in the child (Smythies1956).

2.3. THE EFFECTS OF DAMAGE TO THE BRAIN

A variety of problems could result from damage to the brain. These problems would be of an intellectual, perceptual and motor nature, as well as involving memory, attention and language. It was not within the scope of this thesis to describe all the effects of brain damage and how they could be evaluated, but only those defects which may be classified as of a visual-perceptual-motor nature.

Moses (1983) summarised problems in the sphere of visual-perceptual-motor ability which occurred as a result of cerebral injury as a *"breakdown in* *the integrated model of problem solving that is characteristic of visual-motor performance of the adult. Visual-motor regression ensues and there emerges a more elementary piecemeal attempt at picture analysis"* (Moses 1983:26).

It would appear that the two sides of the brain perform differently as far as this "picture analysis" is concerned. Numerous studies which analysed the function of the two sides of the brain have been described by authors such as Warrington et al (1966), Gazzaniga and Le Doux (1978) and Lezak (1983). It seemed that the left hemisphere is the primary mediator of verbal function and, although involved in the perception of shapes, textures and patterns, spatial relationships, visual imagery, drawing and construction, it was less efficient at these processes than the right hemisphere. It has been found that the right hemisphere processed this type of information in a holistic, synthetic fashion, while the left processed the information analytically (Gazzaniga and Le Doux, 1978). (Appendix 4 briefly summarises the differences). Thus, it appeared that both hemispheres were capable of appreciating the visuo-spatial aspects of a task, but that the right hemisphere is superior to the left in constructing the perceived relations by manipulating the item appropriately (Gazzaniga and Le Doux 1978).

According to Gazzaniga and Le Doux most studies that have examined perceptual function and attempted to localise that function have used only manipulo-spatial activities in the study design. It would seem, therefore, relevant to design a test which would be capable of evaluating perceptual ability through constructional tasks as well as through tasks that required mental manipulation of perceptual concepts without motor output. The V-P-M battery was, therefore, constructed on this basis so that the different types of function could be easily evaluated.

The various types of V-P-M problems that could result from cerebro-vascular accidents and which would be determined by the V-P-M battery are described in Chapters Sixteen to Twenty. Each patient tested is very briefly described in Appendix 3. Due to the fact that V-P-M dysfunction differed in its manifestation between the two sides of the brain and that a language disorder would affect the results of a test, the C.V.A. cases tested for this study were left hemiplegics which assumed right cerebral hemisphere damage.

Christensen (1974) and Luria (1980) attempted to provide information on the anatomical site of a brain lesion through perceptual and cognitive testing. As it is not the function of an occupational therapist to make this type of diagnosis and indeed we are so prevented by legislation, this thesis does not provide information on the site of any brain damage but rather focuses on the type of functional problems that result from V-P-M dysfunction.

2.4. THE EFFECTS OF V-P-M FUNCTION ON ACTIVITIES OF DAILY LIVING

Some investigations have focused on the relationship between perception and self care status (Lorenze and Cancro 1962, Anderson et al 1974, Kaplan and Hier 1982, Carter et al 1988, Arnadottir 1990 and Titus et al 1991). Anderson and Choy (1970) have gone so far as to state that improvement in perceptual function reduced the necessity of training patients in activities of daily living but they offered no empirical evidence to support this.

There is disagreement among the researchers investigating the link between perception and activities of daily living. Some studies have indicated a correlation between visual perceptual defects and self care status (Lorenze and Cancro 1962, Anderson et al 1974, Feigenson et al 1977a, Arnadottir 1990 and Titus et al 1991), whereas some have indicated that there is no relationship (Rosenthal et al 1965 and Taylor et al 1971). Kaplan and Hier (1982) found a significant negative correlation between the improvement of self care and visuo-spatial defects which they interpreted as meaning that patients with *"severe visuo-spatial or visual deficits were less likely to improve than patients without these deficits"* (Kaplan and Hier 1982:320). Arnadottir (1990) has inferred a strong link between performance on ADL tasks and perceptual function and has based her assessment on the analysis of these tasks.

The different findings might be due to the type of tests used and the paucity of standardised procedures. The studies mentioned above used either the performance sub-tests of the Wechsler Adult Intelligence Scale (Lorenz and Cancro 1962, Anderson et al 1974 and Kaplan and Hier 1982), tests formulated for children (Taylor et al 1971) or do not mention test procedures at all (Feigenson et al 1977a). Titus et al (1991) used some standardised and some unstandardised procedures for evaluating five aspects of perception in patients who had suffered a C.V.A. and correlating them with skill in dressing, eating, upper extremity hygiene and communication. Although statistically significant, non of the correlation coefficients were above r=0,55. Titus et al (1991) found perceptual deficits in all the patients tested and concluded that the correlations with ADL tasks indicated that perceptual testing should take place in all C.V.A. patients.

In addition, the above studies did not allow for problems that were experienced in Activities of Daily Living that occurred from paralysis or paresis of one side of the body, nor did they allow for the fact that normal individuals who demonstrated percetual changes as part of the ageing process did not have problems with carrying out Activities of Daily Living.

2.5. CONCLUSIONS

A study of the relevant literature both prior to and after the commencement of this study has led to the conclusion that there is a need for a visual-perceptual-motor battery that:-

a) Evaluated specifically demarcated and defined areas of visual-perceptual-motor function in the adult.

b) Has been shown to be effective in demonstrating problems in the C.V.A. patient.

c) Is standardised in its administration and has norms against which the performance of patients of different ages can be evaluated.

The following chapter describes the research design used to fulfil the above aims.

CHAPTER 3

RESEARCH DESIGN

3.1. AIMS OF THE STUDY

The aims of the study were as follows :

3.1.1. To formulate a model of visual-perceptual-motor function of relevance to occupational therapy and to verify this model through research and to adjust the model if nescessary.

3.1.2. To design and construct a battery of tests for the evaluation of Visual-Perceptual-Motor (V-P-M) function described in the model for the adult who has suffered a Cerebral Vascular Accident (C.V.A.) using existing test material if appropriate. Ultimately only one existing test was used (i.e. the test of Visual Motor Integration by Beery and Buktenica 1967 and Beery 1982). All other tests were designed and constructed by the candidate.

3.1.3. To follow the eleven necessary steps leading to the standardisation of the test battery. This included obtaining data on a normal group of individuals in the age groups for which the test battery was intended.

3.1.4. To try out the tests formulated on a group of patients in order to demonstrate their value in the assessment of V-P-M function in the C.V.A. patient and to determine the way in which the battery could be used to diagnose certain V-P-M syndromes.

For the adequate testing of C.V.A. patients by the occupational therapist it is imperative that standardised measures of V-P-M performance be provided. Certain steps need to be adhered to if a test is to meet the rigorous requirements of standardisation and the research design, therefore, follows the step by step process necessary for standardisation (Lyman 1971, Tuckman 1975, Huysamen 1980 and Thorndike 1982).

3.2. SUMMARY OF THE RESEARCH DESIGN

3.2.1. Materials

A battery of fifteen tests for the evaluation of V-P-M function in the adult C.V.A. patient was constructed. This battery of tests is described in Chapter Four and Appendix 1.

The fifteen tests are as follows:-

Form Matching Test.

Test of Visual-Motor Integration (Beery and Buktenica 1967).

Part-Whole Test.

Perceptual Constancy Test.

Figure-Ground Test.

Arrows Test.

Dot-Line Test.

Spatial Relationships Test.

Position in Space Test.

Block Construction I Test.

Block Construction II Test.

Stick Construction I Test.

Stick Construction II Test.

Design Construction Test.

Colour Matching Test.

3.2.2. Subjects

The subjects tested for the purposes of this study were divided into five groups as follows:-

3.2.2.1. Fifteen occupational therapists (Group One) whose ratings of the various tests and their content provided the data for the content validity of the tests in the battery.

3.2.2.2. Nine C.V.A. patients (Group Two) whose test results and ability to understand the tests' instructions contributed to the formulation of the final test format of each test in the battery.

3.2.2.3. One hundred and eighteen normal individuals who had no neurological dysfunction (Group Three) constituted a control group. The results from testing Group Three on the V-P-M battery provided the normative data on which the psychometric characteristics of the tests were analysed.

3.2.2.4. Thirty three right hemisphere C.V.A. patients (Group Four) whose results on the V-P-M test battery demonstrated the type of problems experienced by C.V.A. patients in the field of V-P-M function.

3.2.2.5. Six normal individuals (Group Five) who were tested by each of three examiners to provide the data on inter-rater reliability.

Thus 181 individuals in total were tested with the V-P-M Battery.

3.2.3. Method

In order to fulfil the aims enumerated in Section 3.1. it was necessary to follow the twelve steps for drawing up a standardised test (Huysamen 1980, Tuckman 1975). In summary these included:

- Defining the domain that the test instrument was to measure and the individual skills constituting the domain.

- Describing the uses planned for the test.

- Listing the test constraints.

- Compiling the tests for each skill.
- Constructing these tests.
- Having the content evaluated by experts.
- Administering the test to a small group of subjects for whom the test is intended.
- Finalising the preliminary test format.
- Testing groups from the population for which the test is intended.
- Analysing the psychometric characteristics of the test.
- *Forming a new* test.
- Standardising the corrected test. This step was not completed for this thesis.

The twelve steps are described in detail in sections (3.3. - 3.14.).

3.3. STEP 1 : DEFINITION OF THE TEST DOMAIN

It was important to provide a definition of the function that the test was intended to *measure* i.e. V-P-M function, so that the user could be sure of the skills that were to be evaluated.

Definition of Visual-Perceptual-Motor Function:

This is the processing and interpretation of visual stimuli as well as the selection and use of the correct motor pattern in response to those visual stimuli when appropriate.

Each test was, therefore, designed to evaluate visual processing as well as the links between the visual process and motor performance, hence the term Visual-Perceptual-Motor function. The details of the model of V-P-M function are described in Chapter Four.

3.4. STEP 2 : USES PLANNED FOR THE TEST BATTERY

This step enumerated the subjects on whom the tests were to be used and the type of decisions to be made as a result of testing.

3.4.1. Subjects

The purpose of the study was to design a battery of tests to evaluate adult individuals between the ages of twenty and seventy-nine years who had suffered a C.V.A. and had resultant V-P-M dysfunction.

3.4.2. Type Of Decisions To Be Made As A Result Of Testing

The test battery itself was designed to determine :

- whether the patient had decreased V-P-M function as a result of a C.V.A.

- the type of dysfunction present. The strategies for therapy would then be based on the type of dysfunction. Thus, the tests were designed for the purpose of identifying particular deficiencies in a patient's performance for remedial action to be taken accordingly. The test battery, therefore, was planned as a diagnostic instrument, as well as a means of monitoring progress in therapy. Chapters Sixteen to Twenty describe the types of dysfunction that can be diagnosed. A discussion of the strategies for treatment is beyond the scope of this thesis.

3.5. STEP 3 : TEST CONSTRAINTS AND VARIABLES

This step described the practical limitations that apply to the administration of the tests such as time, and the way in which the variables are controlled. This was a necessary procedure as it helped the tests' developer plan the items of each test and the administration instructions.

3.5.1 Time

The final version of the test battery should be brief, provide relevant and accurate information and last approximately two sessions of three quarters of an hour each. At least two sessions are recommended as this enables the therapist to evaluate the subject's ability more accurately. She should, however, make herself known to the subject and obtain any background information prior to the test sessions so that the subject would be at ease with the therapist during testing.

3.5.2. Administration

It was planned for the tests to be administered individually to subjects as the observation of test taking behaviour provided additional information regarding the patient's problems. In addition, individual administration meant that the tests did not have to be confined to pencil and paper tasks where the scoring took place after completion, but could include construction tasks where the scoring would be done as the tests progressed. This, however, meant that the correctness of the responses should be evaluated by the tester within a short space of time and necessitated the formulation of strict criteria for passing or failing an item. The testers had to be trained in the evaluation of these criteria to eliminate subjective scoring as far as possible.

3.5.3. Control Of Variables (Independent)

A set method of administration was planned to control the variables that Tuckman (1975) enumerated as influencing test taking behaviour. (Test instructions are provided in Appendix 1).

Sections 3.5.3.1. and 3.5.3.2. deal with variables that may be <u>lasting</u> or <u>permanent</u> and Sections 3.5.3.3. and 3.5.3.4. enumerate variables that may be of a <u>temporary</u> influence on the individuals' behaviour and ability to execute the tests. Section 3.5.3.5. deals with the variables related to the <u>tests</u> themselves.

3.5.3.1. Lasting general characteristics of the individual.

- **General Skills** e.g. reading. This variable was controlled as the subject was not required to read the tests' instructions.

- ***General ability to comprehend instructions.*** In giving instructions for the completion of the tests emphasis was placed on the understanding of the instructions, therefore, these might be repeated using different words.

- **Ability to solve problems of the type presented in the tests.** It was not possible or desirable to control the variation in ability to solve the types of problems encountered in the tests. Obtaining information on normal individuals prior to using the tests on patients took the variation in ability into account and it allowed patient's performance to be compared with that of the normal population

- **Attitudes operating in the test situation** e.g. self confidence. The instructions enabled the tester to ensure that the subject did not become discouraged if he thought that he was not performing well (see general test instructions - Appendix 1).

3.5.3.2. Lasting Specific Characteristics of the Individual.

- **Knowledge and skill required for the particular problems presented in the tests.** As these are tests of visual-perceptual-motor function, the test items were designed to minimise dependence on previous knowledge. However, it has been said that the elderly perform better on familiar tasks (Krauss et al 1980), therefore some test items were constructed with words, letters and numbers.

- **Attitudes and reactions to a particular test.** The type of tests compiled did not deal with emotions or attitudes.

3.5.3.3. Temporary general characteristics of the individual.

- **Health, fatigue and emotional strain.** Testing was only done if the individual felt able, was not too tired and was prepared to continue.

- **Motivation and rapport with examiner.** This was an important aspect of testing and time was spent informing both normal subjects and patients of the purpose of testing, finding out about their problems and putting them at ease (see test instructions Appendix 1).

- **Effects of heat, light, ventilation, etc.** All individuals were tested at a table of a size which accommodated the test materials comfortably and which allowed the subject and the examiner to sit alongside each other (see test layout Appendix 1). All subjects were made as comfortable as possible and were tested in a well lit area.

- **Level of practice or skills required by the test.** No subject had received any practice prior to the testing and no patients had received any occupational therapy for perceptual testing or training which might have given them practice in the type of activities used in the test, prior to evaluation (see selection of patients i.e. group four).

3.5.3.4. Temporary specific characteristics of the individual.

- **Changes in fatigue and motivation produced by the particular test** e.g. discouragement resulting from failure. The tests' instructions attempted to deal with this by informing the normal subjects that there was no "right or wrong way to do the tests". Patients were encouraged to complete as many items as possible. If the examiner observed constant failure the test was discontinued. It was impossible to eliminate some degree of discouragement in the drawing and construction tests as subjects could evaluate their performance against a model.

- **Fluctuations in attention, concentration and judgement.** This could only be controlled by stopping the evaluation when the subject's attention could be seen to be waning and continued at a future date. Fatigue was controlled by giving the subject a ten minute break and a cup of tea half way through the test or by completing the test the next day.

- **Luck in selection of answers by guessing.** This factor could not be controlled in the multiple choice tests. Scoring took this into account by subtracting the total of incorrect answers from the total of correct items.

3.5.3.5. Control of variables within the tests.

- **Number of test items.** Very short tests provide limited examples of performance on which to judge the individuals' ability, whereas long tests produce fatigue and inattention. Long tests however, minimise the effects of guessing and thus improve test reliability. The preliminary test described in this thesis is of necessity long, as items needed to be analysed and refined. At least twice as many items as the number to be included in the final test should be included in the preliminary version (Huysamen 1980). However, in compiling the test, the time constraints listed in step three were taken into account.

- **Quality of test items** i.e. components of each test. Poor items might be too easy or too difficult. Item analyses were conducted so that the poor items of each test could be eliminated from the final form (see analysis of the psychometric characteristics of the test, item 3.12, step 10).

- **Conditions of test administration.** See item 3.5.3.3.

- **Conditions for scoring.** The error variation in scoring was reduced through formulating set criteria for scoring (see Appendix 1).

Thus it can be seen that variables concerning behaviour as well as test items were controlled as far as possible.

3.6. STEP 4 : TEST BLUE-PRINT

The test blue-print is similar to that of defining the test domain. Blue-print was the term used for the detailed description of each skill evaluated by the test battery. A description of each skill is to be found in the chapter describing the model of V-P-M function (see Chapter Four).

3,7. STEP 5 : THE TEST BATTERY

A battery of fifteen tests was devised to isolate all the elements discussed in the Model of V-P-M function which took the previous four steps into consideration. They are described in Chapter Four and Appendix 1. Performance on the tests required a combination of visual-motor skills as used in the test of Visual Motor Integration (Beery and Buktenica 1967, Beery 1982) and visual perceptual skills requiring a minimum of motor action as seen in a matching type test. An attempt was made to test all aspects of the model with tasks that required motor ability as well as tasks that required minimal motor ability (i.e., motor free and motor reduced tasks). Memory is a critical factor in perceptual task performance (Krauss et al 1980) and was eliminated as far as possible.

In the administration of all tests, time was recorded although the emphasis was on accuracy not speed.

All aspects of visual perception were measured through the use of meaningful symbols such as letters and numbers as well as meaningless symbols as it is documented that older subjects perform better with familiar material (Krauss et al 1980, Birren and Schaie 1977). Cognisance was taken of the fact that verbal and non verbal material tapped the function of different parts of the cortex (Concha 1982) i.e. the subject was not required to interpret the verbal cues.

3.8. STEP 6 : EVALUATION OF THE CONTENT BY EXPERTS IN THE FIELD

The appropriateness of each test for measuring the skills named in the model was evaluated by administering the tests to qualified therapists (group one) with more than one year's experience and who had had experience in the field of perception testing and treatment. This part of the study helped to determine the validity of the material for the assessment of the defined function, and also helped to resolve administration problems and ambiguities in the test instructions. Some modifications to the test instructions, therefore, took place after this stage.

3.8.1. Group One

This group consisted of fifteen occupational therapists known to the candidate with more than one year's experience in the profession and who consented to give the time necessary to complete the tests and to evaluate their contents. The occupational therapists evaluated and rated each test on a 0-5 scale according to their assessment of how well each test measured the defined skills. A special form was used for this purpose. Zero on this scale indicated that a test did not evaluate a particular perceptual skill, whereas five indicated that a test evaluated a particular skill extremely well. Intermediate values (1-4) indicated varying degrees of the test's propensity to evaluate that skill.

3.8.2. The Method Of Analysing The Results Of Testing Group One.

The scores of the fifteen Occupational Therapists were extracted from the raw data to give the values for each perceptual modality. These were set out in tables. The mean score for each perceptual modality was calculated for each test. The mean of the series 0-5 is 2,5 and is also the median value. Half the median value is 1,25 which is also the quartile value. Thus 3,75 is the upper quartile. A median score of 2,5 indicated that a particular test provided an average assessment of a particular perceptual modality. A mean score greater than 2,5 indicated a better than average assessment and a mean score of 3,75 and more indicated a very much better than average assessment.

The results of this procedure are reported as part of the content validity studies on each test in the battery. The actual results obtained on each test by the therapists were not used in any other way as it was felt that they were familiar with the tasks. Thus only their ratings of the content were important.

3.9. STEP 7 : ADMINISTERING THE TEST TO A SMALL GROUP OF SUBJECTS FOR WHOM THE TEST IS INTENDED (C.V.A. PATIENTS)

This step was taken in order to solve problems with administration before the test was administered to a larger sample. Nine C.V.A. patients (Group Two) were evaluated on the V-P-M battery for this purpose.

3.9.1. Group Two

The criteria for the selection of this group of subjects were that they had suffered a C.V.A. and were referred to occupational therapy at the Johannesburg Hospital. Patients with both left and right sided brain lesions were evaluated in order to determine the suitability of the instructions and test items for all types of patients and to eliminate as many ambiguities as possible.

3.9.2. The Methods Of Analysing The Results Of Testing Group Two

As a consequence of the results of testing group two, some small changes to the test items were made and the test instructions and administration procedures were finalised. For example, it was decided to leave the stimulus cards in front of the patient during the execution of each item of the Perceptual Constancy test to eliminate memory.

The scores obtained by these nine patients were not included in the study. *These patients were tested for the purposes of altering and* refining test instructions and test items only.

3.10. STEP 8 : PRELIMINARY FORMAT OF THE TEST BATTERY

As a result of the work completed in steps 1-7, the preliminary test format was finalised and printed. (See Appendix 1).

3.11. STEP 9 : TESTING GROUPS FROM THE POPULATION FOR WHOM THE TEST IS INTENDED AND THE TESTING OF HYPOTHESES I AND II

As it was intended to use the test to evaluate the V-P-M function in C.V.A. patients between the ages of twenty and seventy-nine years it was important to obtain information on normal individuals and their ability to execute the test items, prior to using the test on patients. This step was necessary so that dysfunction in patients could be determined by comparing their results with those of the normal individuals. Two groups (i.e. group three and four) were required to provide the necessary data on the tests and to test the following two null hypotheses.

3.11.1. Null Hypothesis I

If normal adults between the ages of twenty and seventy-nine years are tested on the Tests of Visual-Perceptual-Motor function, increased time to complete the tasks and decreased ability will not be found with advancing age in visual-perceptual-motor function.

3.11.2. Null Hypothesis II

That these same Visual-Perceptual-Motor tests will not demonstrate problems of visual-perceptual-motor function in patients who have suffered a right cerebral hemisphere vascular accident resulting in left hemiparesis/hemiplegia.

3.11.3. Group Three

3.11.3.1. Criteria for the Selection of Group Three

To test the first hypothesis the V-P-M test battery was administered to normal men and women between the ages of twenty and seventy-nine years (Group Three). The criteria for the selection of subjects in this population were as follows. The subject should:-

- have no present or past neurological problem.

- have no history of a psychiatric disorder.

- have an education and English language skills at the standard six level and above, therefore, have an IQ within the low normal to above normal range.

- be willing to co-operate throughout testing.

- be Caucasian. A specific ethnic group was selected for this stage of the standardisation process in order to eliminate the variables associated with culture. In apartheid South Africa, the poor schooling of Black South Africans made education a variable of importance. Once the test items of the V-P-M battery have been refined and the final test constructed i.e. when the steps planned for this thesis have been completed, a random sample of individuals from the total South African population will be tested to provide the normative data.

- be drawn as far as possible from the general Caucasian population and were in old age homes in the Johannesburg area, relatives and friends of the tester, relatives accompanying patients to the out patients facility of the occupational therapy department, Johannesburg Hospital.

- have no visual problems that could not be compensated through the use of spectacles.

- have hearing good enough to be able to hear the test instructions.

- be completely independent in all activities of daily living and in the management of their lives.

This group consisted of the following numbers in each age category:-

YEARS	MALE	FEMALE	TOTAL
20-29	11	9	20
30-39	9	12	21
40-49	5	14	19
50-59	6	13	19
60-69	3	16	19
70-79	5	15	20
	n = 39	n = 79	118

This *normal group* was not randomly selected but was a sample of convenience and therefore was biased in favour of co-operative subjects. However, as no subjects were aware of the type of material to be used in the test prior to the testing no bias was shown in the selection of individual skill in test performance. The educational background of the subjects varied from standard eight to matric. Some had received post high school training and education at Technikons and Universities.

3.11.3.2. Differences between men and women.

To determine whether men and women performed differently in the skills evaluated by the test battery, the scores for time and accuracy on all tests for men and women were compared. The U Test of Mann-Whitney (*a non parametric* test for unequal sample sizes) was used to determine the differences in the time scores, while Fishers' Exact test was used for the accuracy scores. No statistical difference was found in the scores obtained by men and women except in the following tests:-

- VMI Time Scores - twenty to twenty-nine year olds ($P=0,0127$, $P<0,05$) and seventy to seventy-nine year olds ($P=0,0093$, $P<0,01$);

- Stick Construction II Time Scores - fifty to fifty-nine year olds ($P=0,0235$, $P<0,05$);

- Colour Accuracy Scores - fifty to fifty-nine year old ($P=0,0173$, $P<0,05$).

The statistician felt that, as there were so few differences it did not warrant carrying out separate analyses for men and women. Therefore their scores were combined for all analyses.

3.11.3.3. Methods of Analysing the Results of Testing Group Three

a) Accuracy Scores

To determine the degree to which the data could be assumed to be normally distributed the following calculations were performed:

- the means and standard deviations of the accuracy scores and time scores were calculated for each age group i.e. twenty to twenty-nine years, thirty to thirty-nine years, forty to forty-nine years, fifty to fifty-nine years, sixty to sixty-nine years and seventy to seventy-nine years.

- the co-efficient of skewness was calculated and used to determine whether the data were skewed and the degree to which it was skewed in order to determine the type of statistics to be used.

To determine if there were any differences in the performance of the older group from that of the younger group, tests for significant differences were carried out between the various age groups (see above), i.e. hypothesis I was tested. Non parametric tests were used as the data could not be assumed to have a normal distribution. Thus value counts were performed and 2 x 2 tables constructed in which the figure equal to and less than the fifty per cent level of accuracy of each test was calculated. When the minimum expected value was equal to or greater than five then the Chi-square Test was used after applying Yates' Continuity Correction (Afifi and Azen 1959). When the minimum expected value was less than five then Fischer's Exact probabilities were calculated. If the null hypothesis of no significant difference among the age groups was rejected, further analyses were conducted to determine if each group differed from the other, or if two of them were similar, both differing from a third. According to the differences found, the data were regrouped with certain ages placed together.

To enable a comparison to be made between the scores of the normal group of individuals and that of the patients, percentile ranks were computed for each of the age groups (Downie and Heath 1974).

Standard z scores of -1.0 are accepted as the lower limit for normality in standardised tests as approximately eighty-five percent of the normal population obtain scores above this point. As the fifteenth percentile is roughly the equivalent of the z score of -1.0 this percentile rank was used to determine whether a patient's score could be considered defective.

b) Time Scores

To determine the dispersion of the data, the mean and standard deviation for the time taken to complete each test were calculated in the previously mentioned age groups.

To determine whether the time scores obtained by the various age groups differed statistically from one another, One Way Analysis of Variance was used, as the data were found to be normally distributed. Levine's test for equal variance was used to establish the between and the within group variance. If Levine's test was not significant (i.e. $P > 0{,}05$) then the F Test was used to determine whether the means differed from one another. If this were the case the pooled variance t with the Bonferroni correction for multiple comparisons (Neter and Wasserman 1974) was used to establish where the differences lay. Thus to be significant at the 0,05 level, the P value had to be less than 0,00333 and when the scores for the different ages were combined, 0.00833. If Levine's test for equal variances were significant ($P < 0{,}05$) i.e. the variances between the groups were found not to be equal, the Welch Test for unequal variances was used to establish whether there were any differences between the means of the groups. When a difference was found the separate variance t with the Bonferroni correction for multiple comparison was used to establish exactly where the differences lay.

According to the differences found, the data were regrouped with certain ages placed together.

To enable a comparison to be made between the scores of the normal group and that of the patients, percentile ranks were computed for each of the age groups.

3.11.4. Group Four.

To test the null hypothesis II the battery of V-P-M tests was administered to a group of patients who had suffered a right hemisphere C.V.A. and whose ages ranged from twenty to seventy-nine years. Only patients with right hemisphere lesions were selected for this study as the right hemisphere is thought to play a more important role in visual perceptual function (Berent 1981). Patients with right hemisphere dysfunction were also selected to eliminate the variables associated with speech problems commonly found in left hemisphere lesions. The criteria for selection of patients in population four were as follows:

The subject should:-

- have had a right sided stroke as determined through medical diagnosis. The exact location of the lesion was not considered.
- have received no prior occupational therapy testing or treatment in the sphere of visual-perceptual-motor function.
- be orientated for person, place and time.
- be well enough to spend at least one hour at a time in the Occupational Therapy Department of the Johannesburg Hospital.
- have had their stroke not more than six months prior to being tested.
- give consent to undergo the battery of tests, which, as was explained, was a longer and more detailed version of a normal therapy assessment.

The goal was to obtain ten patients in each of the age categories twenty to thirty-nine years, forty to fifty-nine years, sixty to sixty-nine years and seventy to seventy-nine years. These age groups were determined through the statistical analysis of the data obtained on the normal individuals.

Patient numbers were as follows (all had a right sided lesion and were left hemiplegic):-

AGE	MALE	FEMALE	TOTAL
20-39	3	4	7
40-59	7	3	10
60-69	5	5	10
70-79	4	2	6
	n =1 9	n = 14	n = 33

The patient group consisted of thirty-three patients. Many difficulties were experienced by the candidate in obtaining suitable patients for evaluation, such as :-

a) the early discharge of patients from the Johannesburg Hospital where the work was carried out. Due to the chronic bed and nursing shortage, patients were discharged as soon as their condition was stabilised and the initial bed rest period was completed. This occurred regardless of their need for therapy.

b) the difficulty in obtaining only left hemiplegic (right hemisphere C.V.A.) patients at the appropriate time.

c) the severity of the defects in the seventy to seventy-nine year old age group prevented some of the patients from completing their tests.

3.11.4.1. The Methods of Analysing the Results of Testing Group Four.

The comparison of the normal group of individuals (group three) with that of the patient group (group four) is discussed as part of the construct validity for each test (see Chapters Five to Fifteen). To determine whether the accuracy and time scores obtained by the normal group and the patient group differed statistically from each other, One Way Analysis of Variance was employed. Levine's test for equal variance was first used to establish whether the variances differed between the groups. If Levine's test was not significant i.e. no differences in the variances were found, the F Test was used to determine whether there was a statistical difference between the mean scores obtained by each group. If Levine's test was significant and the variances were found to be unequal, the Welch Test for unequal variances was used to test for significant differences between the two age groups.

The individual patient scores were converted to percentile scores to obtain an indication of the value of these scores in making a diagnosis. Each chapter, therefore contains information on the number of patients who were thought to experience a problem with the specific V-P-M skill measured by that test.

The results of the observations made during the testing of the patients were discussed for each patient in Appendix 3. These results and observations were also used to provide examples dysfunction (Chapters Five to Fifteen) and to describe the types of V-P-M dysfunction in Chapters Sixteen to Twenty.

3.12. STEP 10 : ANALYSIS OF THE PSYCHOMETRIC CHARACTERISTICS OF THE TEST

In order to determine whether the test was an adequate measure of V-P-M function the results of testing the 118 normal subjects were analysed with a view to determining the psychometric characteristics of the test.

3.12.1. Item Analysis

An analysis of the difficulty level and discriminatory level of each item in each of the fifteen tests was carried out in order to help to determine the final version of each test. The item analysis is used to select items giving a spread of difficulty values in the final test format. The items were also analysed to see if they discriminated between normal subjects and patients. Items that were too easy and that showed no ability to discriminate between those who individuals who could do the test and those who could not do the test, should be eliminated from the final test form.

Two types of item analysis were carried out i.e. the Difficulty Index and the Discrimination Index.

a) <u>The Difficulty Index of each item</u> (i.e. the difficulty level). This is the proportion of individuals who answer the item correctly. This was calculated for all tests except the speed test i.e. the Position in Space (PIS) test. In the PIS test some items were not answered due to time constraints and therefore the difficulty index could not be calculated. A difficulty index can vary between 0 when no subject answers the item correctly, and 1,00 when everybody answers correctly (Huysamen 1980). This index is estimated by p, the population of the testees who have answered the item correctly (Huysamen, 1980). Thus the higher the difficulty p, the easier the item.

b) <u>The Discrimination Index of each item</u>. This is the extent to which the item discriminates between the testees who can do well on the test and those who can not. Two ways of determining the discrimination index were used in this study i.e.:-

- testing the significance of the difference between two proportions. In this technique, the percentage or proportion of individuals who answer an item on a test correctly in the group containing the high scores is tested against the proportion in the group obtaining low scores (Downie and Heath 1974). If the difference is a significant one, the item is accepted as one which discriminates. This technique does not, however, reveal how well each item discriminates. Fishers' Exact Test was used to analyse the data. When the scores for each item were greater than 1 (1 = correct, 0 = incorrect) cut-off points were calculated.

- a correlational approach in which *"a correlation co-efficient is computed that shows the relationship of the responses to the* *total test score"* was used. This approach shows how well the item is doing *"what the test itself is doing"* (Downie and Heath 1974:251). Flaniagan's method of reading the estimated bi-serial co-efficient of correlation between the item and the total test score when the item has been restricted to the lowest and the highest 27% of the total score distribution, was used. The responses of the individuals in the group scoring the highest scores and of those in the group scoring the lowest scores were tallied and the proportions obtained. This step was performed by computer. The correlation co-efficients were then obtained by entering the value obtained for the upper 27% correct on the y axis and the value for the lower 27% correct on the x axis of Figure 3.1. The point of interaction is the estimated value of r.

As *"Flanagan's r is often taken as an estimate of the Pearson r,"* the significance of the Pearson r was used in testing the significance of Flanagan's r (Downie and Heath 1974:255).

Item difficulty and item discrimination levels were also calculated on the data obtained from group four i.e. patients. Differences in the levels of difficulty and discrimination were compared between group three and group four using Fishers' Exact Test.

3.12.2. Test Reliability.

"Each individual score on a test is considered to be made up of two parts, a true score component and an error score component" (Downie and Heath 1974:236). However, when a test is administered, many factors enter into the error component of an individual's score e.g. fluctuations in the individual's concentration, emotional status, health, misreading the item, guessing and the temperature of the surroundings to name but a few (see variables listed in section 3.5.3.). The size of this error component is related to the reliability of the scoring device. The smaller the error, the more reliable the scoring instrument. Reliability coefficients were calculated for each test for

each age group between which no statistical difference was found in the accuracy scores. Reliability coefficients were calculated as follows:-

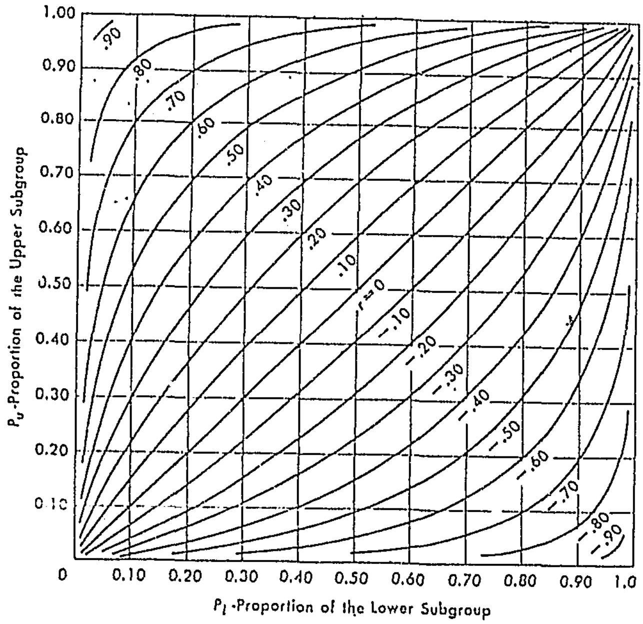


FIGURE 3.1 FLANAGAN'S METHOD

a) The Split-Half Reliability. To determine the consistency of each test with regard to content sampling, the split-half reliability procedure was used. In this method the scores for all the odd and all the even items on the test were totalled. A Pearson product-moment correlation co-efficient was then computed between the two sets of scores. This was calculated by the computer, the resulting reliability co-efficient was corrected for the shortening effect of splitting the test in half by using the Spearman Brown formula. This calculation was done by hand using the following formula:-

$$rtt = \frac{2\,roe}{1 + roe}$$

where rtt = reliability of test

roe = reliability co-efficient obtained by correlating the scores on the odd items with scores on the even items (Downie and Heath 1974:238)

The level of significance of the correlation co-efficient was also determined. If P was <0,050, or less there was a significant correlation between the two halves of the test.

b) Kuder-Richardson Formula 20. To determine the homogeneity of the test or the item consistency of the test, the Kuder-Richardson Formula 20 was used. In this method, the pass fail score for each item is compared with the total test score.

This co-efficient is calculated using the difficulty value of each item and was calculated by hand using the following formula:-

$$rtt = \frac{k}{k-1}\left[1 - \frac{\sum pq}{S^2}\right]$$

k = number of items on test

S^2 = variance of test

p = difficulty value

q = 1 - p

pq = product of p and q (Downie and Heath 1974:240)

"In general, reliability co-efficients tend to be high 0,9 or above - however - there is no hard or fast rule that says that any reliability has to be a certain size before any test or measuring instrument can be useful" and co-efficients of 0,80 or above are traditionally accepted as evidence of adequate test reliability (Downie and Heath 1974:240). The adequacy of the reliability was, therefore, judged according to the degree to which it met the above criterion as well as its level of significance.

c) Inter-Rater Reliability

- Between Examiners.

Three examiners were responsible for the testing of population three i.e. the normal or control group. Permission was given by the Post Graduate Committee of the Medical School of the University of the Witwatersrand for examiners other than the candidate to test this population. The other testers were trained in the administration, provided with test administration instructions and a duplicate test kit. The number of normal subjects evaluated by the three testers were as follows:-

Examiner A (the candidate)	=	56 subjects
Examiner B	=	40 subjects
Examiner C	=	16 subjects

This was thought to be desirable to :-

Firstly, establish inter-rater reliability - the test would be of little value if it could not be reliably administered by therapists other than the candidate. To establish the inter-rater reliability, each of the three examiners initially evaluated the same six individuals.These six subjects formed Group Five.

Group Five - consisted of six normal individuals between the ages of twenty and fifty-five. None of these individuals had been previously subjected to the tests.

A random table for testing the six individuals was drawn up by the statistician and a dice was used to determine which individual (who was allocated a letter of the alphabet) was to be seen by each examiner on the first occasion. Thereafter, the table was followed (see Table 3.1). The testing took place at the same time, on the same day for three consecutive weeks.
Friedman's two-way analysis of variance was used to determine if there were any statistical differences in the scores obtained by each examiner on each test.

TABLE 3.1 PROCEDURE FOR ESTABLISHING INTER-RATER RELIABILITY

	EXAMINER 1	EXAMINER 2	EXAMINER 3
Week 1	C F	A D	B E
Week 2	B D	C E	A F
Week 3	A E	B F	C D

The only difference found was in the time scores obtained on the Form Matching Test (P < 0,05) (see Chapter Five). The statistician could find no realistic basis for this difference as other tests were timed in a similar fashion and showed no difference. He, therefore, recommended that data obtained by the three examiners be combined for analysis on this and all other tests in the V-P-M battery. Subsequent to establishing the inter-rater reliability the subjects belonging to group three were tested by the three examiners.

Secondly, reduce the amount of time taken to collect the data. Due to the fact that this preliminary version of the test takes approximately one and a half to four hours to administer and a further hour to score, the task of collecting data from normal subjects proved to be very time consuming.

- Between Scorers.

Although the tests were administered by three examiners, the paper and pencil tests such as the test of Visual-Motor Integration and the Dot-Line tests were scored by the candidate. The Dot-Line test was scored by measuring the lines with a ruler and a scoring transparency and thus the criteria for pass/fail were objective. However, scoring the Test of Visual Motor Integration involved a degree of subjectivity. Inter-rater reliability in scoring the Test of Visual-Motor

TABLE 3.1 PROCEDURE FOR ESTABLISHING INTER-RATER RELIABILITY

	EXAMINER 1	EXAMINER 2	EXAMINER 3
Week 1	C F	A D	B E
Week 2	B D	C E	A F
Week 3	A E	B F	C D

The only difference found was in the time scores obtained on the Form Matching Test ($P < 0{,}05$) (see Chapter Five). The statistician could find no realistic basis for this difference as other tests were timed in a similar fashion and showed no difference. He, therefore, recommended that data obtained by the three examiners be combined for analysis on this and all other tests in the V-P-M battery. Subsequent to establishing the inter-rater reliability the subjects belonging to group three were tested by the three examiners.

Secondly, reduce the amount of time taken to collect the data. Due to the fact that this preliminary version of the test takes approximately one and a half to four hours to administer and a further hour to score, the task of collecting data from normal subjects proved to be very time consuming.

- Between Scorers.

Although the tests were administered by three examiners, the paper and pencil tests such as the test of Visual-Motor Integration and the Dot-Line tests were scored by the candidate. The Dot-Line test was scored by measuring the lines with a ruler and a scoring transparency and thus the criteria for pass/fail were objective. However, scoring the Test of Visual Motor Integration involved a degree of subjectivity. Inter-rater reliability in scoring the Test of Visual-Motor

Integration (VMI) was therefore sought as consistency of scoring is important for the reliability of the test.

A therapist familiar with the children's version of the VMI was given the scoring criteria for adults. The tests were numbered as they were completed i.e. a single age group was not consecutively numbered. The therapist took every second test for scoring purposes and the scores obtained were compared to those obtained by the candidate. The t Test (Afifi and Azen 1959) was used to establish whether there was a statistical difference between the two sets of scores (for results see Chapter Six).

3.12.3. Test Validity

In this section, the question "Do the tests measure the skill that they are purported to measure and how well do they measure that skill?" was answered. Content and construct validity were thus investigated.

a) Content Validity

Huysamen (1980) named the following steps that should be taken to provide the content validity of any test as :-

- the universe to be tested must be defined (see Step 1 - Section 3.3.).
- the universe divided into its constituent subdivisions (see Chapter Four).
- a random selection of tests constructed for each subdivision (see Step 5 - Section 3.7.).
- the content evaluated by experts in the field (see Step 6 - Section 3.8.2.).
- the scores of each patient analysed with regard to V-P-M function and the areas of dysfunction determined (see Appendix 3). In addition V-P-M syndromes were described

and patients' test results were given as illustrative of the syndrome (see Chapters Sixteen to Twenty).

Thus the content validity of the test was established according to these criteria and discussed for each test.

b) Construct Validity

This *"refers to the extent to which a test's scores correlate with other scores to which we can expect it to be related"* (Tuckman 1975:235). A factor analysis also demonstrated construct validity (Downie and Heath 1974) and helped validate the V-P-M model.

The following relationships were investigated :-

- The relationship of score to age. This was done to determine whether the test was a valid differentiation of age related changes (See Section 3.11.3.1.).

- The relationship of score to diagnosis. This was done to determine whether the test was a valid way of determining the effects of brain damage resulting from a C.V.A.

- The communality and specificity of each of the tests were evaluated to help decide which of the tests should be retained in the battery. "*The proportion of the total variance that is due to factors that the test has in common with other tests is the tests communality, and the proportion due to factors specific to the test itself is the tests specificity....The communality of a test(* in a battery*) is computed as the sum of the squares of the factor loadings on that test...(*Aiken 1996:49). The communalities are then subtracted from their mean reliability's to give the specificity. The specificity is the percentage of the test's total variance that can be accounted for by factors specific to that test.

- The relationship among tests in the battery. This was done by means of a factor analysis which determined the types of tasks that cluster together into various skills. Firstly, a correlation matrix between all the test scores for the different ages combined was computed. This enabled the candidate to judge whether the tests had anything in common. Further to this, a factor analysis was carried out as this is a more objective means of evaluating the relationship. A factor analysis indicates the variances that two variables have in common. Each group of variances was then labelled by the candidate according to the dominant skill being measured and compared to the original model. A new model of V-P-M function was then constructed on the basis of the information provided. (See Chapter 21).

3.13. STEP 11 - THE FORMATION OF A NEW TEST

On the basis of the data collected the final test form will be constructed. Recommendations for this final version are included in Chapter Twenty-One.

3.14. STEP 12 - OBTAIN DATA FROM NORMING GROUP

In this step the final test battery is administered to a very large random sample representative of the population on which the test is to be used. Once this data is collected, the test norms are obtained and final reliability and validity studies carried out. This step is beyond the scope of this thesis.

CHAPTER 4

THE VISUAL-PERCEPTUAL-MOTOR MODEL

4.1. INTRODUCTION

The definition of the domain to be tested and the formulation of a theoretical model on which to base the tests are the first steps in the construction of a battery of tests (see Chapter 3.3.). This chapter describes the Visual-Perceptual-Motor function (V-P-M), the premises on which each aspect of the function is defined and the tests constructed to evaluate this function. The purpose of this chapter, therefore is to describe the framework within which the V-P-M function can be analysed and its components evaluated.

4.2. VISUAL-PERCEPTUAL-MOTOR FUNCTION

4.2.1. Definition

Perception is a broad and encompassing term which incorporates many different types of functions. Gibson (1950:12) stated that:-

> *"... perceiving things depends on first having sensations. Sensations are supposed to be the raw materials of human experience and perception the manufactured product".*

Sensations are only odours, sounds, touches and tastes; colours, shapes, objects and space depend upon perception.

Perception is an essential part of consciousness, it is that part which constitutes reality. The Webster's (1976) dictionary defines perception as:

> *"Consciousness, a result of perceiving: observation, a mental image; concept, awareness of the elements of the environment through physical sensation, physical sensation interpreted in the light of experience; quick, accurate and intuitive cognition."*

To occupational therapists, it has come to mean all of these things and much more, as the undisturbed perceptual process presupposes intact and unaffected means of obtaining the physical sensation, storing it, interpreting it and acting upon it (Concha, 1983).

Perception therefore is an awareness of the elements of the environment interpreted in the light of experience. Information is obtained by exploring the environment, by asking questions, by knowing where to look for information and by making assumptions. There is a vast difference between the real world and the perceived world and the study of perception is to observe the relationship between these two worlds (Hochberg 1978). Perception is thus a word used to describe a complicated phenomenon, a word which means the way the mind looks at and interprets things (De Bono 1976).

Once the information is processed it must be acted upon. However, in order to respond or "act upon" correctly, the information must be correctly processed. Although sensations are generally combined into perceptions, it is possible to separate the functions of the different input channels. This is essential if therapy is to be realistically directed.

In summary, perception is a dynamic process. It is a holistic term referring to meaning attached to an object, event or situation occurring within the spatial or temporal proximity of the individual. It is dependent not only upon the immediate situation, but also on the context in which the event occurs and upon past experience. Perception then is the process of discriminating among stimuli and interpreting their meaning. It intervenes between the sensory process on one hand and behaviour on the other. As all processes are interdependent the occupational therapist has come to include in the testing and treatment of perceptual difficulties the systems by which information is obtained, interpreted and acted upon (Concha 1983).

The sensory process can be measured by physical means and this has been done by the psychophysicists (Hochberg 1978). This is out of the scope of this thesis. This study is concerned with the interpretation and "acting upon" parts of the visual perceptual process. Visual-perceptual-motor function is therefore the processing and the interpretation of visual stimuli as well as the selection and use of the correct motor pattern in response to the stimulus.

4.2.2. Theories Of Visual Perception

As knowledge of perception increased over the last century, many theories regarding visual perceptual organisation were developed. These theories included those of the Empiricists and their structuralist philosophy of the Nineteenth Century in which perception was analysed into component sensations, with learning then taking place through a process of association. By analysing elementary sensations it was hoped that the complexities of human thought could be unravelled. This gave rise to the method of physical measurement of sensations. Nativistic philosophies in which perception was believed to be inborn, vied with those of the Empiricists. However, the Empiricist philosophy dominated early thinking on the subject and was only replaced by Behaviourism in the early nineteen hundreds. The Behaviourists argued that we could never know how other people experienced the world and thus should only observe their behaviour and their responses to various stimuli (Bruce and Green 1985). Both philosophies have merit and the 1960's saw the growth of the new "cognitive psychology" in which the stages between stimulus and response were defined (see Figure 4.1).

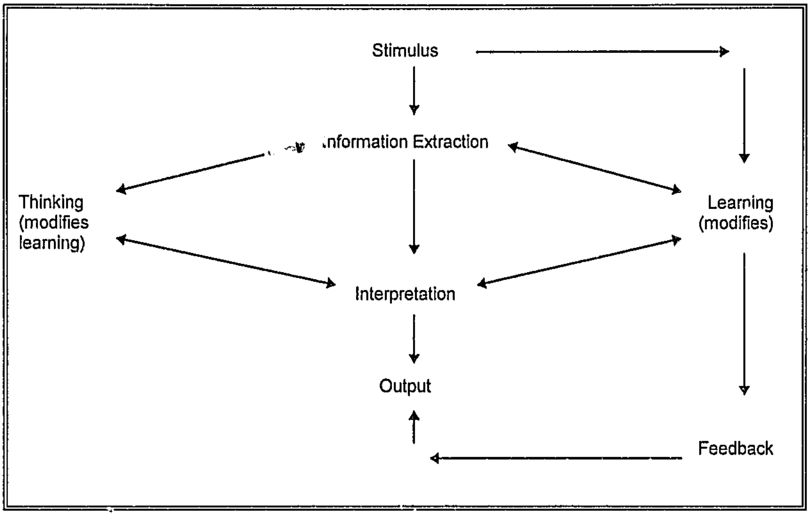


FIGURE 4.1 DIAGRAMMATIC REPRESENTATION OF THE PERCEPTUAL PROCESS

At the same time as the Behaviourists were developing their philosophy the Gestaltists such as Wertheimer showed that the analysis of perception into discrete sensations overlooked some important aspects of form and structure and they formulated a number of principles of perceptual organisation to describe how certain perceptions were more likely to occur than others, thus giving rise to the catch phrase "the whole is greater than the sum of the parts" (Bruce and Green 1985).

From the 1950's texts such as Critchley's famous book on parietal lobe dysfunction (Critchley 1953) began to appear and they described many of the findings regarding the perceptual functions which could be disturbed as a result of disease or injury. Much of the knowledge about the perceptual process stemmed from an analysis of the consequences of disruption to that process through brain injury. The first textbook to appear containing a reference to occupational therapy for patients who had perceptual problems appeared in 1960 (Jones 1960). This referred to adults who had undergone perceptual changes as a result of head injury. In South Africa, therapists tended to work more with children who experienced perceptual problems. The works of well known workers in the paediatric field such as Kephart (1960), Piaget (Flavell 1963), Gesell (1966) and Frostig (1966), greatly influenced the way in which perceptual and motor skills were defined and classified. The sequential acquisition of visual-perceptual-motor skills during the developmental process described by these authors also helped in demarcating the various aspects of visual perception.

This study deals with the consequences of disrupted visual perceptual activity in the adult. Accordingly, we are dealing with an individual who has stored many combinations of sensations and who has already passed through all the processes leading to the mature use of visual images, but whose perceptual function might have deteriorated with the normal ageing process. According to Smythies (1956) the recovery process after brain injury proceeds in a developmental sequence. However, the developmental sequence of childhood is reversed both in the elderly as well as in the brain injured individual (De Ajuriaguerra and Tissot 1969).

The knowledge of perceptual organisation obtained through the investigation of the theories of perception, the results of damage to the brain and the knowledge of the developmental process helped the candidate to understand perception and contributed to its definition.

4.3. DEVELOPMENT OF PERCEPTUAL SKILLS WHICH CONTRIBUTED TO THE MODEL

4.3.1. Form

It is believed that, from a very early age (six months) children are able to perceive form (Cratty 1979). It seems that this early perception is of vague masses without distinct contour, recognisable shape or definite localisation in space. These masses are characterised by cohesiveness and they exist as a totality with no parts and no relationship of parts. Thus, the child distinguishes globular form or form as mass. According to Kephart (1960), this undifferentiated mass is called cyncratic form, and it corresponds to the aspect of form recognition referred to as "silhouette" (Bruce and Green 1985). Infants differentiate faces in this way from the age of fifteen to twenty weeks (Cratty 1979).

Gradually more and more elements of form are distinguished. The early copying behaviour of children reveals how the details are differentiated out of the globular mass (Gesell 1966). Copying involves reproducing the elements to form an outline and thus the "overall contour" becomes distinguished. Once the elements are distinguishable they are recombined to form the "whole" again. This process is referred to as knowledge of the "Gestalt" by the Gestaltists (Bruce and Green 1985). The ability to combine units and sub-units into a perceived whole is developed by the age of seven years.

However, before outlines and parts can be differentiated the child must be able to distinguish horizontal form from vertical i.e. laterality (the awareness that there are two distinct sides to be seen) becomes externalised as directionality. This occurs at approximately two years of age (Cratty 1979).

Directionality is the directional location of a line in relation to the directional line of gravity. The child learns to transpose this line into a one or two dimensional form on paper. Reproducing elements to form an outline of a shape makes use of directional concepts.

In addition, to distinguish form adequately, the child must be able to separate out the object or shape from its surroundings (figure-ground perception). According to Smythics (1956) recovery from brain injury in the adult and the return of form recognition takes place in a sequence markedly similar to the aforementioned process.

The following recovery process is enumerated by Smythies (1956):-

a) First to return is the primitive Gestalt in the visual field i.e. the visual extension without form.

b) In stage two the field becomes differentiated so that qualitatively different stimuli are expressed to left or right without form and without boundary between light and dark.

c) Next the surface area is differentiated but without distinct dimensions. It appears neither horizontal nor vertical, but with the same dimensions on all sides.

d) In the fourth stage, the mass possesses a direction within the visual field.

e) The fifth stage is one of indeterminate forms. There is a definite form differentiation, extension or dimension of the mass which is perceived as somewhat long, small, horizontal etc.

f) Next comes the distinction of several separate masses within the visual fields.

g) Form is then perceived in the strict sense and there is distinction of straight lines, curves and geometric figures.

h) An eighth stage may be described in which three-dimensional *depth is added to a flat picture.*

Disturbance of visual form recognition in the adult is known as visual agnosia (see Chapter Seventeen).

As the recovery pattern described in the literature shows some of the characteristics of the development of form perception it was thought to be helpful to divide form perception into three categories for the purpose of this study i.e.

Form as Mass - where the individual is required to visualise a shape as a whole without necessarily recognising the different elements that go to make up the shape or form.

Overall Contour - in which the individual is required to analyse the elements that make up the form and to reproduce these to form an outline.

Units and Sub-units of Mass - where the individual is required to perceive different elements and parts and to see that they can be combined to make a distinct entity. This is also known as part-whole perception in which the Gestalt principle of closure may be utilised.

Recognition of form includes objects, animals, people (faces), letters of the alphabet, familiar buildings, etc. Directional orientation is also needed for the above skills particularly in the reproduction of the elements of the form, and is a skill evaluated in the battery.

4.3.2. Figure-Ground Perception

As already mentioned, the child must be able to separate out an object from the surroundings before form can be distinguished adequately (Cratty 1979). These elements of figure-ground perception must be present from an early age. However selecting figures that have been embedded into more complex designs can be done from the age of five or six years but only fully matures in the late teens, according to Cratty (1979).

The ability to analyse and synthesise information is inherent in all perceptual tasks and is particularly important in separating the figure from background (Benton 1979).

4.3.3. Perceptual Constancy

"The world we see is by and large a stable one" (Hochberg 1978:72). Thus Hochberg referred to the skill of perceptual constancy which maintained this stability. Hochberg (1978) described the perceptual cons'ancies as being lightness constancy, memory for colour, size and shape constancy. According to Cratty (1979) infants at the age of two years were capable of judgements that indicated that they had acquired both size and distance constancy e.g. they recognised an object as having a stable and consistent shape despite the way it was viewed with the resultant change in stimuli reflected on the retina. The development of "object constancy" as described by Piaget (Flavell 1963) is a precursor to the development of mature perceptual constancy. Perceptual constancy is thus defined as the ability to

perceive an object as possessing invariant properties such as shape, position, size and colour in spite of the variability of the presentation. Perceptual Constancy is a separate, clearly demarcated skill and is evaluated in the V-P-M battery described in this thesis.

4.3.4. Space

The evolution of the child's grasp of space, parallels very closely that of objects (Flavell 1963). In the early stages of development spatial orientation becomes possible as a result of visual, kinaesthetic and vestibular input. *"Only after the establishment of this combined activity, responsible for the processes of inspection, palpitation, head orientation and eye movements, can complex forms of reflection of spatial relationships develop"* (Luria 1980:172).

We have no direct information concerning spatial relationships in our environment, all our information comes to us through some clue which has to be interpreted. Space is therefore essentially a concept which is developed in the brain (Kephart 1960) and which has become defined by words and thus subject to the organising influence of the language system (Luria 1980). Piaget divided the acquisition of spatial concepts into three groups (Flavell 1963) i.e. the practical group, subjective and objective groups. These could be compared to the three spheres of space around an individual described by Kolb and Whishaw (1985) i.e. body space, grasping space and action space. Chaney (Heiniger and Randolph 1981) also formulated a model which corresponded with these three aspects and was called respectively the model of the body, the model of the environment and the model of the universe. Critchley (1953) referred to these space areas as that of touch, grasp and vision.

4.3.4.1. The Practical Group (Body Space)

The first group in which sensory motor actions with respect to objects in space is organised is the practical group. However, the child at this level cannot yet perceive himself to act and can not recognise the organisation inherent in his acts. His spatial world is confined to the major sensory motor spheres of activity i.e. there is a buccal space, visual space, auditory space and tactile space. These space spheres are "centred on his own body" (De Ajuriaguerra and Tissot 1969). Kolb and Whishaw (1985) called this Body Space and Chaney (Heiniger and Randolph 1981) referred to it as the Model of the Body. This model of the body is dependent on an internal awareness which arises initially from the use of postural reflexes.

Thus, the first aspect of a space structure to develop is that of body space, which is also called body schema. Body schema is a physiological and subconscious mechanism in which a self model is constructed on the basis of kinaesthetic and tactile input although other sensations also contribute. Body schema or body space does not form part of the model used for this study as its primary input is not visual. It is recognised, however, that body schema is a basis for other perceptions.

4.3.4.2. The Subjective Group (Grasping Space)

The next group to develop in Piaget's scheme of developing a space structure is the subjective group, this is a group of activities in which the individual becomes aware of his actions with respect to objects in space (Flavell 1963). The acquisition of spatial concepts in this group is based on visual search and the use of the hands to manipulate objects. These activities lead to knowledge of size and shape constancy and the discovery of perspective. Chaney described development at this stage as being dependent on the match between various sensory modalities (Heiniger and Randolph 1981).

The subjective group of space concepts corresponds to that space around the body described by Kolb and Whishaw (1985) as "grasping space" (Figure 3.2). Kolb and Whishaw described disorders of grasping space as being related to the problems of visual attention, eye movements, the ability to perceive more than one stimulus at a time, visual localisation and depth perception.

The activity of the individual within this practical stage helps to develop two space concepts i.e. Directionality and Spatial Relations.

a) Directionality

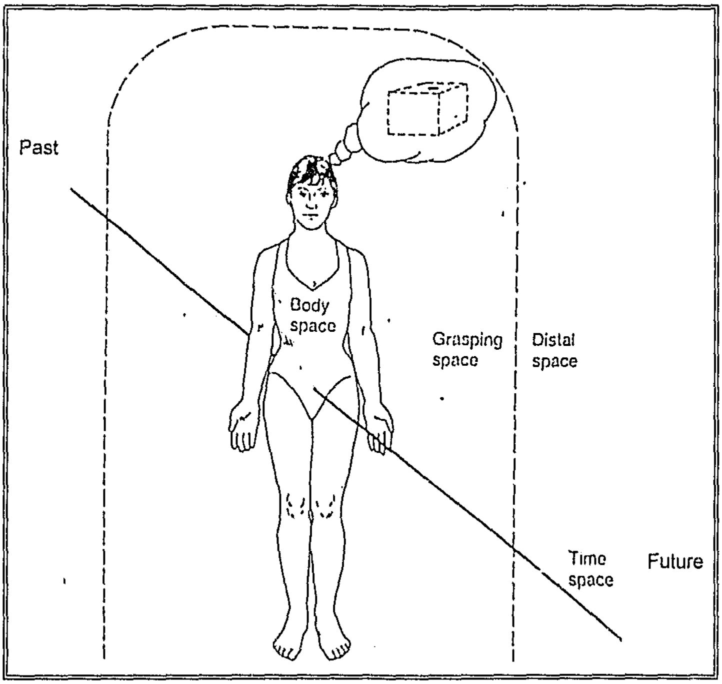
For purposes of this study directional concepts are defined as the ability to perceive the directional location of an object in relation to the line of gravity and to transpose this concept into one dimensional form on paper. It has already been pointed out that this skill is needed for the final stage of the development of form perception and therefore overlaps with form development.

b) Spatial Relationships

The concept of Spatial Relationships develops from that of directionality. Spatial Relations is the ability to perceive two or more objects in relation to oneself and in relation to each other. Spatial relationships and the various clues that are learned to judge these relationships such as movement, parallax, texture and gradient, provide the individual with perception of depth. Sequencing is also a function of spatial relations.

4.3.4.3. The Objective Group (Action Space)

Piaget's third and last stage (Flavell 1963) is that of the objective group of activities. This is a group of developments in which the child takes account of the movements of objects from one spatial position to another, as well as the movements of the body. He develops an organised space structure in which he has internalised displacements and movements. This is the stage in which objective Euclidean space concepts are formed and it corresponds to the space described by Kolb and Whishaw (1985) as Action Space and to Chaney's Spatial Model of the Universe (Heiniger and Randolph 1981). At this stage the child can intellectually manipulate precepts. "*Behaviour in action space requires that we first mentally manipulate objects or places in order to determine the best course of action, or that we draw or otherwise represent things and places in action space*" (Kolb and Whishaw 1985:580). The passage of time is an aspect of objective or action space.



<u>**FIGURE 4.2 MODEL OF SPACE (KOLB AND WHISHAW 1985)**</u>

It is interesting to note that Piaget (Flavell 1963) discussed the way in which the representation of space developed as a separate skill and emphasised that spatial representations were built up through the organisation of actions performed on objects and later internalised. Most spatial perceptual tests used representations in a two dimensional form to evaluate the aforementioned skills.

The aspects of action space are topographical orientation and spatial thought.

a) Topographical Orientation is the ability to perceive and remember the location of objects and places in space.

b) Spatial Thought is the ability to manipulate spatial concepts in the mind and has been termed position in space for the purpose of labelling the tests constructed for this thesis.

Thus, the third spatial concept to develop is that of objective Euclidean space in which the above mentioned groups are combined. In this way the geometric categories of extensity (i.e. vertical and horizontal concepts), location in 2D (i.e. location of points in the visual field corresponding to the X and Y co-ordinates of geometry), shape and form (corresponding to the abstract form of Greek geometry), depth and distances (the third dimension) give us a three dimensional world (Gibson 1950). Position in space is the term used by the candidate to describe the ability to perceive the different positions in space that one object may hold without reference to other objects. Space visualisation makes this possible.

Each type of space can be represented cognitively. Three dimensional spatial relationships can only be represented on a two dimensional surface through the use of perspective. Perspective is a learned concept (Concha 1984b).

4.3.5. Colour

The world is a colourful world and the ability to discriminate colour is present from the age of three years. Colour adds meaning to objects and contributes to depth perception. Colour perception is defined as the ability to discriminate between colours of different hue.

4.3.6. Praxias

Form, direction and spatial concepts are only important in an individual's life in that these concepts are needed to carry out motor tasks effectively. Motor activity consists of a series of learned motor combinations called praxias which, over the period of development, become more and more complicated. Perception, therefore, does not exist in isolation but is incorporated into perceptual-motor function.

The Behaviourists suggested that observing behaviour gives an indication of the way in which perceptions have been formed. It seems that consciousness first occurs at this point in the perceptual process i.e. awareness of the stimulus occurs when an output pattern has been generated (Kephart 1960).

Praxias, therefore, may be defined as movement combinations which are learned and stored as motor engrams in the brain and which subsequently become used automatically as part of the individual's repertoire of movement e.g. throwing a ball. The formation of movement combinations are discussed in more detail in Chapter Sixteen.

A defect in the perceptual process may influence the motor (output) component and the correct formation of motor engrams. It is, therefore, necessary to determine whether the defect is in the perceptual processing of sensory stimuli or in the motor planning and execution component as treatment will vary according to the location of the defect.

4.4. VISUAL PERCEPTUAL PROBLEMS OF ADULT BRAIN INJURED PATIENTS WHICH CONTRIBUTED TO THE MODEL

Various authors (Critchley 1953 Siev and Freishtat 1976, , Lezak 1976 and 1983, Luria 1980, Kolb and Whishaw 1985 Siev, Freishtat and Zoltan 1986) have enumerated and classified several visual perceptual skills in the adult and have described the problems that could occur when perceptual dysfunction was disrupted by brain injury. The correlation between types of lost function and localisation of lesion has led to the enumeration of different aspects of visual perceptual function. The correlation of different functions with parts of the brain has also contributed to knowledge regarding the different functions of the left and right hemispheres. (A summary of the function of the different sides of the cerebral hemispheres is given in Appendix 4). It is, however, beyond the scope of the occupational therapist and of this thesis to determine localisation of function. It is, however, important to be able to diagnose areas of V-P-M dysfunction. Table 4.1 summarises the areas in which adult dysfunction has been found to occur. The V-P-M areas already described are compared to these areas of dysfunction. Components of dysfunction that can be diagnosed with the V-P-M battery such as apraxia, agnosia, colour agnosia, problems with spatial discrimination, hemianopia, visual inattention and unilateral visual neglect are described in chapters sixteen to twenty. Several problem areas described by the above named authors have not been covered in the model of perception for reasons which are described in Section 4.4.1. - 4.4.7.

4.4.1. Problems Associated With Visual Perception Not Covered In The Model

4.4.1.1. Face Recognition

Face recognition appears to be an aspect of visual perception not often affected in C.V.A. The candidate has never seen a case of prosopagnosia in clinical practice.

Critchley (1953) queried the presence of this as a separate syndrome and stated that -

> *"It seems scarcely credible that human faces should occupy a perceptual category which is different from all other objects in space ... can there be any attribute of size, shape, colouring or mobility which distinguishes a face from other objects in such a way as to preclude identification. Presumably in those patients, a human face is seen, and is also realised for what it is, i.e. a human face; but the ownership of the face is not recognised. In other words, as the result of cerebral disease, the visual Gestalt formed by the combination of component features - whereby*

one individual differs from every other person - is no longer recognised." (pg. 196).

Although a face is the first object recognised by the infant, a test of facial recognition was not included in the battery or as a separate entity in the perceptual model for the reasons enumerated by Critchley. Face recognition is a complicated process encompassing the relationships between affect, physical structure, voice, memory etc. (Bruce 1988) and therefore uses many different perceptual skills.

4.4.2. Ambiguous Visual Stimuli

The ability to interpret ambiguous visual stimuli is for the most part learned or it requires individual interpretation of the stimulus seen. For this reason ambiguous visual stimuli are often used in personality tests e.g. the Rorschach Test. Because of the possibility of individual interpretation arising from this aspect of perception it was felt to be inappropriate in a model of visual-perceptual-mptor ability.

4.4.3. Visual Search, Scanning Accommodation And Convergence

Visual perception is based on the integrity of the visual system and a pre-requisite for visual processing is the detection and localisation of visual stimuli. To detect stimuli the visual world should be explored through eye and head movements. Accommodation and convergence also provide sensory input from the contraction of the muscles of the eye which contributes to the interpretation of distance and space. However, detailed analysis of the eye movements and the extent of convergence and accommodation is more adequately carried out by optometrists and is beyond the scope of this thesis.

4.4.4. Body Orientation

There are a number of terms used to describe aspects of body orientation such as body concept (the intellectual knowledge of the body including functions and names), body image (which is concerned with feelings about the body and is related to self image) and body schema (which is the physiological and subconscious mechanism).

The term generally accepted as the one that encompasses all aspects, is body awareness.

There is some inconsistency in the literature in the use of the above terminology. For example, Frederiks (1969(a . 20) defined normal body schema as *"... the perceptual, schematically conscious, structured, elastically bordered spatial perception of one's own body, constructed from previous and current (especially somaesthetic) sensory information".* This definition is comparable with that given above, however, Critchley (1953) regards body schema to be synonymous with body image. Lezak (1983) viewed body orientation as the organisation of personal space i.e. it includes all aspects of body awareness but may be more related to the actual position in space held by the body. Kolb and Whishaw (1985) as already discussed, described the body surface as one of the three subspaces and referred to the body surface on which stimuli could be localised to a place. As all types of body orientation, regardless of the terminology, deal with the body and not the visual perceptual system, this aspect of space is not included in the visual-perceptual-motor model described in this thesis.

4.4.5. Spatial Calculia

This aspect of perception involves the spatial ordering of numbers in an arithmetical problem. Dyscalculia of this nature may demonstrate *"errors of numbers relative to one another, confusion of columns or rows of numbers and neglect of one or more numbers ..."* (Lezak 1983:544). This skill is thought to be a function of spatial relations.

TABLE 4.1 CATEGORIES OF PERCEPTUAL -MOTOR FUNCTION

FUNCTION	PROBLEM	CORRESPONDING V-P-M SKILL
Visual Attention	Visual Inattention Neglect and Extinction	Visual Attention Concentration
Visual Recognition of:-		
Objects, figures and designs	Visual agnosias	Form perception
Faces	Prosopagnosia *	
Angulation	Poor perception of angulation	Directionality
Colour	Achromatopsia	Colour recognition
Visual Organisation	Inability to perceive:- Distorted visual stimuli * Ambiguous visual stimuli Fragmented visual stimuli * Incomplete visual stimuli	Form constancy Part-whole perception
Discrimination with Interference	Inability to perceive:- Embedded figures * Masked figures	Figure-Ground
Visual discrimination between complex configurations	Disturbed visual discrimination	Visual discrimination
Space:-		
Body Orientation	Asomatognosia *	Practical group i.e. space centred on the body
Directional orientation	Directional discrimination	Directionality
Distance estimates	Poor distance estimates	Position in space
Mental manipulation	Poor Mental transformations	Position in space
Space Representation:		
Topographical representations	Poor topographical orientation *	
Space Relations	Disturbed space relations	Space relations
Motor Function:-		
Eye Movements	Disturbed visual search, tracking, accommodation, convergence and fixation *	
Praxias	Apraxia	Praxias
Manual dexterity	Inco-ordination	Co-ordination

* not covered in the model.

4.4.6. Topographical Orientation

Benton (1969) defined topographical orientation as memory for familiar routes and for the location of objects and places in space. Difficulty with topographical orientation is due to an impaired ability to revisualise and to retrieve established visuo-spatial knowledge of a topographical nature. Every individual, therefore, has different topographical memories so it is difficult to standardise a test for this aspect of spatial perception. Topographical ability utilises the ability to revisualise and perform mental manipulations in space. The skills of revisualisation and spatial mental manipulation are important attributes of Visual Spatial Perception and are included in the Position in Space test of the V-P-M battery.

4.5. THE VISUAL-PERCEPTUAL-MOTOR MODEL

A model of visual-perceptual-motor skills was constructed on the basis of the foregoing. A circular design (see Figure 4.3) was selected for the graphic form of this model as it was felt that the perceptual skills mentioned are interdependent not hierarchical in the adult (Concha 1982). A summary of the perceptual skills selected for evaluation and the tests that were constructed for this purpose, are presented in Table 4.2. The definition of each V-P-M skill follows.

4.5.1. Form - Perception Of Form/Object Can Be Divided Into Three Areas:-

a) Form as Mass - the form is visualised as a whole, the different elements that go to make up the form are not perceived.

b) Overall Contour - the elements that make up the form are perceived and can be reproduced as in drawing the outline of a shape.

c) Units and sub-units of mass - elements of forms are again perceived and if separated can be combined to form a whole. This skill is also known as part-whole perceptual ability. Should an element be absent the skill of visual closure is needed to form a "gestalt". Part-whole perception is the ability to "make sense" out of incomplete or fragmented visual stimuli.

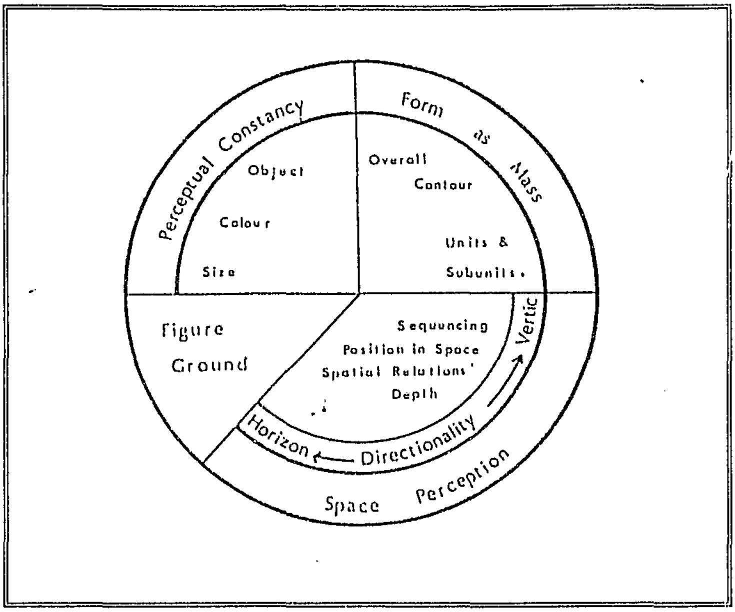


FIGURE 4.3 THE V-P-M MODEL

4.5.2. Figure-Ground - is the ability to select out of the mass of incoming stimuli only those that are meaningful and require attention. The selected stimulus forms the figure in the perceptual field while the majority of stimuli form the dimly perceived background.

4.5.3. Perceptual Constancy - is the ability to perceive an object as possessing invariant properties such as shape, position, size and colour in spite of the variability of the presentation.

4.5.4. Space - this is the ability to perceive space and consists of the following skills:-

a) Directionality - the ability to perceive the directional location of an object in relation to the direction of the line of gravity and to be able to transpose this perception into a one dimensional form i.e. on to paper.

b) Spatial Relations - the ability to perceive the position of two or more objects in relation to oneself and to each other.

Sequencing is a function of spatial relations. Spatial relations and the various clues that are learned to judge spatial relations such as movement, image, size, parallax etc. provide us with the perception of depth.

c) Position in Space - the ability to perceive the different positions in space that one object may hold without reference to other objects and to manipulate the positions that each object holds in the mind.

In the perception of space, the ability to transpose from the three dimensional to the two dimensional is important as it is the way in which the skill can be evaluated.

TABLE 4.2 PERCEPTUAL SKILLS IN THE VISUAL-PERCEPTUAL-MOTOR MODEL AND THE CORRESPONDING TESTS

PERCEPTUAL SKILL	TEST
Form:-	
Form as mass	Form Matching test*
Overall contour	Test of Visual-Motor Integration (Beery and Buktenica 1976)
Units and Sub-units	Part-Whole Test**
Figure-Ground	Figure-Ground Test*
Perceptual Constancy	Perceptual Constancy Test**
Space:-	
Directionality	Dot Line Test Arrows Test *
Spatial Relations	Spatial Relations Test** Block I } Stick I } Construction tests Design }
Position In Space (Space Visualisation)	Position in Space Test** Block I } Construction Stick II tests
Colour	Colour Matching Test*

* Motor Free Tests. ** Motor Reduced Tests.

4.5.5. Colour - the ability to discriminate between colours of slightly different hue.

4.5.6. Praxis - is the ability to plan and execute a movement sequence.

4.6. THE DESCRIPTION OF THE VISUAL-PERCEPTUAL-MOTOR TEST BATTERY

The following is a description of the fifteen tests, fourteen of which were constructed by the candidate to evaluate the individuals' skills described in the model The tests and their instructions are to be found in Appendix 1.

4.6.1. The Form Matching Test (Appendix 1 Section 2.1.)

The subject is presented with a multiple choice sheet on which twelve forms are represented. Twelve test cards with replicas of the forms are presented individually to the subject who is required to point to the duplicate on the multiple choice sheet. This test was designed to evaluate the ability to distinguish "form as mass" with minimal motor output (motor free). The term motor free is used for those tests in which the subject is only required to point to the correct answer. Visual inattention, the ability to scan a page and hemianopia can be noted from the individuals response pattern.

4.6.2. The Test Of Visual-Motor Integration (Beery and Buktenica 1967 and Beery 1982) (Appendix 1. Section 2.2.).

This test is the only previously published test used in the V-P-M battery described in this thesis. The subject is presented with a booklet containing a sequence of twenty-four drawn geometrical forms arranged from the simple to the complex which he is required to copy with a pencil as accurately as possible. This test evaluates visual-perceptual-motor-skills such as overall contour, units and sub-units of form, as well as constructional (drawing) praxis. Directionality, perspective, the ability to form a "Gestalt" and spatial relationship skills as well as the extent of the visual field can also be observed.

4.6.3. The Part-Whole Test (Appendix 1. Section 2.3.)

The subject is presented with a booklet containing twelve test items. Each page consists of one of the forms from the Form Matching test, an outline of the same form with a piece missing and an outline of three alternatives for the missing piece. The subject is required to indicate which of the three alternatives is the piece missing from the incomplete form thus evaluating part-whole ability. The form matching test is presented prior to the part-whole test to ensure that lack of form recognition does not influence the results on the part-whole test.

4.6.4. The Perceptual Constancy Test (Appendix 1. Section 2.4.)

This test is presented in a booklet of twenty pages with twenty separate stimulus cards. The test is divided into three sections as follows:-

a) Section 1 - consists of eight items each depicting objects or animals in different positions with different shadings. The subject is required to indicate the objects or animals which are the same as that depicted on the stimulus card. For adequate performance on this test the subject must have intact object recognition.

b) Section II - contains six items in which the subject is required to match the shape on the stimulus card with appropriate shapes on the multiple choice sheet taking into account constancy of size as well as shape, thus testing size constancy.

c) Section III - contains six items in which letter constancy is evaluated.

4.6.5. The Figure-Ground Test (Appendix 1. Section 2.5.)

This test has twenty test plates each of which consists of a multiple choice section of eight pictures at the bottom of the page and a "hatched" over section containing five of the pictures, at the top of the page. The subject is required to indicate on the multiple choice plate the figures depicted in the "hatched" section.

4.6.6. The Arrows Test (Appendix 1. Section 2.6.)

This test consists of a multiple choice sheet comprising twenty sets of arrows which are combined in different directions. The subject is presented with two stimulus cards from which he is required to select the one which has a duplicate on the multiple choice sheet. This test was constructed to evaluate motor free directional concepts.

4.6.7. The Dot-Line Test (Appendix 1. Section 2.7.)

This test consists of twelve test designs arranged from simple to complex. Each model consists of a series of dots linked by lines. The subject is required to link the dots in the test design at the bottom of the page so that it resembles the model. This test evaluates directionality through motor output i.e. drawing (praxis) and the estimate of distance.

4.6.8. The Spatial Relations Test (Appendix 1. Section 2.8.)

In this forty item test the subject is required to determine which figure out of five figures in each item matches the stimulus with respect to the constituent parts holding the same relationship to each other. This test is a motor reduced test of spatial relationships (i.e. the subject places a cross on the design of his choice). Visual discriminative ability can be noted.

4.6.9. The Position In Space Test (Appendix 1. Section 2.9.)

The test requires the subject to mentally rotate figures to the same position as the model in order to evaluate whether its orientation is the same as the model and to place a cross on the design of his choice. This test consists of forty items which evaluate position in space, space visualisation and mental rotations in space. This is probably the most difficult test in the series.

4.6.10. The Block Construction I Test (Appendix 1. Section 2.10.1.)

This test consists of six items in which the subject is required to copy a design of blocks from a model constructed behind a screen by the examiner. It is a motor enhanced test requiring skilled motor planning. The test is designed to evaluate spatial relationships in the three dimensional sphere as well as constructional praxis.

4.6.11. The Block Construction II Test (Appendix 1. Section 2.10.2.)

This is a six item test in which the subject is required to construct a three dimensional design from a two dimensional model depicted on a card. This test evaluates the ability to visualise space as well as spatial relationships.

4.6.12. The Stick Construction I Test (Appendix 1. Section 2.10.3.)

Stick construction is a test of six items in which the subject is required to copy a design constructed by the examiner. This is a two dimensional test of spatial relationships in comparison to the block construction as it is not necessary to take the different sides of the stick into account (Lezak 1983).

4.6.13. The Stick Construction II Test (Appendix 1. Section 2.10.4.)

This evaluates the ability to perceive spatial relationships through space visualisation as the subject is required to construct a design from a model depicted on a card. The test contains six items.

4.6.14. The Design Construction Test (Appendix 1. Section 2.10.5.)

This test consists of six items in which a design is constructed from a model using thin plastic pieces. This evaluates spatial relationships and constructional ability as the relationship of different pieces as well as the overlapping must be taken into account. Visual closure can be noted.

4.6.15. The Colour Matching Test (Appendix 1. Section 2.11.)

This test, like that of form matching, requires the subject to match a test card with its duplicate on the multiple choice sheet of sixteen colours. This test was designed to evaluate the discrimination of colour hues.

4.7. CONCLUSIONS

A model of visual perception has been constructed taking into account the perceptual skills that develop during childhood, the theories of perceptual function and the problems that can be experienced as a result of brain injury. A battery of fifteen tests was compiled fourteen of which were constructed by the candidate to evaluate the function of each of the areas defined. Some skills were tested through both motor enhanced and motor reduced or free tasks. Thus steps one to five of standardising the test have been completed. Steps six to ten will be described in the following chapters.

CHAPTER 5

THE FORM MATCHING TEST

5.1. INTRODUCTION

One of the earliest tasks given to young children is that of placing forms in a form board. Gesell (1966) stated that children were able to do a three holed form board by the age of three years. Testing the ability to perceive form is, therefore, basic to using other more complicated tests of visual perception.

The Form Matching test was drawn up by the candidate. The forms employed are similar to those used by Ayres in the Southern California Manual Form test (Ayres 1972 and 1980a). Unlike the Southern California Manual Form test the Form Matching test of the V-P-M battery is only a test of visual form discrimination and does not include tactile discrimination.

5.2. REVIEW OF SIMILAR TESTS

Previously published tests for adults are of three types i.e. recognition and matching of drawn designs, recognition and matching of real objects and pictures of real objects and recognition and matching of three dimensional shapes.

5.2.1. Recognition and Matching of Drawn Designs

The recognition and matching of drawn designs featured in tests such as the Visual Form Discrimination test (Benton et al 1983), the Discrimination of Forms (Level IV) of the Standford Binet Intelligence test (Terman and Merrill 1973) and in part of Raven's Coloured Matrices (Raven 1965). The Standford Binet and Raven's Coloured Matrices are unsuitable for use by the occupational therapist as they were designed and are used by psychologists for the purpose of measuring intelligence and general ability. However, Benton's Visual Form Discrimination test (Benton et al 1983) was designed to assess the capacity for complex visual form discrimination. The subject is required to select from four alternatives, the design that matched a two form stimulus. As the test evaluated the capacity to *"... discriminate between complex visual stimulus configurations differing in one or other minor characteristic"* (Benton et al 1983:55), it is a more complex task than the Form Matching test of the Visual-Perceptual-Motor Battery designed by the candidate to evaluate form as mass. It was therefore considered not compatible with the definition of form used in this thesis. The Benton test is, however, a suitable test for occupational therapists to use in the evaluation of visual discriminative function in the adult.

The Visual Identification of Shapes test of the LOTCA (Itzkovich et al 1990) is comparable to the Form Matching test in the V-P-M battery and the LOTCA, like the V-P-M battery, is not sensitive to age related changes.

5.2.2. <u>Recognition and Matching of Real Objects and Pictures of Real Objects</u>

Recognition and matching of real objects has been described by Siev and Freishtat (1976), Lezak (1983), Siev et al (1986), and Whiting et al (1985). The tests described by Siev and Freishtat (1976) and Siev et al (1986) were not standardised whereas the Object Matching test of the Rivermead Perceptual Assessment Battery (RPAB) (Whiting et al 1985) has some normative data and is, therefore, a more acceptable test of object matching, as is the most recent version of the Beta Examination. This latter test was originally developed by the U.S. Army to evaluate the intellectual ability of World War I recruits who were illiterate. The Beta Examination contains a sub-test which requires the identification of pictures of tools, birds and geometric figures as well as numbers (Lezak 1983). Both the RPAB test and the Beta Examination sub-test are suitable for the evaluation of the ability to identify real objects, although one uses real objects and the other pictures. Likewise the LOTCA has a section that requires the subject to identify pictures of real objects.

The identification of real objects appeared to be a simple task, object agnosia being rarely seen in patients. Whiting et al (1985), obtained a mean score of 3,25 (maximum score 4) when testing twenty right hemisphere stroke patients on the Object Matching test, of the RPAB. However, the difference between the performance of the normal individuals and that of the patients was significant at the P<0,05 level of significance, but, failure in Whiting's patient group appeared to be due to field neglect rather than difficulty in identifying the objects. The component of the LOTCA (Itzkovich et al 1990) which requires the patient to name or match illustrations of everyday objects, unlike the RPAB, did not differentiate between normal individuals and patients which makes its usefulness doubtful.

Tests for the recognition and matching of real objects and pictures of real objects are not directly comparable to the Form Matching test of the V-P-M Battery as the Form Matching test evaluates basic shape recognition.

5.2.3. <u>Recognition and Matching of Three Dimensional Shapes</u>

Recognition and matching of three dimensional shapes is usually found in form board tests. A form board test was used to evaluate Form Perception in the Cognitive Perceptual Motor Evaluation (CPM) for head injuries in the Occupational Therapy Services at Santa Clara Valley Medical Centre (Meeder 1982). The Triang Toy Post Box was similarly used by Adams (1974). Both these tests are unstandardised and patient performance is not evaluated against norms. Form board tests are nevertheless useful tools for the evaluation of 'form as mass', but they require manual manipulation for the correct placement of the pieces. The Form Matching test of the V-P-M battery, although a two dimensional test, was designed to eliminate manual manipulation and the effects of disturbed motor ability.

5.3. REASONS FOR COMPILING A NEW TEST

The Form Matching test was constructed for use in the V-P-M battery as no motor free test existed for the evaluation of 'form as mass'. The RPAB Object Matching test and the LOTCA Visual Identification of Shapes test were not available at the time that the V-P-M battery was constructed and the Benton Visual Form Discrimination test was thought to be more suitable for visual discrimination. The Form Matching test consists of a response sheet containing twelve shapes. Test cards with identical shapes are placed in front of the patient who is required to point to its duplicate on the multiple choice sheet. It is classified as a motor free test by the candidate. (See Appendix 1 Section 2.1.).

5.4. RESULTS OF TESTING NORMAL INDIVIDUALS

5.4.1. Accuracy Scores

The means and the standard deviations were calculated for the accuracy scores for the normal group as a whole (Table 5.1) and for each age group (Table 5.2).

TABLE 5.1 MEAN AND SD OF ACCURACY SCORES - FORM

n	MAX SCORE	MEAN	SD	RANGE
118	12	11,96	0,20	11-12

TABLE 5.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - FORM

AGE	n	MEAN	SD	RANGE
20-29	20	12	0,0	12 only
30-39	21	11,95	0,22	11-12
40-49	19	11,95	0,23	11-12
50-59	19	12	0	12 only
60-69	19	11,84	0,38	11-12
70-79	20	12	0	12 only

No further analysis was carried out on the accuracy scores as very few errors were made. Only five subjects out of 118 score : eleven out of a possible twelve points.

5.4.2. Time Scores

The means and the standard deviations were calculated for the time scores for the group as a whole (Table 5.3) and for each age group (Table 5.4).

TABLE 5.3 MEAN AND SD OF TIME SCORES - FORM

n	MEAN/SEC	SD	RANGE
118	22,73	8,11	11-56

TABLE 5.4 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - FORM

AGE	n	MEAN/SEC	SD	RANGE
20-29	20	17,15	5,03	11-28
30-39	21	17,24	3,18	11-23
40-49	19	20,63	5,47	12-29
50-59	19	22,53	5,39	14-35
60-69	19	28,1	9,27	16-56
70-79	20	31,2	7,55	19-46

Analysis of variance was used to analyse the time scores. Levine's test for equal variance was significant (P=0,0170) therefore, the variances between the groups were not equal. The Welch Test for unequal variances demonstrated a significant difference between the means of the groups (P=0,0000), therefore, the separate variance t was used with the Bonferroni correction for multiple comparisons to establish where the differences lay (see Table 5.5).

Thus it can be seen that the age groups varied from one another in the time taken to complete the test. The twenty to twenty-nine year olds differed from the age groups fifty through to seventy-nine years as did the thirty to thirty-nine year olds. The forty to forty-nine year olds differed only from the seventy year old group.

On this basis and on the basis of the results found in other tests that make up the battery the time scores were combined within the age groups twenty to thirty-nine, forty to fifty-nine, sixty to sixty-nine and seventy to seventy-nine, there being no significant difference in the time scores within these age groups. The combined scores thus determined were similarly analysed (see Table 5.6).

TABLE 5.5 DIFFERENCES IN TIME SCORES - FORM

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,9472
20-29	40-49	0,460
20-29	50-59	0,0027 *
20-29	60-69	0,0001 **
20-29	70-79	0,0000 ***
30-39	40-49	0,0250
30-39	50-59	0,0008 *
30-39	60-69	0,0001 **
30-39	70-79	0,0000 ***
40-49	50-59	0,2892
40-49	60-69	0,0051
40-49	70-79	0,0000 ***
50-59	60-69	0,0309
50-59	70-79	0,0002 **
60-69	70-79	0,2621

* = P<0,05 ** = P<0,01 *** = P<0,001

By combining the time scores it can be seen that each age group differed from the others except in the case of the forty to forty-nine year olds and that of the sixty to sixty-nine year olds.

TABLE 5.6 DIFFERENCES IN RE-GROUPED TIME SCORES - FORM

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-59	0,0002 ***
20-39	60-69	0,0001 ***
20-39	70-79	0,0000 ***
40-59	60-69	0,0091
40-59	70-79	0,0000 ***
60-69	70-79	0,2621

*** = P<0,0001

5.4.3. Discussion Of The Results Of Testing Normal Individuals

The Form Matching test of the V-P-M Battery is a very simple test, showing no age related changes in test performance when accuracy is analysed. The null hypothesis is, therefore, retained in this instance i.e. if adults between the ages of twenty and seventy-nine years are evaluated with the V-P-M battery, a decrease in the ability to match forms will not be found with advancing age.

Benton et al (1983), Whiting et al (1985) and Itzkovich et al (1990) also did not describe age related changes in the corresponding tests. The Visual Form Discrimination Test (Benton et al 1983) was administered to a group of eighty-five normal subjects ranging in age from nineteen years to seventy-four years, although this test is far more complex than the Form Matching test of the V-P-M Battery, it still cannot be used to measure age related changes in normal individuals.

Howell (1972) found that in the recognition of relatively meaningless and unfamiliar patterns and colours, subjects in the age group fifty-six to eighty-three years did not perform as accurately as the younger subjects in the age group eighteen to forty years. The accuracy of the young was also significantly superior to the old with regard to complex photographs of common, present day objects. However, old and young subjects exhibited the same level of perceptual accuracy with respect to photographs of single antique objects. It is concluded that the accuracy of perceptual recognition of the elderly is related to the meaningfulness and the familiarity of the stimulus material. It seems, therefore, that lack of age related changes in the Form Matching test of the V-P-M battery could be due to the level of familiarity with the forms used in the test.

Although there was no difference in the accuracy scores obtained for each age group for the Form Matching test, the differences found in the time taken to complete the test, along with the results obtained on the other tests in the battery confirmed that the time taken to complete the test increases with age. The increase in the amount of time taken to complete the Form Matching test occurred from forty years of age, thus the null hypothesis was rejected at the 0,01% level of significance.

5.5. RESULTS OF TESTING THE PATIENT GROUP

5.5.1. Accuracy Scores

The mean and the standard deviation was calculated for the accuracy scores for the group as a whole (Table 5.7) as there were no differences in the performance of the test demonstrated by the normal group. It should be noted that only thirty-one of the thirty-three patients were evaluated by means of this test. Testing was discontinued on one seventy-five year old patient due to his inability to concentrate after having performed poorly on the test of Visual-Motor Integration. Another patient in this age category was untestable on all the tests in the battery.

TABLE 5.7 MEAN AND SD FOR ACCURACY SCORES - FOR PATIENTS - FORM

AGE	n	MEAN	SD	RANGE
20-79	31	11,48	1,39	6-12

5.5.2. Time Scores

The means and the standard deviations were calculated for the age group twenty to thirty-nine, forty to sixty-nine and seventy to seventy-nine years as these were the age groups within which no statistical difference occurred in the time scores obtained by the normal group (see Table 5.8).

TABLE 5.8 MEAN AND SD FOR TIME SCORES - FOR PATIENTS - FORM

AGE	n	MEAN/SECS	SD	RANGE
20-39	7	33,29	20,56	16-72
40-69	20	43,60	21,54	20-88
70-79	4	135,5	114,67	37-297

5.5.3. Comparison Between Patients And Normal Subjects

5.5.3.1. Accuracy Scores

No statistical difference was found between the scores obtained by the patients and those obtained by the normal group. Only two patients obtained a score lower than eleven i.e. scores of six (patient number twenty-seven) and seven (patient number nine). Both these patients were found to be untestable on the other tests in the V-P-M Battery. The results of testing the patient group on the relevant component of the LOTCA demonstrated similar findings. This is contrary to the findings of Whiting et al (1985) on the Matching of Real Objects test of the RPAB. However, failure in Whiting's group, as already mentioned, appeared to be due to field neglect rather than identifying objects.

5.5.3.2. Time Scores

The mean of the time scores of the patients within the three age groups were compared with that of the normal group. Levin's test for equal variances was found to be significant therefore the Welch Test was used for testing significant differences between the means of the two groups (see Table 5.9).

TABLE 5.9 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - FORM

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20-39	0,0000	0,0836
40-69	0,0000	0,0006 ***
70-79	-	-

*** P<0,001

It can be seen that in the age group forty to sixty-nine there was a highly significant difference between the time scores obtained by the patients versus the normal group. The scores obtained for the patients aged seventy to seventy-nine were not compared with those of the normal group due to the small number of patients in that age group.

5.5.4. Analysis Of Tests Executed By Patients

In the age group twenty to thirty-nine all patients obtained an accuracy score of twelve. Time scores fell into percentiles of between fifty-one and 100 thus indicating that many of the patients performed at a much slower speed than the normal group although they experienced no difficulty with matching the forms. Time percentiles must be interpreted in reverse, higher scores indicate that the subject takes longer to complete the test and thus high scores are poor scores.

In the age group forty to fifty-nine - three of the patients obtained accuracy scores of eleven. The time percentiles ranged from fifty-seven to 100.

In the age group sixty to sixty-nine - two of the patients obtained a score of eleven and one a score of seven (patient nine). The patient who received this score of seven was untestable on other perceptual measures. The time percentiles ranged from the fifty second percentile to 100.

In the age group seventy to seventy-nine - one patient was untestable on all tests of the V-P-M Battery. The zero scores of this patient were not included in the analysis. One of the patients in this age group obtained a score of six and the other three obtained a full score, the patient who received the score of six was untestable on other measures. Time scores ranged from the eighty-fifth percentile to 100, the patients were therefore very slow.

5.5.5. Problems That May Be Diagnosed From Performance On The Form Matching Test

The problems which affected the patient group and caused failure on the Form Matching test were as follows:-

5.5.5.1. Left visual field neglect - this contributed to scores of below twelve in seven of the patients.

5.5.5.2. Visual Inattention - the failure of patient number nine could be attributed, in part, to visual inattention as she used only one clue to help her select the correct shape. For example, if the test card had a shape with a point e.g. the triangle, she chose the first shape that she saw that had points i.e. the star. However, if told to look at the whole multiple choice sheet she could match the shape correctly although failure still occurred in the left visual field.

5.5.5.3. Impulsivity - there was a tendency among those who made only one mistake, to be impulsive in their selection.

It appeared that most of the patients who obtained a score below the maximum on the Form Matching test in this battery did so, not because of difficulty with matching of forms as such, but due to hemianopia and loss of the left visual field. Eighty-one percent of the mistakes made were located in the two left columns of figures. However, the size of the response sheet was such that only patients with a severe left field neglect omitted items on the left side of the response sheet.

There is evidence in the literature (Meier and French 1965, Newcombe and Russel 1969, Bisiach et al 1975 and Benton et al 1983,) to support the view that visual discrimination of complex visual stimuli differentiates between patients with lesions of the left and right side of the brain. Those patients with right hemisphere dysfunction tend to perform more poorly than those with left hemisphere dysfunction. Right brain damaged patients with visual field defects demonstrated the worst performance (Bisiach et al 1975). Bisiach et al's (1975) study showed that patients with left hemisphere dysfunction demonstrated poorer performance when the task of matching random shapes involved memory.

Wapner et al (1978) and Kertesz (1979) showed that objects not recognised visually were recognised with 100% accuracy when tactile recognition was employed. The Form Recognition test in the V-P-M Battery is presented in such a way that it is impossible to evaluate the presence of a true visual agnosia as the forms cannot be manipulated and, therefore, recognised and matched through touch and stereognosis.

Patients with posterior cerebral lesions identified simple visual items considerably better than compound items consisting of several components (Levine and Calvanio 1978). This might explain why the patient group performed better on the Form Matching test than they did on the Space Relations test of the V-P-M Battery. The Space Relations test requires complex form matching and visual discriminative ability. Visual inattention may also contribute to a reduced performance in the Spatial Relations test.

In the normal individual, simultaneous perception of multiple elements occurs sequentially item by item. *"In the patients, the perceptual acts occur normally but the interval between such acts is inordinately prolonged"* (Levine and Calvanio 1978:66). This could explain why the time taken to complete the Form Matching test by the patients who achieved a maximum accuracy score was much longer than that taken by the normal subjects.

As the two patients who scored less than eleven points on the test were found to be untestable on the other perceptual measures, it seems possible that the Form Matching test could be used as an initial screening procedure to identify those patients who are at a recovery stage which enable detailed evaluation to take place.

5.6. ITEM ANALYSIS

As there was very little variation in the accuracy scores for the various age groups (only five subjects out of 118 scored less than the maximum score) no item analysis was done. The five mistakes were as follows:- item three one mistake, item four two mistakes, item six one mistake, item nine one mistake. The mistakes made were due to carelessness and were corrected by the testees in the course of test execution although the initial response was incorrect. Thus, it appears that all items were equally easy.

5.7. TEST RELIABILITY

5.7.1. Split-Half Reliability

As there was very little variation in the test scores it was not possible to carry out Split-half studies or to calculate the Kuder-Richardson formula 20.

5.7.2. Inter-rater Reliability

No statistical difference was found in the accuracy scores between the different examiners. There was however a difference p=0,042 (P<0,05) between raters in recording the time taken to complete the test. No explanation could be found for this as no difference existed in the time scores on any of the other similarly administered tests. The statistician therefore recommended that as the difference was only at the 5% level of significance, that the time scores obtained on the Form Matching test by the three examiners be combined for all the analyses described in section 5.4.2.

5.8. TEST VALIDITY

5.8.1. Content Validity

From the ratings given by the fifteen occupational therapists a mean score of 4,20 in favour of the test measuring the perception of "form as mass" was calculated. This value in the upper quartile indicated a very much better than average test of measuring form as mass.

5.8.2. Construct Validity

Three types of construct validity were examined.

5.8.2.1. Age differentiation

It appears that the Form Matching test does not differentiate ability at the different ages for accuracy and is not sensitive to age related changes. This test is, therefore, not a valid test of decreasing function with age.

5.8.2.2. Correlation with other tests in the Battery

The statistician advised that this test should not be included in the correlation studies due to its poor discriminative ability.

5.8.2.3. Results of testing patients

The accuracy scores of thirty-three patients between the ages of twenty and seventy-nine although not statistically different from those of the normal individuals, were useful for determining patients who were untestable on the other tests in the V-P-M battery.

5.9. CONCLUSIONS

From the studies carried out on the Form Matching test it appears that this test is a poor discriminator of age related changes and perceptual changes resulting from brain damage. Its value appears to be in the diagnosis of a left field defect and as a screening test.

In summary this test-

- demonstrated no age related changes in the accuracy scores of the normal subject;
- showed no statistical difference between the accuracy scores obtained by the normal group and the patient group;
- demonstrated similar results to those found by other researchers on form matching tests;
- showed very little variation in the scores obtained by the normal group, therefore, no item analysis was carried out. No reliability studies were carried out for a similar reason;
- was a valid measure of "form as mass".

For the above reasons this test should not be included as a diagnostic instrument in the final test format. It, however, could be used as a quick screening test to determine when a patient is ready for more comprehensive assessment or if there is a need to evaluate basic form perception.

The results of the psychometric evaluations of all the tests of the V-P-M battery will be compared in Chapter Twenty-One. Their comparative usefulness and the need to retain them in the battery will then be discussed.

CHAPTER 6

THE TEST OF VISUAL - MOTOR INTEGRATION

6.1. INTRODUCTION

A classical means of testing the ability of a patient to perceive shape (overall contour) is through the copying of forms and designs. Precise copying is dependent on the ability to accurately perceive the designs to be copied as well as an intact motor system. A copy forms test, therefore, evaluates a combination of visual input and motor output and can be called a test of visual-motor integration. This method of evaluating form perception is more complex than the method described in the previous chapter as it requires the subject to analyse the components making up the form.

The Developmental Test of Visual-Motor Integration (VMI) is a copy forms test made up of a sequence of twenty-four geometric forms arranged from the simple to the complex. It was first published in 1967 (Beery and Buktenica 1967). Revised normative data were published in 1982, 1989 and 1997 (Beery 1982, 1989 and 1997). Beery has stated, that the test could be used with adults as well as children, but was primarily designed for assessing sensory-motor development in pre-school and early grade level children (Beery 1982). The 1982 version of the test provided standard scores, percentiles and age equivalent scores for ages four through to thirteen years. The 1989 and 1997 versions extended the upper age limit to seventeen years. The figures given for the seventeen year olds are said to serve for older age groups as well. Although occupational therapists in South Africa have found the Developmental Test of Visual-Motor Integration to be an extremely valuable instrument in the evaluation of visual-motor function in both children and adults, no scores for the adult population or information on the clinical use of the test with adults has been described in the literature. For a description of the test see Chapter Four (4.6.2.) and Appendix 1 (Section 2.2).

6.2. REVIEW OF SIMILAR TESTS

Several other paper and pencil tests are available for the evaluation of the adult, some of which give information on normal adult function . Many of these tests have a memory component, for example, the Memory for Designs Test by Graham and Kendall (1960) (see Figure 6.1) and the Visual Retention test by Benton (1974). The designs in these two tests could well be used to analyse visual-motor problems. The Visual Retention test does purport to assess visual perception and visual constructive abilities in Form C of the test (Benton 1974). However, the memory component in the Forms A and B, as in the Graham and Kendall test, distorts the analysis of visual-motor problems.

Other form copying tests that do not include memory are available, the most well known and well researched of these being the Bender-Gestalt test (Bender 1938) (see Figure 6.2). Bender herself called this a Visual-Motor Gestalt test, but over time it has come to be known as the Bender-Gestalt test (Lezak 1983). There are different versions for the administration of this test and the "*... variables that can be scored are numerous and equivocal and their dimensions are often difficult to define*" (Lezak 1983:388). This has resulted in a profusion of scoring systems, which together with the fact that the test is a well known psychological instrument used by psychologists for diagnostic purposes, makes it unsuitable for use by the occupational therapist.

The Minnesota Perceptuo-Diagnostic Test (Fuller 1969) (see Figure 6.3) makes use of two of Wertheimer's Gestalt designs which are also used in the Bender-Gestalt Test. These specific designs were used because of their ability to produce rotations in pathological conditions as well as yielding some information on visual-motor performance. The standardisation data available are similar to that of the VMI but using only two designs is restrictive in terms of the information that they can provide about the perception of overall contour of form, motor planning and spatial concepts.

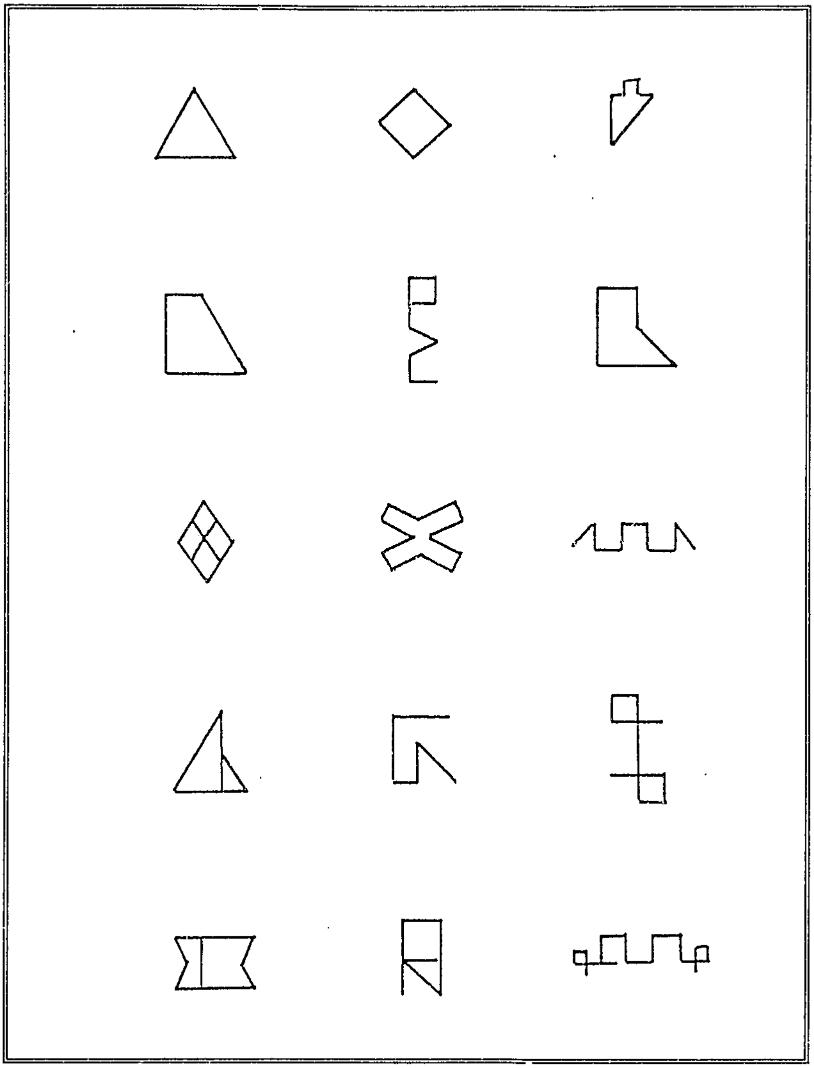


FIGURE 6.1 DESIGNS FROM THE GRAHAM AND KENDALL MEMORY FOR DESIGNS TEST (reduced in size)

A Complex Figure Test devised by Rey in 1941 and standardised by Osterrieth in 1944 (Osterrieth 1944, Lezak 1983 and 1995) (see Figure 6.4) is one of the more suitable tests for occupational therapists to use in the evaluation of visual-motor performance in the adult as it contains normative data on sixty adults in the age range sixteen to sixty years. Taylor (1979) drew up an alternative figure for use in re-testing. There are, however, no normative data for Taylor's figure. The combination of figures and direction of lines in the Rey-Osterrieth figure make the evaluation of directional confusions, fragmentation and spatial problems possible. However, the test consists of only one design, so one mistake can be compounded and the desirability of a test beginning with simple items and ending with more complicated and difficult ones is lost. The published data on the Rey-Osterrieth Complex Figure Test does show a differentiation between performance in different age groups, but a study carried out by Kaplan and Hier (1982) found that normal subjects younger than sixty-eight years performed better on the test than those older than sixty-eight years ($P<.01$). Normative data, however, are not provided by Kaplan and Hier (1982).

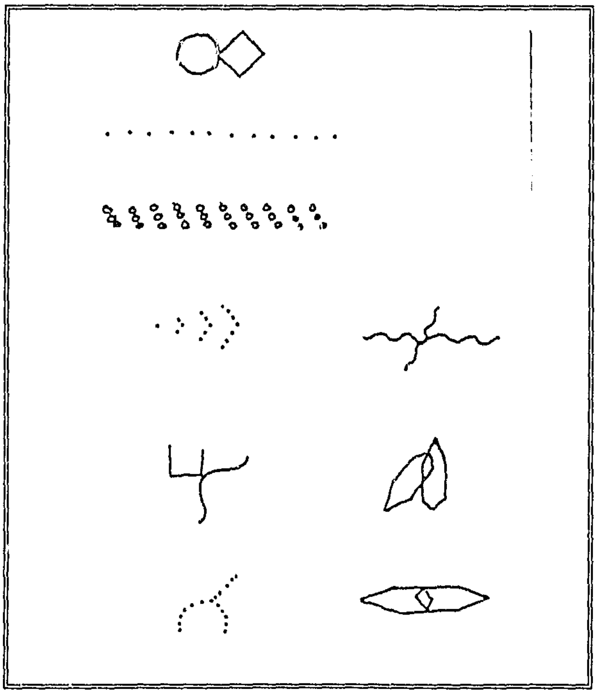


FIGURE 6.2 THE BENDER -GESTALT FIGURES (reduced in size)

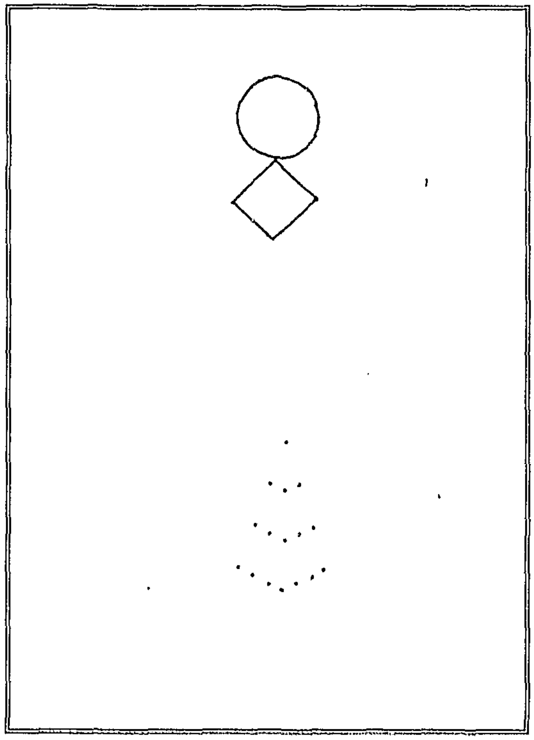


FIGURE 6.3 THE MINNESOTA PERCEPTUO-DIAGNOSTIC FIGURES

The Slosson Drawing Co-ordination Test (Slosson 1980) (Figure 6.5) was designed to identify individuals with perceptual defects as well as disorders of eye-hand co-ordination. The test used several figures from the Standford-Binet Intelligence Test and scoring is based on three repetitions of each of these designs. Data on normal individuals for ages one to twelve years are provided as well as a reliability study which included data from individuals aged four to fifty-two years. However, it was felt to be inappropriate for occupational therapists to use figures from the Standford-Binet Test as this is a test specifically related to the assessment of intelligence.

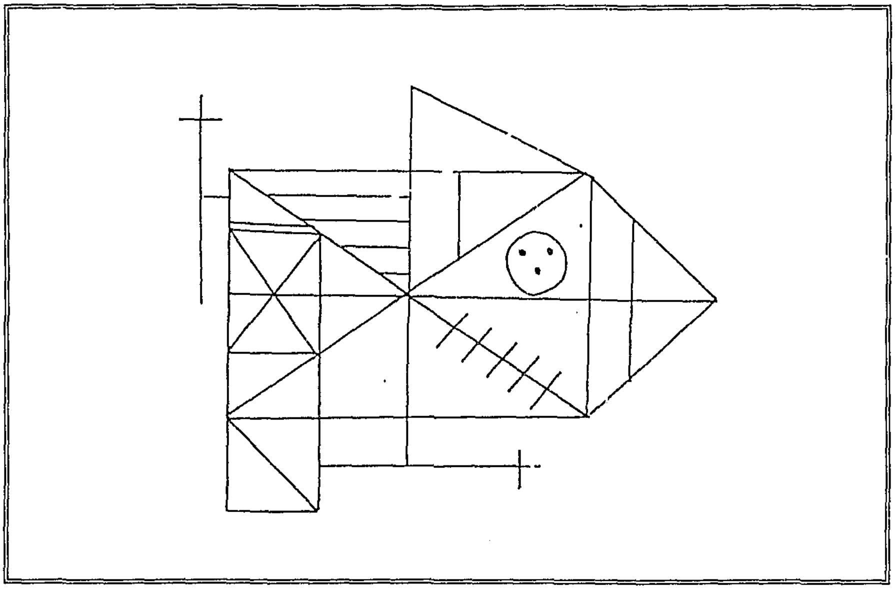


FIGURE 6.4 THE REY-OSTERRIETH COMPLEX FIGURE TEST (reduced)

The most recently published tests which were specifically designed for use by occupational therapists are the R/L Copy Shapes test of the Rivermead Perceptual Assessment Battery (Whiting et al 1985) (see Figure 6.6) and the Copy Geometric Forms test of the LOTCA (Itzkovich et al 1990). The R/L Copy Shapes test (Whiting et al 1985) was designed to evaluate inattention through the ability to copy correctly to the left and to the right of the mid-line. In this test no significant relationship was found to age but a correlation was established between the R/L Copy Shapes and the Mill Hill Vocabulary Scale, as well as with the Synonyms Test, a fact that the authors attributed to the possible influence of IQ on perceptual performance.

Accordingly, the final determination of a perceptual deficit is based on estimating the patient's pre-morbid intelligence level. This requirement is not always easily fulfilled, however, this test could be used for the Visual-Perceptual-Motor assessment described in this thesis but it had not been formulated or published at the time this thesis was begun. The candidate has since had extensive experience in the clinical use of the R/L Copy Shape test, but has found the VMI test to be of greate diagnostic value due to the greater variety of shapes in the VMI test.

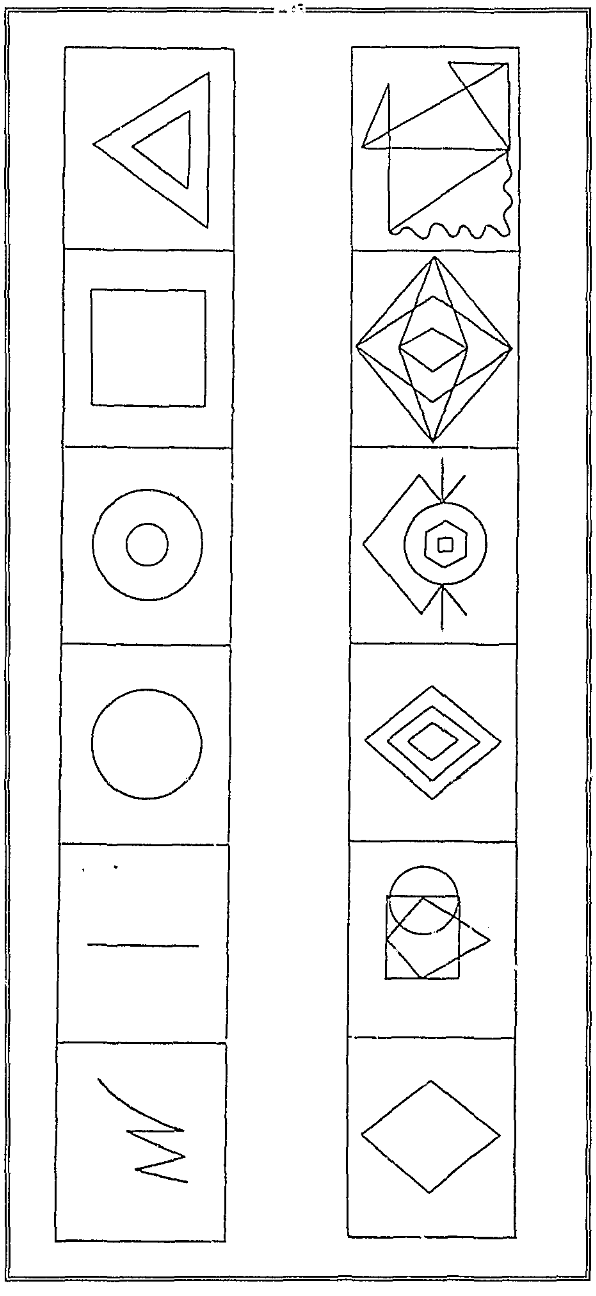


FIGURE 6.5 FIGURES USED IN THE SLOSSON DRAWING CO-ORDINATION TEST (reduced)

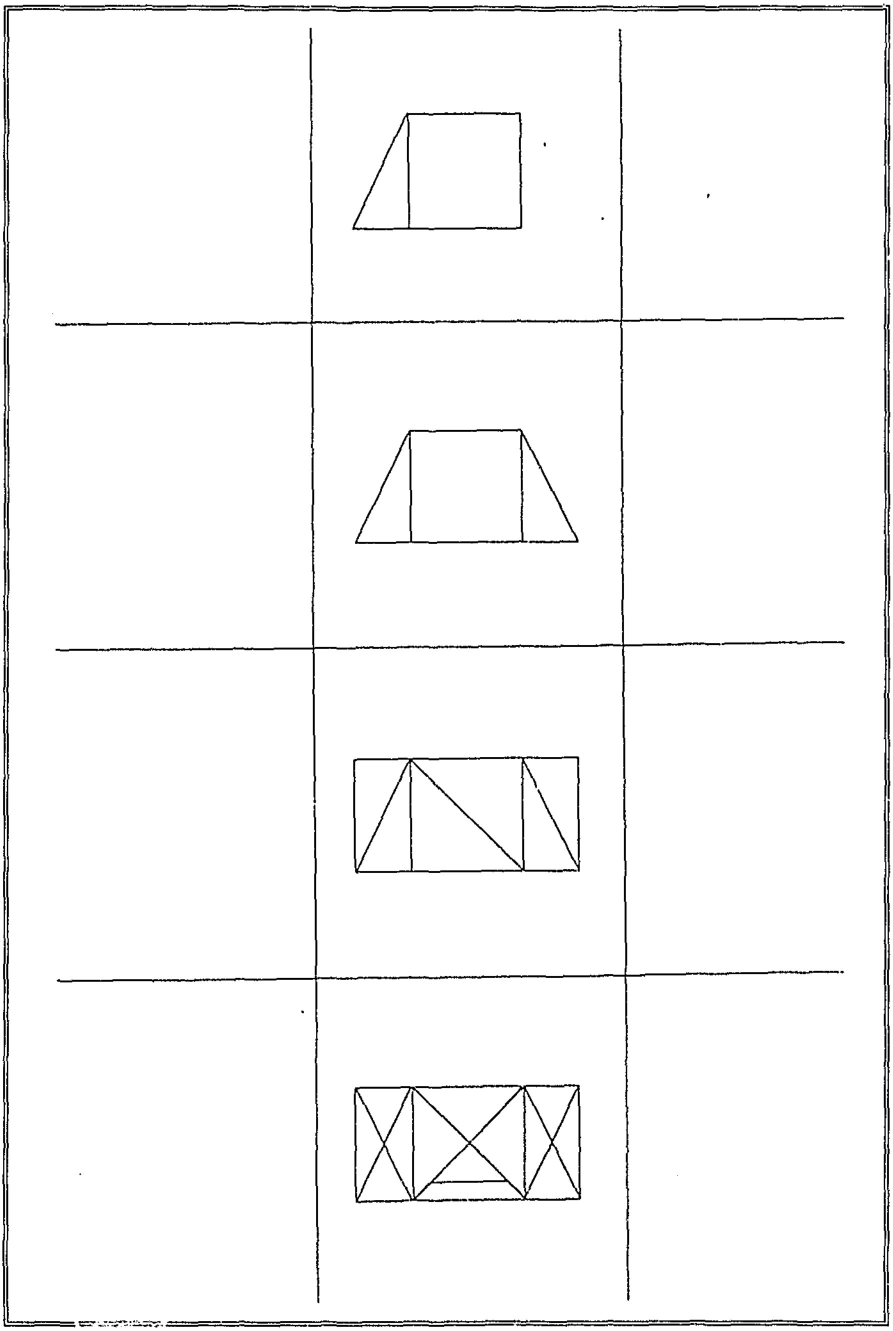


FIGURE 6.6 FIGURES FROM THE RIVERMEAD PERCEPTUAL ASSESSMENT BATTERY (reduced)

The Copy Geometric Forms in the LOTCA battery (Itzkovich 1990) requires the subject to copy five simple forms. No difference was found in the performance of the different age groups on any of the tests in the LOTCA battery.

Numerous tests based on free drawings are described by Critchley (1953), Luria (1980) and Lezak (1983) and include such items as drawing a daisy, a bicycle, the sun and a clock. Lezak (1983) has developed a twenty point scoring system for the bicycle drawing test which could be used as a pre and post treatment evaluation. Kolb and Whishaw (1985 and 1990) provided means and standard deviations for the bicycle drawing test for ages two to nineteen years, twenty to twenty-nine years and thirty to fifty years.

Drawing a daisy, a bicycle, or the sun is a quick way to determine the presence of a perceptual defect. However, the absence of a model changes the test from that of perceiving something to be copied into one of recalling a picture in the mind. This type of assessment cannot, therefore, be compared to that of the VMI test.

6.3. REASONS FOR SELECTING THE VMI TEST

The analysis of the various tests available showed that most of them were unsuitable for use in the battery described in this thesis and many appeared on the market well after this thesis was begun. The VMI test, however, seemed to be a valuable assessment tool for the following reasons:-

- It taps visual-motor integrity in the sphere of visual analysis of overall contour of shape.
- It provides different designs for copying presented in a graded way from the simple to the complex thus meeting one of the criteria for good test design.
- Some of its designs are fairly complex thus making them suitable and challenging for adults.
- The variety of designs in the test makes it possible to evaluate different aspects of perceptual function such as directionality, spatial problems, unilateral neglect as well as overall contour.

- Comparisons can be made between the results of this test and motor free tests which can assist in the diagnosis of a drawing apraxia.

- The designs are not widely used in other tests.

- The full potential of the VMI test as an assessment tool has not been fully utilised for all age groups.

The provision of normative data for the different age groups between twenty and seventy-nine years could make the VMI test even more useful for the evaluation of function in the adult patient. (See Appendix 1, section 2.2).

6.4. RESULTS OF TESTING NORMAL INDIVIDUALS (POPULATION THREE)

6.4.1. Accuracy Scores

The means and the standard deviations were calculated for the normal group as a whole (Table 6.1) and for each age group (Table 6.2). The scoring described by Beery and Buktenica (1967) was used with additional criteria added by the candidate to make the scoring easier (see Appendix 1, section 2.2).

TABLE 6.1 MEAN AND SD OF ACCURACY SCORES - VMI

n	MAX SCORE	MEAN	SD	RANGE
116	24	19,6	2,96	11-24

The number of test results used in the calculation of the means and standard deviations and for testing the significant difference between the age groups was 116 as compared to 118 for the other tests. This was due to the fact that one test completed by a seventy year old subject had a score that was much lower than the other scores for the same age group and much lower in comparison to his other test results. It was felt that his performance of the VMI test was not typical of his usual performance and the statistician suggested that the result should not be included. Also the test was not administered to one subject in the twenty to twenty-nine year old group as she was familiar with the VMI, having accompanied her mother to occupational therapy where she had watched the VMI being administered.

TABLE 6.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - VMI

AGE	n	MEAN	SD	RANGE
20-29	19	20,68	2,81	11-24
30-39	21	20,38	2,48	14-24
40-49	19	20,90	2,42	17-24
50-59	19	20,68	1,97	16-24
60-69	19	18,37	3,11	11-22
70-79	19	16,95	2,80	12-22

The Pearson Chi-square Test showed that there was a statistically significant difference between the performance of the various age groups (P <0,001) (Table 6.3) a cut point score of nineteen being used for this analysis.

TABLE 6.3 COMPARISONS OF ACCURACY SCORES - VMI

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 19	21,1	23,8	26,3	21,1	57,9	84,2	38,8
Greater than 19	78,9	76,2	73,7	78,9	42,1	15,8	61,2

Pearson Chi-square Value = 27,694 P < 0,001

On the basis of this finding, further analysis of the differences between the age groups was carried out.

Yates' corrected Chi-square demonstrated no difference between the accuracy scores of ages sixty to sixty-nine and seventy to seventy-nine (P = 0,1525). These two age categories were therefore grouped together for further analysis.

The Pearson Chi-square Test showed that there was no difference between the scores for each of the age categories twenty to twenty-nine, thirty to thirty-nine, forty to forty-nine, and fifty to fifty-nine years. Thus the scores of these age categories were also grouped together for further analysis. A highly significant difference between the accuracy scores of the two groupings now formed i.e. the twenty to fifty-nine year olds and the sixty to seventy-nine year olds at the P < 0,001 level of significance, was found using Fisher's Exact test (see Table 6.4).

It would appear that there was a "drop off" in the ability to copy forms on to paper at about sixty years of age. However, as other tests in the battery of which the VMI test forms a part, showed a slightly different picture, it was decided to group the ages according to the groupings used in the other tests in the battery for the calculation of percentile ranks.

TABLE 6.4 COMPARISON OF ACCURACY SCORES 20 - 59 AND 60 - 79 YEARS - VMI

AGE	20-59	60-79	TOTAL
Less than 19	23,1	71,1	38,8
Greater than 19	76,9	28,9	61,2

Fisher's Exact test P < 0,001

Percentile ranks were computed from the raw scores for the age categories twenty to thirty-nine, forty to fifty-nine, sixty to sixty-nine and seventy to seventy-nine (see Appendix 2).

6.4.2. Time Scores

The means and the standard deviations were calculated for the time scores for the group as a whole (see Table 6.5) and for each group (see Table 6.6).

TABLE 6.5 MEAN AND SD OF TIME SCORES - VMI

n	MEAN/SECS	SD	RANGE
116	578,90	328,94	106-1918

TABLE 6.6 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - VMI

AGE	n	MEAN/SEC	SD	RANGE
20-29	19	418,95	158,65	106- 780
30-39	21	510,80	348,47	198-1616
40-49	19	683,90	379,64	214-1508
50-59	19	528,20	205,22	238-1152
60-69	19	551,95	256,53	113-1091
70-79	19	786,79	431,28	364-1918

In carrying out an analysis of variance, Levine's test for equal variance showed a significant difference among the different age groups (P = 0,0031), the Welch Test was, therefore, used to determine if a difference existed in the mean scores obtained by each age group. A difference was found (P = 0,0079) therefore the separate variance t with the Bonferroni correction was used to establish exactly where the differences lay (Table 6.7).

The analysis of the time scores showed that only the scores obtained for the twenty to twenty-nine year old group differed significantly from that of the seventy to seventy-nine year old group and only at the 5% level of significance (P = 0,002).

However, the age groupings used for all the tests in the battery were used for the compilation of percentiles (see Appendix 2). In re-grouping the scores, although the younger age group now included the thirty to thirty-nine year olds, the time scores still differed from that of the seventy to seventy-nine year olds at the 5 % level (P = 0,0067) (see Table 6.8).

TABLE 6.7 DIFFERENCES IN TIME SCORES - VMI

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,2850
20-29	40-49	0,0098
20-29	50-59	0,075
20-29	60-69	0,0641
20-29	70-79	0,0020 *
30-39	40-49	0,1429
30-39	50-59	0,8496
30-39	60-69	0,6713
30-39	70-79	0,0337
40-49	50-59	0,1272
40-49	60-69	0,2186
40-49	70-79	0,4402
50-59	60-69	0,7547
50-59	70-79	0,0261
60-69	70-79	0,0505

* P < 0,05

TABLE 6.8 DIFFERENCES IN RE-GROUPED TIME SCORES - VMI

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-59	0,0408
20-39	60-69	0,2544
20-39	70-79	0,0067 *
40-59	60-69	0,44891
40-59	70-79	0,1150
60-69	70-79	0,0505

* P <0,05

6.4.3. Discussion Of The Results Of Testing Normal Individuals

The results of testing normal individuals across a wide age spectrum, showed that older adults tend to score less well than younger adults on tests of perceptual function. This is demonstrated in Table 6.2 where it can be seen that there is a decrease in the accuracy scores with increasing age. Statistical analysis of the results of the VMI test revealed a significant difference between the scores obtained for the twenty to fifty-nine year old group and those obtained for the sixty to seventy-nine year old group. This finding was confirmed by Botwinick (1981) who stated that older adults tend to score less well in tasks requiring fluid intelligence i.e. perceptual function, as discussed in Chapter Two of this thesis.

It appeared that tests that measured crystallised intelligence were similar to those that measured the left hemisphere function and those that measured fluid intelligence were similar to tests for the right hemisphere function. Klisz (1978) found that tests that best differentiate the young from the older adult were also those that were sensitive to right hemisphere damage. The VMI test appeared to be a test for the right hemisphere function in that it evaluated overall contour of form as well as drawing and constructive ability.

Schaie and Schaie (1977) maintained that the tests of old people might look like those of the right hemisphere patients and this was confirmed in the examples of the VMI test completed by a normal seventy year old (see Figure 6.7) when compared to that completed by a right hemisphere C.V.A. patient (see Figure 6.8). The tests completed by normal elderly individuals showed fragmentation, loss of perspective and incomplete drawings just as did those of the patients. These types of errors were also found by Farver and Farver (1982) who found statistically significant age effects in drawings done from command. Although drawing from command was not directly comparable to the VMI test, Farver and Farver (1982) found that there was poor reproduction of angles, loss of perspective, rotation and errors of spatial relations as shown in Figure 6.7.

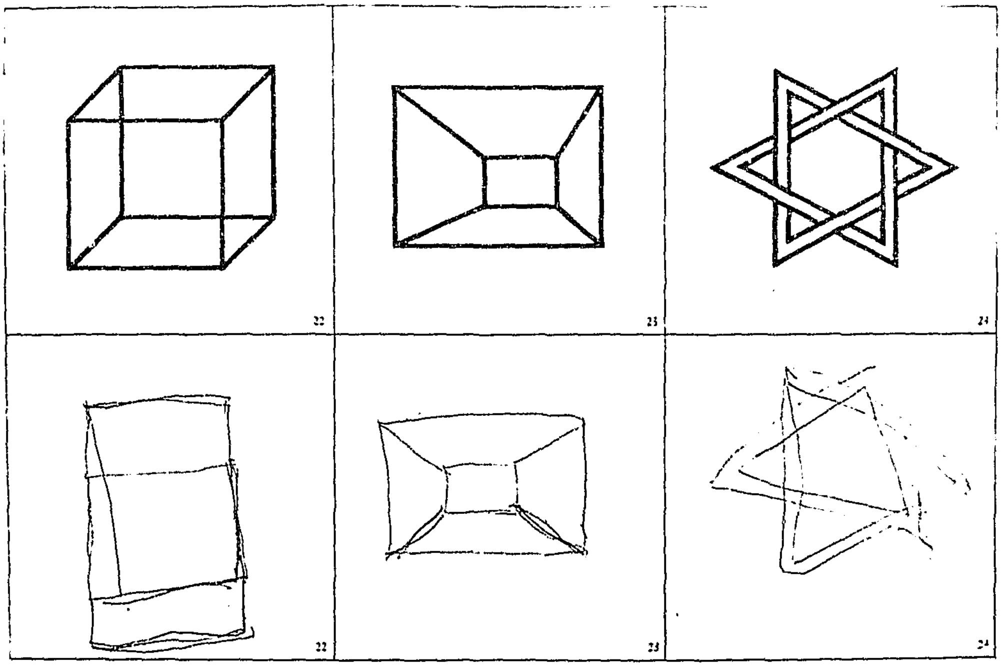


FIGURE 6.7 THE NORMAL INDIVIDUAL AGED SEVENTY YEARS

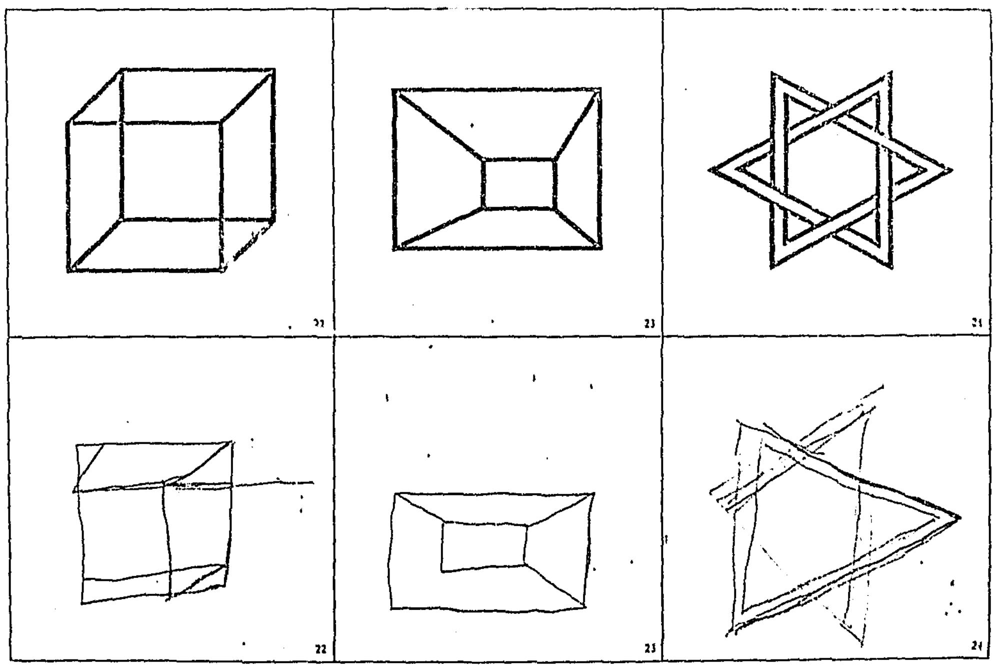


FIGURE 6.8 THE C.V.A. PATIENT

Studies which have appeared in the occupational therapy literature are inconsistent in assessing the effects of ageing on tasks involving skills of visual-motor integration. Taylor (1968) and Whiting et al (1985) found no age related changes in their copy forms tests whereas Kaplan and Hier (1982) found that there were age related changes, from the age of sixty-eight years, on Reys Complex Figure Test. Kaplan and Hier also found that there was no evidence of an age x diagnosis effect on this test, thus suggesting that *"the greater visuo-spatial deficits found in the older right hemisphere damaged patients are due to the effect of age on visuo-spatial ability ..."* (Kaplan and Hier 1982:319). Having examined the published tests that showed no age related changes it appeared that the simplistic nature of the test forms made them less sensitive to changes with age. If this were the case then they would probably be also be less sensitive to changes from brain injury.

6.5. RESULTS OF TESTING PATIENTS (POPULATION FOUR)

6.5.1. Accuracy Scores

The means and the standard deviations were calculated for the age groups twenty to fifty-nine years and sixty to seventy-nine years as these were the age groups within which no statistical difference occurred in the accuracy scores obtained by the normal individuals (see Table 6.9).

TABLE 6.9 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - VMI

AGE	n	MEAN/SECS	SD	RANGE
20-59	17	15,12	3,55	7 - 19
60-79	14	11,25	5,35	5 - 18

6.5.2. Time Scores

The means and the standard deviations were calculated for the age groups twenty to sixty-nine and seventy to seventy-nine years as these were the ages within which no statistical difference occurred in the time scores obtained by the normal group (see Table 6.10).

TABLE 6.10 MEAN AND SD FOR TIME SCORES FOR PATIENTS - VMI

AGE	n	MEAN/SECS	SD	RANGE
20-69	27	613,15	233,88	227-1214
70-79	4	940,0	296,37	570-1180

6.5.3. Comparison Between Patients And Normal Subjects

6.5.3.1. Accuracy Scores

The mean of the accuracy scores obtained by the patients was then compared with that obtained by the normal individuals to determine whether there was a statistical difference between the scores obtained by the two groups. If Levine's test for equal variance was significant then the Welch Test was used for testing the significance of the differences between the two groups. If Levine's test was not significant, the variances could be assumed to be equal and the F Test was used. In the case of the VMI, Levine's test was significant therefore the Welch Test was used. This showed that the differences in scores obtained between the patients and the normal subjects for both age groups was significant (see Table 6.11).

TABLE 6.11 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - VMI

AGE	LEVINE'S TEST P VALUE	F TEST P VALUE
20 - 59	0,0129	P= 0,000***
60 - 69	0,2340	P= 0,0003***

*** P < 0,001

6.5.3.2. Time

The mean of the time scores obtained by the patients was similarly compared with that obtained by the normal group (see Table 6.12). In both age groups the differences in the time taken to complete the test did not reach significance.

TABLE 6.12 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - VMI

AGE	LEVINE'S TEST P VALUE	THE WELCH TEST P VALUE
20 - 69	0,4406	P= 0,1389
70 - 79	0,4566	P= 0,4695

6.5.4. Analysis Of Tests Executed By Patients

a) Accuracy Scores Ages Twenty to Thirty-Nine Years - The accuracy scores ranged from ten to nineteen i.e. between the first percentile and the 22,5 percentile with only a score of nineteen approaching normality. It was interesting to note that if the percentile tables for the seventy to seventy-nine year olds were used, this group of patients would have achieved scores ranging from 2nd percentile to 84th percentile. Thus, a score of nineteen obtained by the twenty to thirty-nine year olds indicated performance just within that expected of a normal individual but the same score obtained by a seventy to seventy-nine year old would put him in the 84th percentile i.e. in the high normal range. Mistakes made by the twenty to thirty-nine year olds revealed spatial disorders and poor planning of direction, but they were not as severe as those seen in the other age groups (see Figure 6.9).

b) <u>Ages Forty to Fifty-Nine Years</u> - The accuracy scores of this age group ranged from seven to nineteen i.e. from below the 1st percentile to the 24th percentile, thus a patient receiving the highest score fell into the low normal range. In this age range the errors became more obvious.

c) <u>Ages Sixty to Sixty-Nine Years</u> - The accuracy scores ranged from fifteen to twenty-two i.e. from below the 1st percentile to the 100th percentile. This was the only group in which one patient showed above average performance and could be said to have no difficulty in executing the test. This group as a whole showed a high incidence of the type of errors that could indicate apraxia (see Figure 6.10).

d) <u>Ages Seventy to Seventy-Nine Years</u> - The accuracy scores of this age group ranged from eight to fourteen. However, two of the six patients were unable to do the test, one of them was untestable on all measures and the other had extreme visual neglect, therefore, testing was discontinued. The percentiles ranged from below the first percentile to the twenty-first percentile thus a score of fourteen fell within the low normal range. Mistakes made by this age group were severe (see Figure 6.11) and showed problems of left neglect which could not be compensated for by moving the page within the field of vision. Severe problems with overall contour, spatial relations and directionality were also evident.

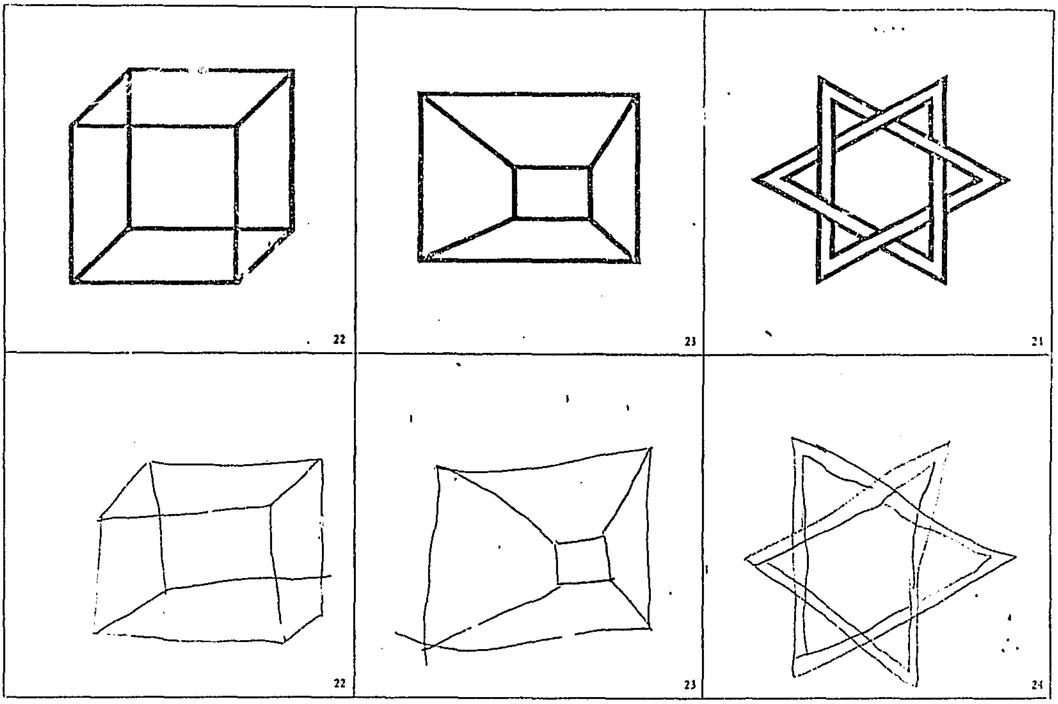


FIGURE 6.9 TWENTY YEAR OLD PATIENT

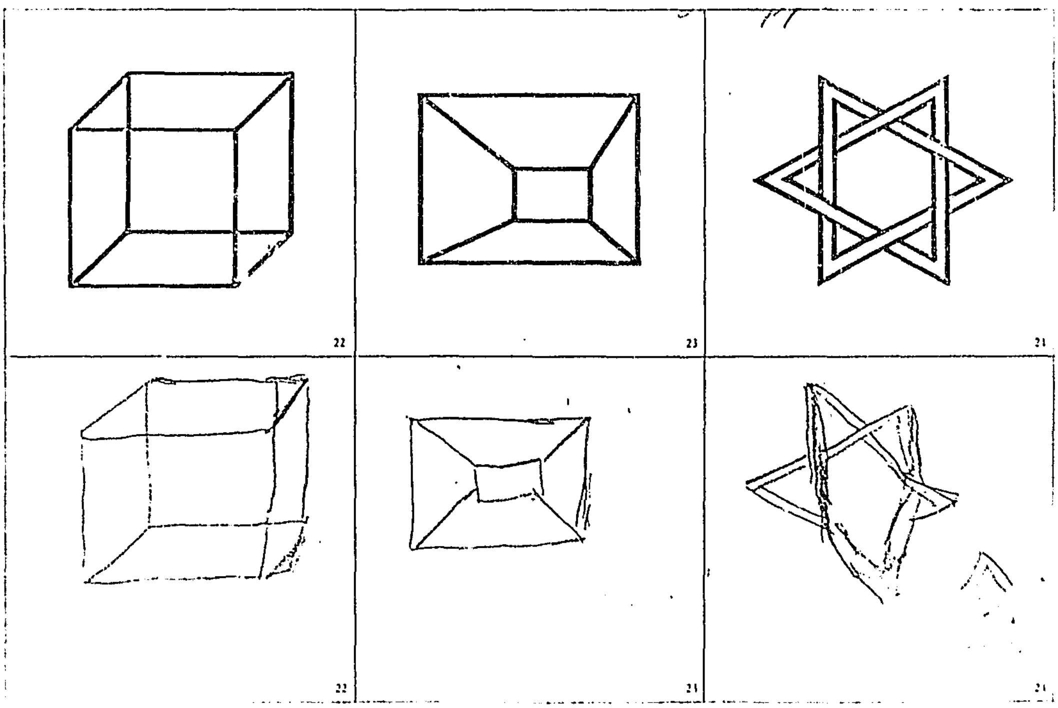


FIGURE 6.10 SIXTY YEAR OLD PATIENT

6.5.5. Problems That May Be Diagnosed From Performance On The VMI Test

Once it has been determined that the patient is dysfunctional from the score achieved on the test, the features of the patients' drawings can be examined to help determine the specific type of problem being experienced. The types of problems that could be determined through the mistakes made on the drawings were as follows:-

- Overall Contour. The VMI was included for the specific purpose of evaluating this skill. The subject was required to analyse the elements that make up the form and to reproduce these to form an outline. A patient's inability to do this is seen in Figure 6.12

- Directionality. The VMI test was said to measure this skill as rated by the fifteen occupational therapists. Problems in this area are seen in figure 6.13 and 6.14.

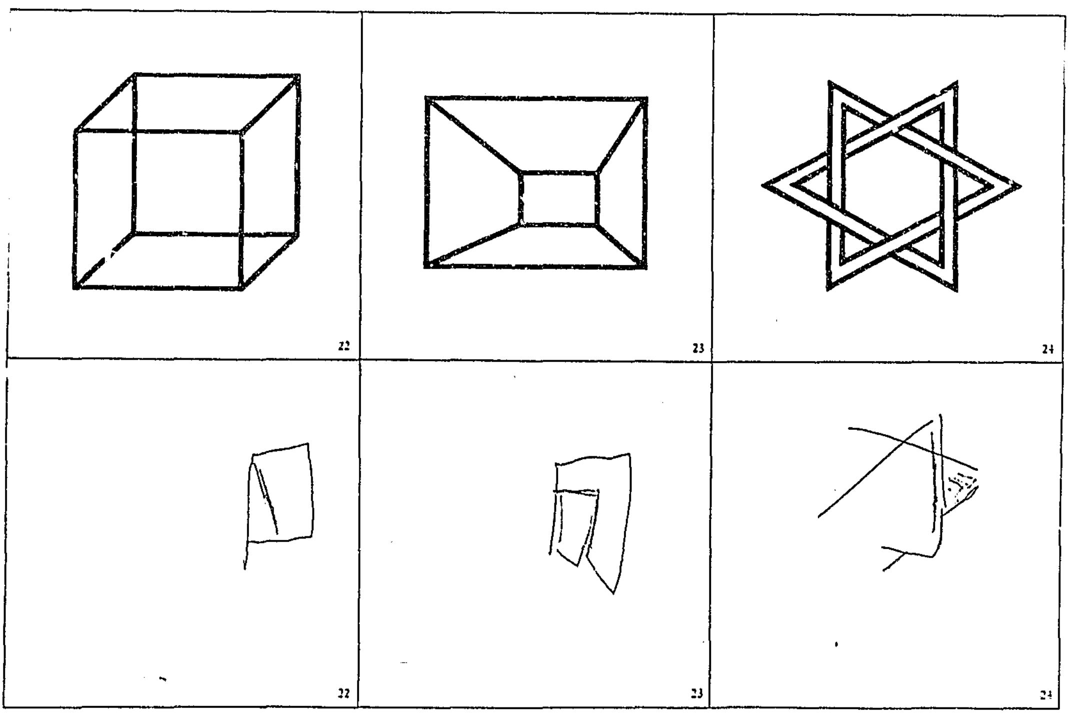


FIGURE 6.11 SEVENTY YEAR OLD PATIENT

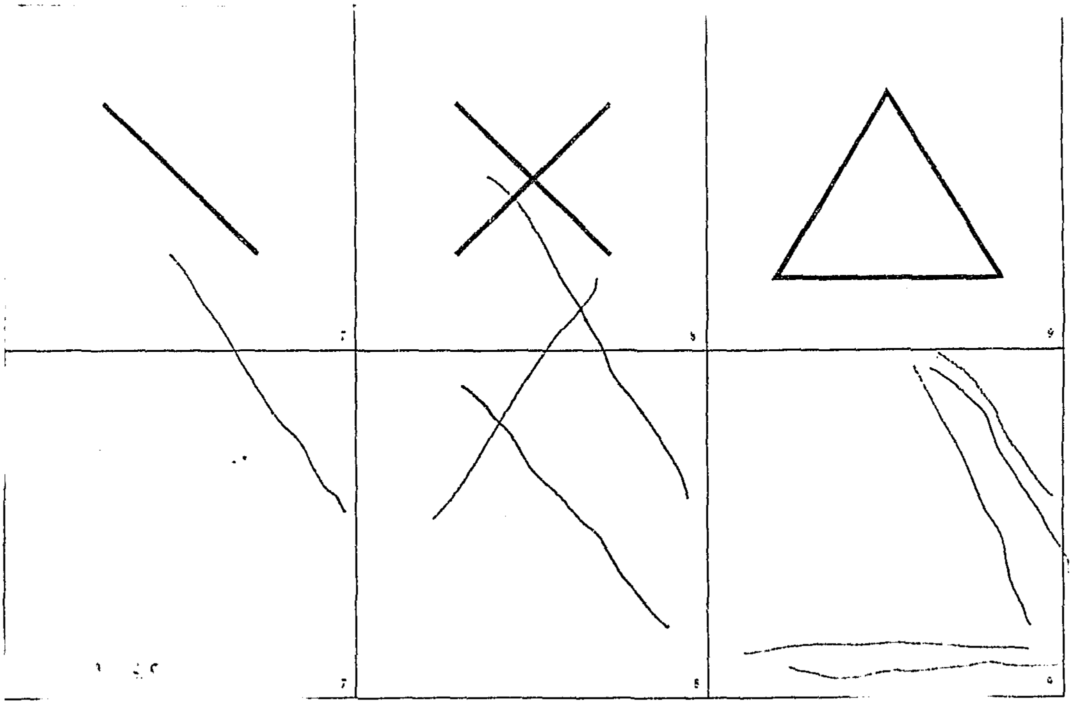


FIGURE 6.12 PROBLEMS IN OVERALL CONTOUR

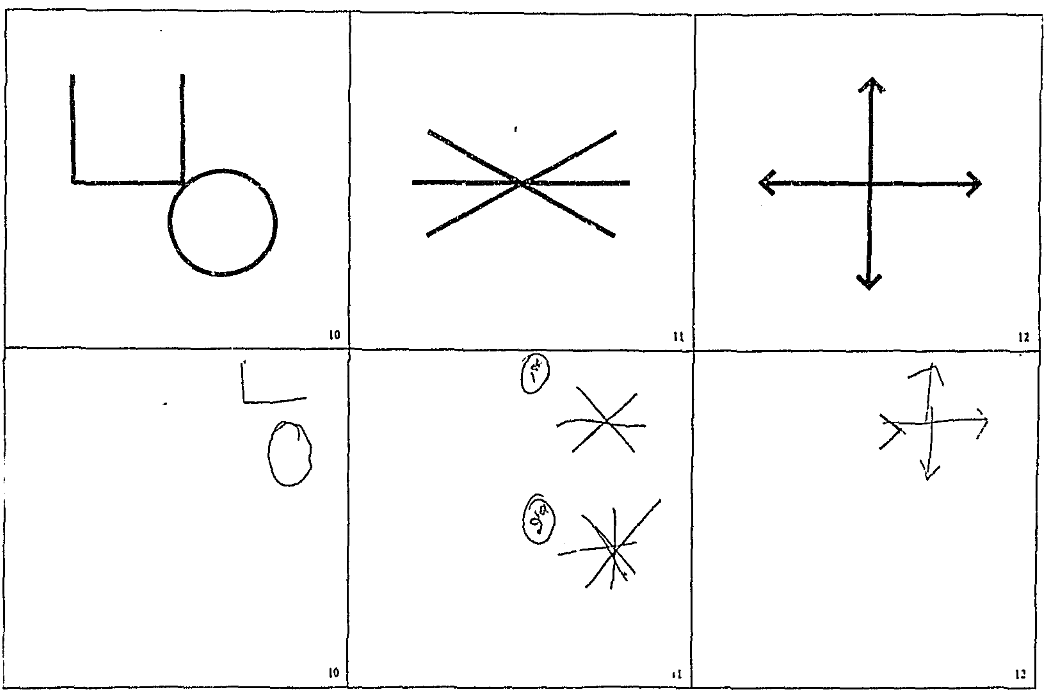


FIGURE 6.13 PROBLEMS IN DIRECTIONALITY

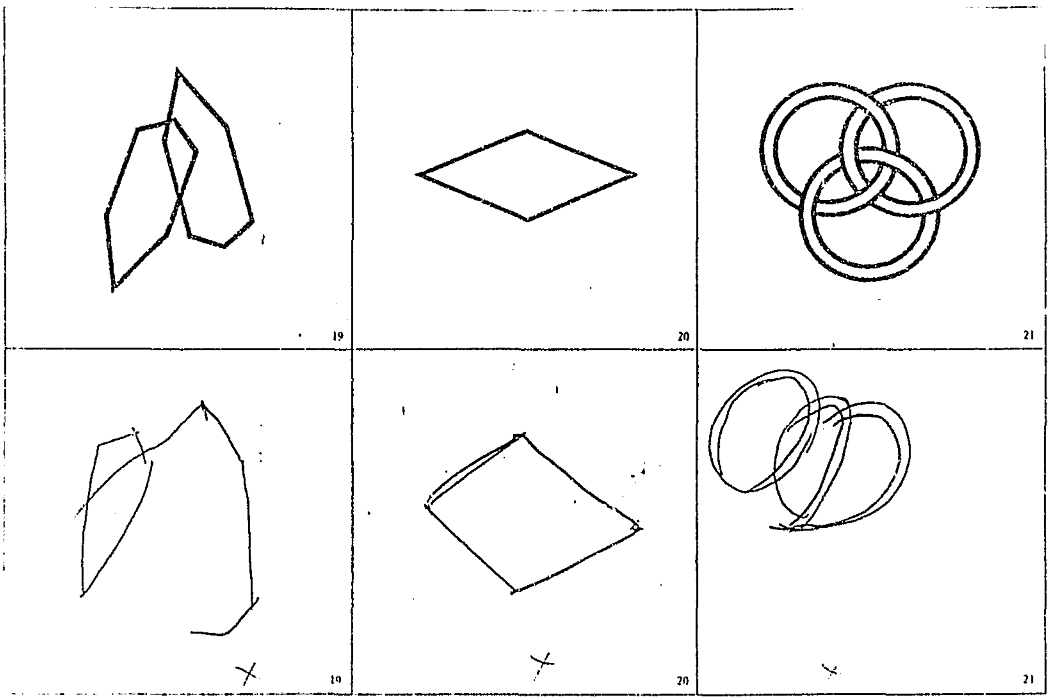


FIGURE 6.14 PROBLEMS IN DIRECTIONALITY

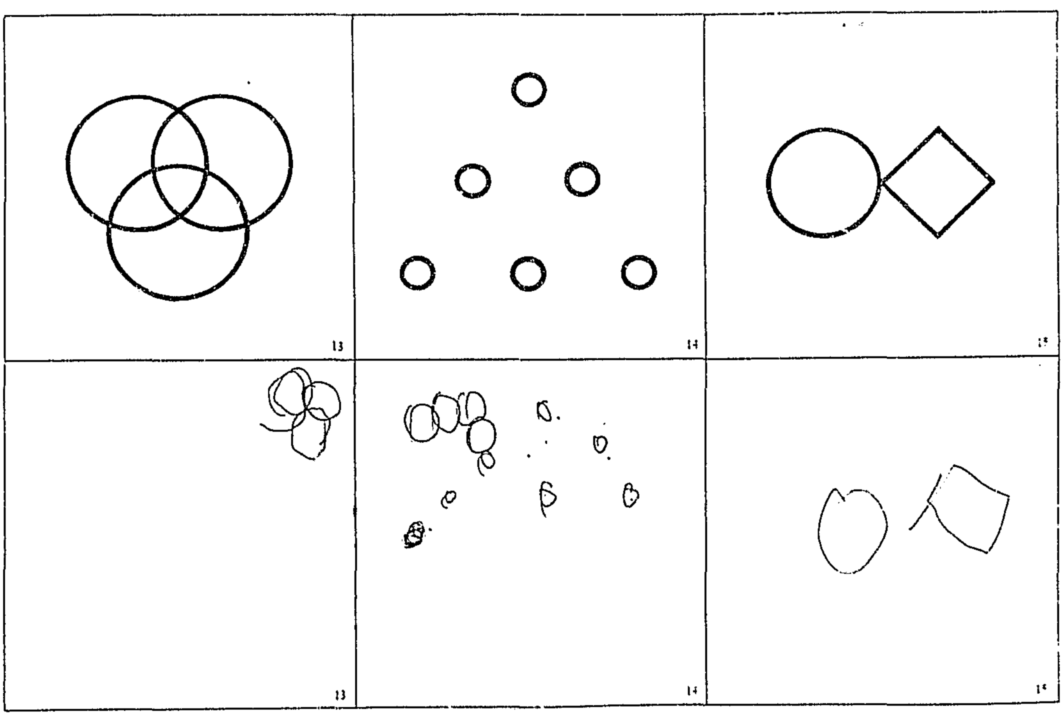


FIGURE 6.15 PROBLEMS IN SPACE RELATIONS

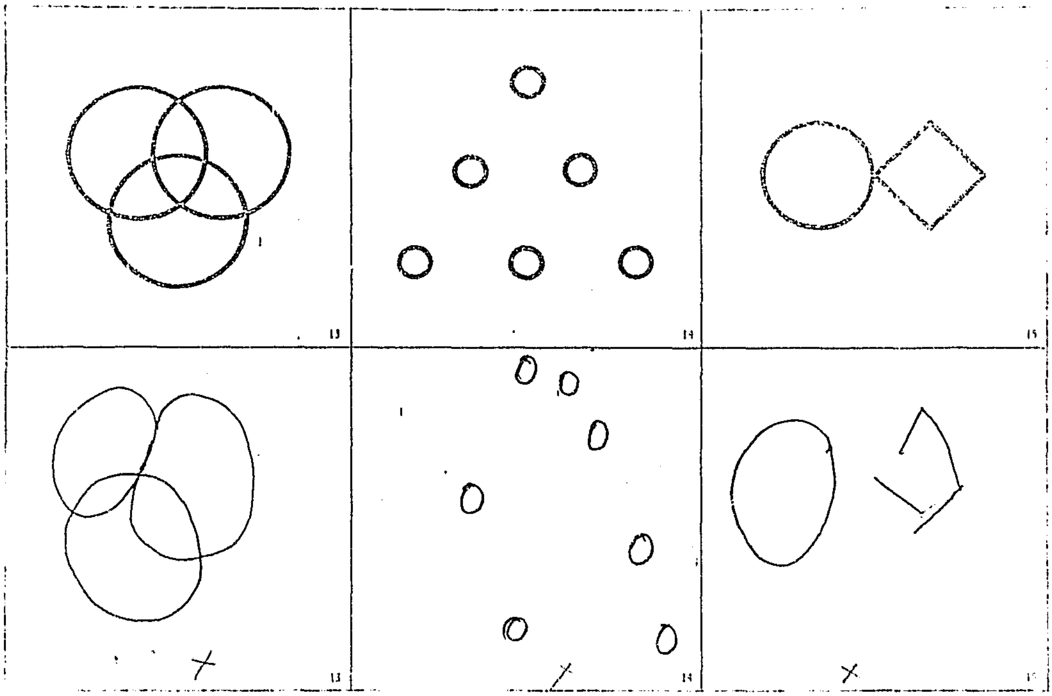


FIGURE 6.16 PROBLEMS IN SPACE RELATIONS

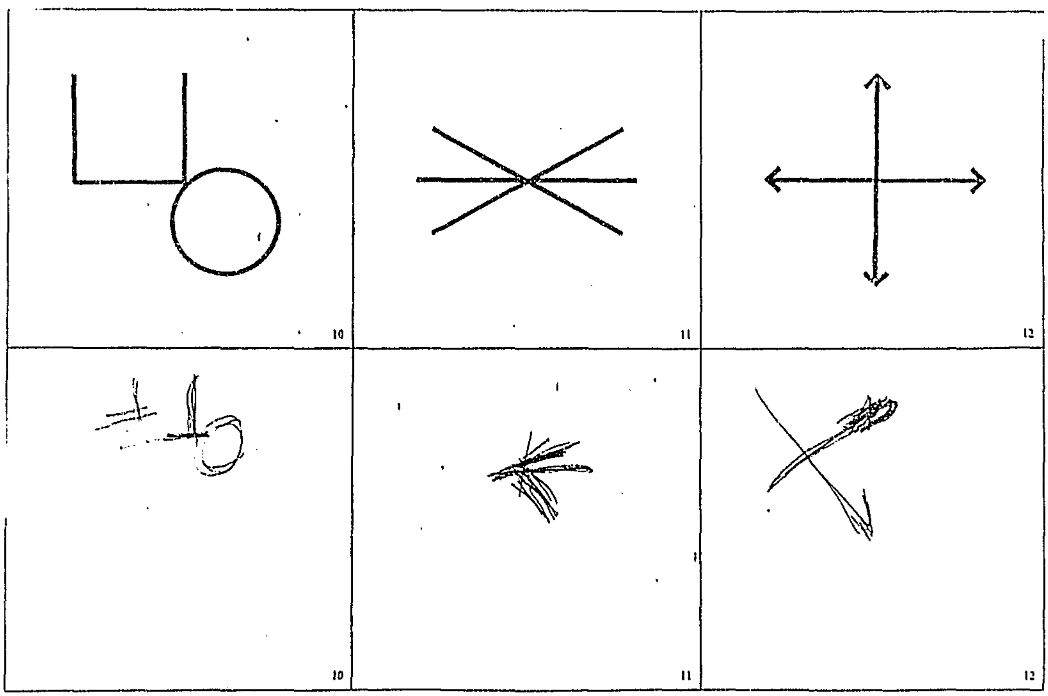


FIGURE 6.17 DRAWING APRAXIA

- Spatial Relationships. The VMI test was said to measure this skill as rated by the fifteen occupational therapists. Problems in this area are seen in Figures 6.15 and 6.16.
- Apraxia. Constructional apraxia in the field of drawing is clearly seen in Figure 6.17, which shows pronounced perseveration.

6.6. ITEM ANALYSIS

6.6.1. Item Difficulty Index

The difficulty index of an item *"is the proportion of individuals who answered the item correctly"* (Downie and Heath 1974: 249). Thus, the higher the difficulty index the easier the item. The difficulty index was calculated for the normal subjects for the age groups between which no statistical differences were found in the accuracy scores i.e. twenty to fifty-nine year olds and sixty to seventy-nine year olds (see Table 6.13).

Items in order of difficulty for the normal subjects were as follows:-

Ages Twenty to Fifty-Nine - (one to seven), (nine, thirteen, eighteen), (ten, fourteen), seventeen, eight, eleven, fifteen, twenty-two, twenty-three, nineteen, twelve, twenty, twenty-one, sixteen and twenty-four.

Ages Sixty to Seventy-nine - (One to seven), nine, thirteen, eighteen, fourteen, ten, eleven, eight, seventeen, fifteen, twenty-three, twelve, twenty-two, twenty, nineteen, twenty-one, sixteen, twenty-four. The numbers bracketed together have the same difficulty index.

The poor ability of all age groups to reproduce the seemingly simple figure of the diamond (item sixteen) was noticeable. Performance on the horizontal diamond (item twenty) was only marginally better. Failure in both these figures was attributed to the lack of ability to accurately estimate the degree of each angle, a factor which contributed to the scoring, whereas this was not the case in other figures that contained angles.

The difficulty index for the responses obtained from the patient group, was also calculated (see Table 6.14). Fisher's Exact test was used to calculate whether there were statistical differences in the responses obtained on each item for the two groups i.e. the normal and the patient group (see Table 6.14). This was done to help to determine which items should be retained in the final version.

TABLE 6.13 DIFFICULTY INDEX - VMI

ITEM NO.	DIFFICULTY p 20-59 YRS	DIFFICULTY p 60-79 YRS
1-7	1,00	1,00
8	0,89	,076
9	0,99	1,00
10	0,94	0,82
11	0,87	0,79
12	0,65	0,53
13	0,99	0,97
14	0,94	0,84
15	0,82	0,63
16	0,56	0,29
17	0,90	0,66
18	0,99	0,92
19	0,71	0,34
20	0,64	0,45
21	0,63	0,34
22	0,86	0,53
23	0,78	0,55
24	0,51	0,21

Items in order of difficulty for the patients were as follows:

Ages twenty to fifty-nine years - (one, two, three, five, seven, nine), (four, six), thirteen, (eight, ten seventeen), eleven, fourteen, eighteen, (twelve, twenty-two), fifteen, (sixteen, twenty-three), (nineteen, twenty, twenty-one), twenty-four.

Ages sixty to seventy-nine years - (one, two, five, seven), (three, four nine), six, eight, thirteen, (eleven, fourteen), eighteen, twelve, (seventeen, twenty-three), fifteen, (ten, twenty-two), twenty-one, (sixteen, twenty), twenty-four.

Only items ten, eleven, thirteen, fourteen, fifteen, seventeen, eighteen, nineteen, twenty, twenty-one, twenty-two, twenty-three and twenty-four showed statistical differences in the performance of the patients and of the normal subjects and were therefore important for the analysis of test results. It was interesting to note that there were no significant differences between the performance of the patient group and the normal group for items twenty to twenty-three in the ages sixty to seventy-nine years. This might have been due to the fact that normal individuals sixty to seventy-nine years old did not perform as well as the younger age group on these items. Should an individual be brain damaged at the age of sixty to seventy-nine years, his ability to execute these items would not be very much more impaired.

6.6.2. Item Discrimination

This is the extent to which the item discriminates between those who can do well on the test and those who cannot. Two ways of determining item discrimination on the normal subjects were used.

6.6.2.1. Testing the differences between two proportions

The responses obtained from those individuals whose overall score was in the top 27% of answers were compared with those in the lowest 27%.

TABLE 6.14 DIFFICULTY INDEX - PATIENT GROUP - VMI

ITEM	DIFFICULTY p 20 - 59 YRS	ITEM NO.	DIFFICULTY p 60 - 79 YRS
1-3	1,0	1-2	1,0
4	0,94	3	0,93
5	1,0	4	0,93
6	0,94	5	1,0
7	1,0	6	0,86
8	0,71	7	1,0
9	1,0	8	0,64
10	0,71 *	9	0,93
11	,53 **	10	0,21 ***
12	0,47	11	0,50 *
13	0,77 **	12	0,37
14	0,65 **	13	0,64 **
15	0,41 ***	14	0,50 *
16	0,35	15	0,40 **
17	0,71 *	16	0,07
18	0,58 ***	17	0,36
19	0,18 ***	18	0,50 ***
20	0,18 ***	19	0,29 *
21	0,18 ***	20	0,07
22	0,47 **	21	0,14
23	0,35 **	22	0,21
24	0,00 ***	23	0,36
		24	0,00 ***

* P <0,05 ** P < 0,01 *** P <0,001

Table 6.15 for ages twenty to fifty-nine years and sixty to seventy-nine years gives the result of the comparison between the scores, using Fisher's Exact test.

On examining table 6.15 it was apparent that items eight, twelve, fifteen, sixteen, seventeen, nineteen, twenty, twenty-one, twenty-two, and twenty-four were the items that successfully discriminated between those individuals who could do the test and those who could not and should therefore be included in the final test form.

TABLE 6.15 ITEM DISCRIMINATION - VMI

ITEM	20-59 YRS	ITEM	60-79 YRS
1-7	-	1-7	-
8	0,3433	8	0,0325 *
9	-	9	-
10	0,6060	10	0,4737
11	1,000	11	0,5820
12	0,002 ***	12	0,0055 **
13	-	13	-
14	0,1069	14	-
15	0,0670	15	0,0198 *
16	0,000 ***	16	0,0198 *
17	0,3433	17	0,0108
18	-	18	0,4737
19	0,0014 **	19	0,0198 *
20	0,0000 ***	20	0,1698
21	0,0122 *	21	0,0055 **
22	0,0013 **	22	0,0698
23	0,6965	23	0,0573
24	0,0000 ***	24	0,0007 **

* P = <0,05 ** P = <0,01 *** P = <0,001

6.6.2.2. Correlational Approach

This approach demonstrated the relationship of the individual responses to the total test score using Flannigan's method i.e. it reflected how well the item is doing what the test itself is doing (Downie and Heath 1974). For the results see table 6.16 and 6.17.

From the results it appeared that for the age group twenty to fifty-nine years the scores obtained on items eight, ten, twelve, fourteen, fifteen, sixteen, seventeen, nineteen, twenty, twenty-one, twenty-two, and twenty-four correlated significantly with the total test score.

TABLE 6.16 BI-SERIAL CO-EFFICIENT OF CORRELATION BETWEEN ITEM AND TOTAL SCORE 20-59 YEARS - VMI

ITEM NO.	UPPER 27%	LOWER 27%	r VALUE
1-7	-	-	-
8	,95	,81	,51 ***
9	-	-	-
10	,95	,86	,52 ***
11	,81	,76	,09
12	,95	,38	,78 ***
13	-	-	-
14	1,00	,81	,55 ***
15	,91	,62	,68 ***
16	,86	,14	,78 ***
17	,95	,81	,51 ***
18	-	-	-
19	,95	,48	,72 ***
20	,95	,29	,82 ***
21	,76	,33	,45 ***
22	1,00	,57	,69 ***
23	,86	,76	,30 *
24	,81	,14	,68 ***

* P = < 0,05 ** P = <0,01 *** P < 0,001

Correlation coefficients for the sixty to seventy-nine year olds were significant for items eight, ten, eleven, twelve, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, twenty-one, twenty-two, twenty-three and twenty-four. Items eleven and twenty were significant at the 0.01% level of significance, all other items correlated at the 0,001% level of significance with the total test score.

TABLE 6.17 BI-SERIAL CO-EFFICIENT OF CORRELATION BETWEEN ITEM AND TOTAL SCORE 60 - 79 Years - VMI

ITEM	UPPER 27%	LOWER 27%	r VALUE
1-7	-	-	-
8	1,00	,50	,73 ***
9	-	-	-
10	1,00	,80	,51 ***
11	,90	,70	,40 ***
12	,80	,10	,70 ***
13	-	-	-
14	-	-	-
15	,90	,30	,63 ***
16	,70	,10	,63 ***
17	1,00	,40	,74 ***
18	1,00	,80	,51 ***
19	,70	,10	,63 ***
20	,60	,20	,42 ***
21	,80	,10	,70 ***
22	,80	,30	,51 ***
23	,90	,40	,56 ***
24	,80	,00	,84 ***

* P = <0,05 ** P < 0,01 *** P < 0,001

Thus it appeared that items eight, ten, eleven, twelve, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, twenty-one, twenty-two, twenty-three and twenty-four were important as they showed a high discriminative ability. As this is a published test, the items may not be changed, but note should be made of the items difficult for older adults.

6.7. TEST RELIABILITY

6.7.1. Split-Half Reliability

This determined the internal consistency of the VMI test on the responses of the normal group. The co-efficient of correlation was r = 0,72 when age group twenty to fifty-nine was considered and r = 0,81 when age group sixty to seventy-nine was considered. Beery (1982) reported split-half correlations on children that ranged from r = 0,66 - r = 0,93 with a median of r = 0,79. Reliability studies were not repeated by Beery for the 1989 version of the test.

6.7.2. Kuder-Richardson Formula 20

Using this formula, the resulting co-efficient of internal consistency on the normal group responses was r = 0,53 for the age groups twenty to fifty-nine years and r = 0,69 for the age groups sixty to seventy-nine years.

Thus the co-efficient of correlation ranged between r = 0.53 and r = 0.81, with all correlations being significant at the P<0,001 level of significance (see Table 6.18).

TABLE 6.18 RELIABILITY CO-EFFICIENTS - VMI

AGE	SPLIT-HALF	KUDER-RICHARDSON
20-59	0,72 ***	0,53 ***
60-79	0,81 ***	0,69 ***

*** P < 0,001

6.7.3. Inter-Rater Reliability

6.7.3.1. Examiners (Population Five)

Friedman's two-way analysis of variance showed that there were no statistical differences between either the accuracy or time scores obtained by the three examiners on each of the six normal individuals tested i.e. population five (P = 0,7470 for the accuracy scores and P = 0,3114 for the time scores).

6.7.3.2. Scorers

Due to the possible difficulty in scoring this test (i.e. each item had to be carefully scrutinised to determine whether the individual completing the test met all the criteria for passing each item), inter-scorer reliability was sought for the normal group.

An independent occupational therapist was taught the scoring criteria for pass and fail on each item, and scored fifty-seven tests from the total of the tests completed. The raw scores obtained by the therapist and those obtained by the candidate were compared by using the t test to establish whether there was a statistical difference between the two scorers. No statistically significant difference between the two scorers was found (P = 0,3082). The test could, therefore, be said to have inter-rater reliability. Beery (1982) found inter-rater reliability correlations to range from r = 0,93 to r = 0,98 which indicated a high degree of reliability. Correlation coefficients were not used to determine inter-rater reliability in this thesis as the t test is more accurate.

6.8. TEST VALIDITY

6.8.1. Content Validity

This was provided by a group of fifteen occupational therapists (population one) who rated the test on a zero to five scale.

The VMI test was shown to be a very much better than average test for evaluating overall contour and a better than average method of evaluating directionality and spatial relations with a motor base (see Table 6.19). The VMI test could, therefore, be said to be a valid method of evaluating visual-motor integration in the field of form (overall contour) and space.

TABLE 6.19 CONTENT VALIDITY - VMI

PERCEPTUAL SKILL	MEAN RATING
Form as mass	1,6
Overall contour	3,9 ***
Sub-units of Form	2,4
Perceptual Constancy	0,86
Figure Ground	1,3
Directionality	2,9 **
Space Relationships	2,6 **
Position in Space	2,0
Colour	0,0
Praxis	Present

** a better than average assessment *** a very much better than average assessment

6.8.2. Construct Validity

Three types of construct validity were considered i.e. age differentiation, correlation with other tests in the battery and performance of patients.

6.8.2.1. Age Differentiation

Discussion already presented, indicated that some perceptual skills showed a progressive decrease with age (Chapter Two and Chapter Six, Section 6.4.8). In the case of the VMI test the "drop off" in performance occurred at approximately sixty years of age, a statistical difference being found in the accuracy scores of the twenty to fifty-nine year olds when compared to those of the sixty to seventy-nine year olds. The VMI test would seem, therefore, to be a valid differentiation of function between older and younger age groups.

6.8.2.2. Correlation with other tests in the Battery

The VMI test correlated highest with the Dot-Line test (r = 0,71) (see Table 6.20). This is the only other paper and pencil test in the battery and some correlation between the two tests would be expected. It was interesting to note that the VMI test had a low correlation with the construction tests such as block building and stick construction thus demonstrating that drawing and construction tests do not measure the same skills (see discussion on apraxia in Chapter Seventeen). There was, however, a higher correlation between the VMI test and the Design Construction test (r = 0,61). This link appeared to exist as both tests dealt with form manipulation.

A further analysis was carried out on the tests in the battery to determine whether the tests had related factors. The tests that appeared to measure the same factor as the VMI test (see Table 6.21) were: Part-Whole, Figure-Ground, Dot-Line, Position in Space and Block Construction II.

As the factor analysis carried out was an exploratory analysis rather than a confirmatory analysis and no marker variables were included, the results could not be readily used for the purpose of construct validity.

However, it was apparent that Factor I, of which the VMI is a part could be a space factor, in which all the tests require some space visualisation for their execution.

6.8.2.3. Results of Testing Patients

The patient group performed significantly ($P<0{,}001$) poorer than the normal group in accuracy. The difference between the two groups in the time taken to complete the test was not significant. The test is therefore a valid means of testing for perceptual change in C.V.A. patients but not a valid means of evaluating changes in performance speed.

TABLE 6.20 CORRELATION WITH OTHER TESTS IN THE BATTERY - VMI

TEST	CORRELATION WITH VMI
Part-Whole	0,51
Figure-Ground	0,51
Perceptual Constancy	0,46
Dot-Line Test	0,71
Arrows	0,27
Space Relations	0,43
Position in Space	0,56
Block Construction I	0,39
Block Construction II	0,35
Stick Construction I	0,36
Stick Construction II	0,42
Design Construction	0,61
Colour	0,14

TABLE 6.21 SORTED ROTATED FACTOR LOADINGS - VMI

TEST	FACTOR I
Position in Space	0,80
Block Construction II	0,74
Part-Whole	0,73
VMI	0,63
Dot-Line	0,61
Figure-Ground	0,58

6.9. CONCLUSIONS

It was demonstrated that the test of Visual-motor Integ.ation (Beery 1967) was a valuable tool for the evaluation of visual-perceptual-motor problems in the brain damaged adult as it indicated problems in perceptual motor ability. Scores alone, although reflecting poor performance are not the only guide to dysfunction, the patient's performance and the type of mistakes that are made must be analysed so that the different aspects of V-P-M function can be confirmed. The VMI test was particularly useful in this process.

It is concluded from all the information provided in the chapter that the VMI test should be retained in the final format of the V-P-M test battery for the following reasons:-

- it was easy to administer and if the instructions were followed it was easy to score (the inter-rater reliability was good).

- it was challenging and enjoyable for adults. They did not feel that they were being given childish tests.

- it was sensitive to changes in the normal elderly person.

- it was sensitive to perceptual problems in patients of all ages.

- the items on the test can be graded from the simple to the more complex and difficult. Item analysis showed that different ages found different items difficult. This indicated that the entire test should be administered to adults, unlike the children's version where a cut off of three consecutive failures could be used. In addition this was a published test and therefore could not be changed. Certain items, however, were particularly useful for diagnosis as they discriminate well between normal people and C.V.A patients i.e. items eight, twelve, fifteen, sixteen, nineteen, twenty, twenty-one, twenty-two and twenty-four.

- it was reliable. Reliability coefficients ranged from r=0,53 to r = 0,81. These compared with reliability coefficients of the published data on children.

- it was a valid measure of V-P-M function in the area of "overall contour".

- correlation with other tests in the battery was generally low. The highest correlation being with the only other drawing test (i.e. the Dot-Line test) r=0,71 and with the Design Construction test, r=0,61. This might indicate that the VMI test provided enough information without the additional testing time needed for the other two tests.

The results of the psychometric evaluations of all the tests will be compared in Chapter Twenty-One. Their comparative usefulness and therefore the need to retain them in the battery will be discussed.

Aspects of this chapter were published in the South African Journal of Occupational Therapy 1986 16:1, 11-19 and that of 1989 19:2, 17-25 (see Appendix 5).

CHAPTER 7

THE PART-WHOLE TEST

7.1. INTRODUCTION

Part-whole perception requires that the subject "makes sense" out of incomplete and fragmented visual stimuli. This ability develops from the perception of simple form recognition and is a further progression of the skills evaluated in Chapters Five and Six. Kephart (1960) referred to this skill as constructive form i.e. the individual must construct an integrated form out of the elements provided. This skill is essential as we do not have time to respond to all the elements in our environment. For example, when reading we do not see each letter, much less the parts of each letter, yet we construct words from the number of elements that we see.

In part-whole perception, the parts are available to be seen, none are hidden, fragmented pieces are drawn together so that the whole can be interpreted.

Visual closure in comparison is a slightly more difficult skill in that parts of a form or picture, scene, or word may be missing, and the individual has to visualise the missing part i.e. choose the missing part in his mind and interpret it correctly. The two terms i.e. part-whole and visual closure are sometimes used incorrectly in the literature to denote the same thing.

Part-whole perception for the purpose of this thesis is defined as the ability to "make sense" out of incomplete and fragmented stimuli. All the component parts are presented to the person taking the test, but in a fragmented way. For a description of the Part-Whole test see Chapter Four (4.6.3.) and Appendix 1, Section 2.3.

7.2. REVIEW OF SIMILAR TESTS

7.2.1. Tests Involving Fragmented Visual Stimuli

The Hooper Visual Organisation Test (Hooper 1957) consists of drawings of objects which have been cut up. The subject is asked to name and or write the name of the object in a space provided below each drawing, thus requiring verbal skills. The test was originally constructed to identify patients with organic brain conditions and published data relate to its capacity to do this. The Hooper Visual Organisation Test was included as part of the Sensory-Motor Integration Test Battery for C.V.A. Clients (Jongbloed et al 1986). Jongbloed and her co-workers, however, evaluated this test as a measure of visual ability and form-space perception.

The Minnesota Paper Form Board Test (Likert and Quasha 1970) makes use of fragmented drawings of geometrical figures such as circles. Like the Part-Whole test of the V-P-M battery, this is a multiple choice test but with a time limit. There are average scores for various adult groups. Statistical analysis of the data collected on the Minnesota Paper Form Board Test indicated a strong association between this test and construction tasks.

7.2.2. Tests Involving Incomplete Visual Stimuli

The Visual Closure sub-test of the Illinois Test of Psycholinguistic Abilities (Kirk et al 1968) is a well known test of visual closure. However, normative data are only available for ages two years and four months to ten years and three months. More recently Hammill et al (1993), (Visual Closure sub-test of the Developmental Test of Visual Perception) and Gardner (1982, 1988 and 1992) (Visual Closure sub-test of the Test of Visual Perceptual Skills - Non-Motor) have provided sub-tests to measure part-whole ability. Normative data for individuals older than seventeen years and eleven months were not provided by either of these tests.

The Incomplete Letters test of the Visual Object and Space Perception Battery (VOSP) (Warrington and James, 1991) is a more recently published test that was standardised for ages twenty to sixty-nine years. The scores obtained by the non-brain damaged group revealed changes after the age of fifty years on this test. The Incomplete Letters test in the battery is a test involving incomplete visual stimuli.

7.3. Reasons for compiling a new test

It can be seen from the above that there are few activities and tests available to evaluate part-whole perception and they are generally not included in perceptual test batteries. The Incomplete Letters test of the VOSP (Warrington and James 1991) became available after the commencement of this thesis. The Hooper Visual Organisation Test was not suitable for the V-P-M Battery as the verbal responses alter the output processes required in the V-P-M Battery.

The Part-Whole test of the V-P-M Battery requires the subject to choose from three alternatives, the piece that is missing from a shape. The shapes of the Form Matching test were used for this purpose (see Appendix 1, section 2.3.).

7.4. Results of testing normal individuals

7.4.1. Accuracy Scores

The means and the standard deviations were calculated for the normal group as a whole (Table 7.1) and for each individual age group (Table 7.2). The score of one subject in the sixty to sixty-nine year old group was excluded due to a missing score for one item.

A demonstrable change in the scores with increasing age is noted. The Pearson Chi-square Test was used to determine whether this change was statistically significant. A cut point score of nine was used with the percentage obtaining a score greater than and lesser than the cut point being used to calculate the Chi-square value (Table 7.3).

As a significant difference at the $P<0,05$ level of significance was found to occur among the age groups constituting the normal group, further analysis using Fisher's Exact test was undertaken to determine where the differences lay. The age groups twenty to fifty-nine years and sixty to seventy-nine years were used for this analysis as a natural division of the scores occurred between those two groups (Table 7.4).

TABLE 7.1 MEAN AND SD OF ACCURACY SCORES - P-W

n	MAX SCORE	MEAN	SD	RANGE
117	12	9,39	2,14	3-12

TABLE 7.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - P-W

AGE	n	MEAN	SD	RANGE
20-29	20	10,35	1,46	7-12
30-39	21	10,10	1,64	8-12
40-49	19	10,0	1,56	7-12
50-59	19	9,42	2,29	3-12
60-69	18	8,22	2,18	4-11
70-79	20	8,15	2,58	3-12

TABLE 7.3 COMPARISON OF ACCURACY SCORES - P-W

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 9	30,0	38,1	31,6	42,1	72,2	60,0	45,3
Greater than 9	70,0	61,9	68,4	57,9	27,8	40,0	55,7

Pearson Chi-square Value 10,860 P= 0,0542 P< 0,05

TABLE 7.4 COMPARISON OF ACCURACY SCORES AGES 20 - 59 AND 60 -79 YEARS - P-W

AGE	20-59	60-79	TOTAL
Less than 9	35,4	65,8	45,3
Greater than 9	64,6	34,2	54,7

Fisher's Exact Test P = 0,0028 P < 0,01

A significant difference was found between the above age groups and a drop off in function appeared to occur at sixty years of age. Accordingly the null hypothesis that decreased ability in V-P-M function will not be found with advancing age was rejected at the P=0,01 level of significance. However, the age groups used for other tests were used for the purpose of compiling percentiles i.e. twenty to thirty-nine years, forty to fifty-nine years, sixty to sixty-nine years and seventy to seventy-nine years (see Appendix 2).

7.4.2. Time Scores

The means and the standard deviations were calculated for the time scores for the group as a whole (Table 7.5) and for each age group individually (Table 7.6).

TABLE 7.5 MEAN AND SD OF TIME SCORES - P-W

n	MEAN/SECS	SD	RANGE
117	225,02	99,21	9,60 - 607,0

Levine's test for equal variance established that there was a significant difference in the variances in the groups (P=0,0006) and the Welch Test demonstrated that there was a significant difference in the means obtained by the different ages (P=0,0000). The separate variance t was, therefore, used with the Bonferroni correction for multiple comparisons, to establish where the differences lay (See Table 7.7).

TABLE 7.6 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - P-W

AGE	n	MEAN/SECS	SD	RANGE
20-29	20	174,45	56,73	96-290
30-39	21	166,91	46,0	99-286
40-49	19	217,95	80,14	132-457
50-59	19	215,21	54,01	133-307
60-69	18	253,0	117,75	149-607
70-79	20	327,45	122,64	169-587

On the basis of the differences found between the various age groups it was possible to regroup the data, as previously described, for the purposes of compiling percentiles (Table 7.8), although no differences were shown to exist between the forty to fifty-nine year olds and the sixty to sixty-nine year olds.

TABLE 7.7 DIFFERENCES IN TIME SCORES - P-W

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,6437
20-29	40-49	0,0602
20-29	50-59	0,0273
20-29	60-69	0,0167
20-29	70-79	0,000 ***
30-39	40-49	0,0214
30-39	50-59	0,0046
30-39	60-69	0,0081
30-39	70-79	0,0000 ***
40-49	50-59	0,9205
40-49	60-69	0,3008
40-49	70-79	0,0022 *
50-59	60-69	0,2260
50-59	70-79	0,0009 *
60-69	70-79	0,0644

* $P<0{,}05$ ** $P<0{,}01$ *** $P<0{,}001$

As can be seen from the above the null hypothesis of no increase in the speed taken to complete the test with advancing age can be rejected.

TABLE 7.8 DIFFERENCES IN RE-GROUPED TIME SCORES - P-W

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-59	0,0011 **
20-39	60-69	0,0099 **
20-39	70-79	0,0000 ***
40-59	60-69	0,2348
40-59	70-79	0,0009 **
60-69	70-79	0,0644

*P< 0,05 ** P <0,01 *** P <0,001

7.4.3. Discussion Of The Results Of Testing Normal Individuals

The review of the literature revealed that there were few tests for the assessment of part-whole ability but also that there were no studies available on age related changes apart from those reported by Warrington and James (1991) on the VOSP. A study carried out by Shan-Shan et al (1987) compared the ability of learning disabled and normal young men on the Test of Visual Perceptual Skills (Gardner, 1982). This test included part-whole ability, however, no conclusions about age related changes were drawn from the study.

7.5. RESULTS OF TESTING PATIENTS

7.5.1. Accuracy Scores

The means and the standard deviations were calculated for age groups twenty to fifty-nine and sixty to seventy-nine as those were the age groups within which no statistical difference occurred in the scores obtained by the normal group (Table 7.9).

TABLE 7.9 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - P-W

AGE	n	MEAN	SD	RANGE
20-59	17	8,18	1,89	5 - 11
60-79	13	7,15	2,23	3 - 11

7.5.2. Time Scores

The means and the standard deviations were calculated for the time scores in the age groups within which no statistical difference was found in the normal group (Table 7.10).

TABLE 7.10 MEAN AND SD FOR TIME SCORES FOR PATIENTS - P-W

AGE	n	MEAN/SECS	SD	RANGE
20-39	7	294,57	213,75	116-745
40-69	19	316,26	148,30	62-606
70-79	4	502,75	151,14	376-691

7.5.3. Comparison Between Patients And Normal Subjects

7.5.3.1. Accuracy Scores

The mean of the accuracy scores obtained by the patients were compared to those obtained by the normal subjects (Table 7.11).

It was interesting to note that there was no difference between the scores obtained by the normal subjects and the patients in the sixty to seventy-nine year old group.

TABLE 7.11 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - P-W

AGE	LEVINE'S TEST P-VALUE	F TEST P-VALUE
20-59	0,3931	0,0003 ***
60-79	0,9235	0,2454

*** P < 0,001

7.5.3.2. Time Scores

As with the accuracy scores, the time scores for the normal group and for the patients were compared (Table 7.12). There was a difference in the time taken to complete the test between patients and normal subjects in the forty to sixty-nine year group as well as in the older seventy year old group but there were no differences between the two groups in the younger age bracket.

TABLE 7.12 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - P-W

AGE	LEVINE'S TEST P VALUE	P VALUE
20-39	0,0002	0,1797 Welch Test
40-69	0,0056	0,0249* Welch Test
70-79	0,5388	0,0194* F Test

*: P < 0,05

7.5.4. Analysis Of Tests Executed By Patients

In analysing the percentile equivalent scores obtained by the patients, the following was observed:-

a) Out of the seven patients in the twenty to thirty-nine year old group, two obtained scores below the fifteenth percentile and could, therefore, be judged to have problems with part-whole perception.

b) Five out of the ten patients in the forty to fifty-nine year old group obtained scores below the fifteenth percentile.

c) Only one of the nine patients tested in the sixty to sixty-nine year old group fell below the fifteenth percentile. However, one patient was untestable and his zero scores were not included in the calculations.

d) Only one patient out of the four patients tested in the seventy to seventy-nine year old group fell below the fifteenth percentile. However, two patients in this group were untestable.

7.5.5. Problems That May Be Diagnosed From Performance On The Part-Whole Test

As this is a test that requires the individual to put parts together to make a whole, inability to complete the test would indicate a problem with part-whole perception. The factor analysis indicated that the rotation required of the forms before correct identification could take place also indicated difficulty in space perception. The test does, however, also determine the presence of a hemianopia and left neglect of visual space when answers to the left of the response section are neglected.

7.6. ITEM ANALYSIS

7.6.1. Difficulty Index

The proportion of individuals who answered each item on the test correctly was calculated using the responses from the normal group of individuals in the ages between which no statistical differences were found in the accuracy scores i.e. twenty to fifty-nine and sixty to seventy-nine (Table 7.13).

Items in order of difficulty are therefore, twelve, four (two, ten), three, seven, nine, five, one, six,, eleven in the age group twenty to fifty-nine and two, twelve, ten, nine (seven, four), three, one, six, five, eight, eleven in the age group sixty to seventy-nine. A lower difficulty index for each item in the older group was demonstrated. This is to be expected in view of the significant difference in the overall scores between the two groups.

The difficulty index was also calculated for each item for the patient group (Table 7.14). Fisher's Exact Test was used to calculate whether there was a statistical difference between the normal and the patient group in the responses obtained for each item. As can be seen from Table 7.14 items two, five, seven, ten and twelve revealed a difference between the responses of the patients and the normal subjects in the younger age group. But only item two showed a difference between the responses in the older age group. Thus confirming that the test items were not very sensitive to problems experienced by the older C.V.A. patients in the area of part-whole perception

TABLE 7.13 DIFFICULTY INDEX - P-W

ITEM	DIFFICULTY p 20-59 YRS	DIFFICULTY p 60-79 YRS
1	0,79	0,61
2	0,94	0,92
3	0,91	0,68
4	0,95	0.71
6	0,77	0,58
7	0,89	0,71
8	0,67	0,50
9	0,82	0,76
10	0,94	0,82
11	0,52	0,37
12	0,98	0,91

7.6.2. Item Discrimination

This is the extent to which the item discriminates between those who do well on each test item and those who do not. Two ways of determining the discriminative ability of each item were used.

7.6.2.1. Testing the Significance of the Difference Between Two Proportions

The responses obtained from those normal individuals with overall scores in the upper 27% were compared to those with scores in the lower 27%. Table 7.15 gives the results of Fisher's Exact test in making this comparison.

TABLE 7.14 DIFFICULTY INDEX - PATIENT GROUP - P-W

ITEM	DIFFICULTY p 20-59 YRS	DIFFICULTY p 60-79 YRS
1	0,65	0,77
2	0,35 ***	0,54 **
3	0,82	0,69
4	1,0	0,85
5	0,47 *	0,39
6	0,71	0,62
7	0,65 *	0,46
8	0,65	0,62
9	0,82	0,62
10	0,77 *	0,69
11	0,47	0,23
12	0,82 *	0,77

*P < 0,05 ** P < 0,01 *** P < 0,001

All items that showed a significant difference at the 5% level of significance and greater could be retained in the final version of the test. It can be seen from Table 7.15 that item four and twelve could be dropped from the test as they do not discriminate between those normal individuals who can do the test and those who cannot in the age group sixty to seventy-nine. No calculations could be carried out on items two, four and twelve in the younger age group and item two in the older age group as the difficulty index was high. It is interesting that, although items two and twelve do not discriminate between the performance levels of normal subjects, they do discriminate between patient and normal performance.

7.6.2.2. Correlational Approach

This approach of analysing the discriminative ability demonstrated the relationship of the individual item response to the total test score where the sample is restricted to the higher and lower 27% using Flanagan's method (Table 7.16).

TABLE 7.15 ITEM DISCRIMINATION - P-W

ITEM	20-59 YRS	60-79 YRS
1	0,0001 ***	0,0055 **
2	-	-
3	0,0478 *	0,0055 **
4	-	0,1409
5	0,0034 ***	0,0055 **
6	0,0002 ***	0,0055 **
7	0,0478 *	0,1698
8	0,0009 ***	0,0001 ***
9	,0013 **	0,0031 **
10	0,1069	0,0325 *
11	0,0000 ***	0,0573
12	-	0,2105

* P < 0,05 ** P < 0,01 *** P <0,001

It can be seen that only in one instance was there a positive correlation between the upper and lower figures, all other correlations being negative. Thus as the upper value increases, so the lower value decreases. This indicated substantial discrimination, but in the wrong direction and suggested a scoring or a criterion problem. In view of the method of scoring which gave credit for a correct response, a method which is not different from other tests in the battery, it appeared that Flanagan's method is not a useful tool for this test.

TABLE 7.16 BI-SERIAL CO-EFFICIENT OF CORRELATION BETWEEN ITEM AND TOTAL SCORE - P-W

	20-59 YRS			60-79 YRS		
ITEM	UPPER 27%	LOWER 27%	r VALUE	UPPER 27%	LOWER 27%	r VALUE
1	0,00	0,57	-0,70	0,10	0,78	-0,70
2	-	-	-	-	-	-
3	0,00	0,24	-0,50	0,10	0,78	-0,70
4	-	-	-	0,10	0,44	-0,40
5	0,00	0,38	-0,60	0,10	0,78	-0,70
6	0,00	0,52	-0,70	0,10	0,78	-0,70
7	0,00	0,24	-0,50	0,20	0,56	-0,40
8	0,09	0,62	0,40	0,00	0,89	-0,90
9	0,00	0,43	-0,60	0,00	0,67	-0,80
10	0,00	0,19	-0,50	0,00	0,44	-0,70
11	0,05	0,76	-0,30	0,40	0,89	-0,90
12	-	-	-	0,00	0,22	-0,60

7.7. TEST RELIABILITY

7.7.1. Split-Half Reliability

The Split-half reliability determines the internal consistency of a test. In this instance the co-efficient of correlation was r = 0,54 for the age group twenty to fifty-nine years and r = 0,68 for the age group sixty to seventy-nine years with a mean of r = 0,61.

Although both values were significant (P<0,001) the internal consistency of the test was only moderate for both age groups (Table 7.17) when the r value was considered.

7.7.2. Kuder-Richardson Formula 20

This formula determines the homogeneity of the test items or the inter-item consistency. By using this formula the resulting co-efficient of internal consistency was found to be r = 0,57 for the age group twenty to fifty-nine years and r = 0,81 for the age group sixty to seventy-nine years with a mean of r = 0,69. It is clear that there was a high degree of correlation between items and the total test score in the older age group a correlation which was only moderate in the younger age group. (See Table 7.17). However, both correlation values were most certainly significant at the P<0,001 level of significance.

TABLE 7.17 RELIABILITY CO-EFFICIENTS - P-W

AGE	SPLIT-HALF	KUDER-RICHARDSON
20-59	0,54 ***	0,57 ***
60-79	0,68 ***	0,81 ***

*** P < 0,001

7.7.3. Inter-Rater Reliability

Friedman's two-way analysis of variance showed that there were no statistical differences between the accuracy (P=0,8825) or time (P=0,8465) scores obtained by the three examiners on each of the six individuals tested. Therefore, the test was found to have inter-rater reliability.

7.8. TEST VALIDITY

Validity refers to the ability of the test to measure the perceptual skill that it was designed to measure. Two types of validity were examined i.e. content and construct validity.

7.8.1. Content Validity

This was provided by the fifteen occupational therapists who rated the test on a 0-5 scale. Content validity was demonstrated as the therapists rated the test as a very much better than average method of evaluating part-whole ability (Table 7.18).

TABLE 7.18 CONTENT VALIDITY - P-W

PERCEPTUAL SKILL	MEAN RATING
Form (form as mass)	1,5
Overall Contour	1,7
Sub-units of Form	4,0 ***
Perceptual Constancy	1,8
Figure-Ground	1,0
Directionality	0,7
Space Relations	1,6
Position in Space	1,4
Colour	0,2
Praxis	not used

*** a very much better than average assessment

7.8.2. Construct Validity

Three types of construct validity were determined.

7.8.2.1. Age Differentiation

The Part-Whole test was shown to differentiate among the different age groups and a "drop off" in scores occurred at the sixty year old level, a statistical difference in the scores being found between the twenty to fifty-nine year olds and the sixty to seventy-nine year olds. This test would appear, therefore, to be a valid method of determining age related changes in individuals.

7.8.2.2. Correlation with Other Tests in the Battery

As can be seen from Table 7.19 the highest correlation of the Part-Whole test was with the following tests:- Position in Space test (r =0,60), the Design Construction test (r = 0,57), the Dot-Line test (r = 0,53), the VMI test (r = 0,51) and the Figure-Ground test (r = 0,50). All of the tests, except for the Figure-Ground test, have a space component and according to the factor analysis are part of Factor I (Table 7.20). It is apparent that the mental rotation of the forms needed to enable the individual to choose the correct answer in the Part-Whole test could be the source of the relationship between this test and the Position in Space test.

TABLE 7.19 CORRELATION WITH OTHER TESTS IN THE BATTERY - P-W

TEST	CORRELATION WITH PW
VMI	0,51
Perceptual Constancy	0,48
Figure-Ground	0,50
Arrows	0,25
Dot Line	0,53
Space relations	0,27
Position in Space	0,60
Block Construction I	0,40
Block Construction II	0,42
Stick Construction I	0,47
Stick Construction II	0,36
Design Copying	0,57
Colour	0,16

7.8.2.3. Results of Testing Patients

The Part-Whole test has not proved to be an effective measure of defective perceptual function in the older group of patients. This was contrary to the finding in the younger group where a highly significant difference was found between the performance of the normal and the patient groups (P < 0,001). Therefore, the test could be said to be a valid means of determining dysfunction in the younger patient i.e. twenty to fifty-nine years only More extensive testing of older C.V.A. patients needs to take place before it can be said that this test is suitable for determining dysfunction in that age category.

<u>**TABLE 7.20 SORTED ROTATED FACTOR LOADINGS - P-W**</u>

TEST	FACTOR I
Position in Space	0,80
Part-Whole	0,73
Block Construction I	0,74
[illegible]	0,63
[illegible]-Line	0,61
Figure-Ground	0,58
Design Copying	0,51

7.9. CONCLUSIONS

The Part-Whole test revealed some characteristics which would suggest that it should be retained in the V-P-M Battery. These characteristics were as follows:-

- it was easy to administer and to score;
- it demonstrated differences in performance between the older and the younger normal person;
- some of the items were easy and some were difficult, it is therefore possible to grade the items in the test;
- it has good inter-rater reliability;

- it has been shown to be a valid measure of part-whole abilities rated by the occupational therapists;

- there are very few other tests available for this type of function;

However, less than 50% of the items revealed a difference in performance between the normal group and the patients in spite of the fact that the overall score did discriminate between the two groups at the younger age. The correlational approach of testing the ability of each item to discriminate between those who could do the test and those who could not, demonstrated negative correlation values, the only test in the battery to do this. However, when Fisher's Exact Test was used to determine the discriminatory level of each item, most items demonstrated significant differences between the upper and lower scores. Fisher's Exact procedure is the more robust statistical technique and the results, when using this technique, indicated that most items are discriminatory between those who passed and those who failed the test. Most importantly, the test did not differentiate between patient performance and normal group performance at the sixty to seventy-nine year age level, but more patients need to be tested in this age band before the test's usefulness for the older person can be decided.

The factor analysis seemed to indicate that this test was more a test of spatial function than it was of part-whole ability. Therefore, it is recommended that the instructions be changed to eliminate the use of spatial knowledge.

The test could also be used in its present form as a screening tool as all patients who performed below the cut off point for adequate function on this test performed extremely badly or were untestable on the Position in Space test. However, as a test of part-whole ability is needed, the instructions should be changed to eliminate the need for spatial ability. See Chapter Twenty-One for the comparison of this test with the others in the battery.

CHAPTER 8

THE PERCEPTUAL CONSTANCY TEST

8.1. INTRODUCTION

Perceptual constancy has been defined as the ability to perceive an object which possesses invariant properties such as shape, position, size and colour in spite of the variability of the presentation. This skill has also been referred to as "form constancy". Sieve and Freishtat (1976:43) defined "form constancy" as *"the (in)ability to attend to subtle variations in form to the extent that the patient may confuse his water pitcher and his urinal since they have similar shapes".* Hochberg (1978) described the perceptual constancies as being brightness constancy and memory for colour and size. He also stated that for every example of constancy there was an illusion or perceptual error such as the Mueller-Leyer illusion. Benton (1979) named altered thresholds for the discrimination of size, brightness or length as features of visual discriminative ability. This inferred that aspects of perceptual constancy could also be termed visual discrimination. Visual discriminative ability was obviously necessary to perceive size, brightness, etc. Benton (1979) stated that impairment of visual discrimination is not a particularly frequent occurrence in patients with brain disease. For a description of the Perceptual Constancy test see Chapter Four (4.6.4.) and Appendix 1, Section 2.4.

8.2. REVIEW OF SIMILAR TESTS

Perceptual constancy is discussed by workers dealing with children, such as Frostig (1966) and Hammill, et al (1993), it does not feature extensively in the literature dealing with adult perceptual problems. Consequently tests for this perceptual modality are not generally available. The Constancy of Shape test of the Developmental Test of Visual Perception (Frostig 1966) which is standardised for children from four to nine years, has been used in adult studies (Taylor 1968, Gane 1976, Harlowe and van Deusen 1984) but no information exists on its usefulness for the detection of perceptual problems in the adult and no adult norms exist.

The Rivermead Perceptual Assessment Battery (RPAB) has a Size Recognition test which allows the testing *"of constancy through the ability to recognise and match objects in 2D, despite dissimilarity in size"* (Whiting et al 1985:29). As already discussed, the RPAB tests showed no age differentials, the normative group scoring a mean of four points (which is the maximum score for the test) for the size recognition and 3,49 for the Series tests. The tests, however, do appear to discriminate between normal individuals and patients at the 5% level of significance.

The Silhouettes test of the Visual Object and Space Perception Battery (VOSP) (Warrington and James 1991) uses common objects photographed from an unusual view and extended by using 3D shadow images. Warrington and James (1991) stated that this was a test of object perception but the candidate felt that the test required the skill of perceptual constancy for the interpretation of the objects.

The Lowenstein Occupational Therapy Cognitive Assessment (LOTCA) contains an Object Constancy test (Itzkovich et al 1990) similar to that of the Silhouettes test of the VOSP. The subject is also required to identify objects from photographs taken at unusual angles. Although showing no age related changes, the test does discriminate between normal individuals and patients ($P < 0,001$) thus making it a more sensitive test than the RPAB tests.

8.3. REASONS FOR COMPILING A NEW TEST

As can be seen in the review of tests evaluating perceptual constancy, no test existed for the evaluation of this skill at the time that this study was begun. Tests for children such as the Perceptual Constancy component of the Developmental Test of Visual Perception (Frostig 1966) were considered too easy for the evaluation of adult function. The RPAB evaluated size constancy only and was, therefore, not as comprehensive as the Form Constancy test in the V-P-M Battery.

The Perceptual Constancy test requires the subject to locate objects, shapes and letters similar to those presented on a demonstration card, on the test sheet.

8.4. RESULTS OF TESTING NORMAL INDIVIDUALS

8.4.1. Accuracy Scores

The means and the standard deviations were calculated for the normal group as a whole (Table 8.1) and for each group individually (Table 8.2).

TABLE 8.1 MEAN AND SD OF ACCURACY SCORES - PC

n	MAX SCORE	MEAN	SD	RANGE
118	70	57,88	4,92	38-66

TABLE 8.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - PC

AGE	n	MEAN	SD	RANGE
20-29	20	60,5	3,07	55-66
30-39	21	59,48	3,88	51-66
40-49	19	59,47	4,38	47-64
50-59	19	57,47	3,39	50-63
60-69	19	55,95	5,91	38-64
70-79	20	54,3	5,62	42-64

The Pearson Chi-square Test was used to determine whether any difference existed within the group as a whole. A cut point score of fifty-eight was used with the percentage obtaining a score greater than and lesser than the cut point being used to calculate the Chi-square value (see Table 8.3).

TABLE 8.3 COMPARISON OF ACCURACY SCORES - PC

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 58	20,0	33,3	31,6	63,2	63,2	80,0	48,3
Greater than 58	80,0	66,7	68,4	36,8	36,8	20,0	51,7

Pearson Chi-square Value = 21,833 = P<0,001

As a highly significant difference was found to be present, further analysis was carried out to determine where the differences lay. The results of this analysis are as follows:-

a) the Pearson Chi-square Test revealed that there were no differences in the accuracy scores of the age groups twenty to twenty-nine, thirty to thirty-nine and forty to forty-nine years, (P=0,5942). These groups were, therefore, combined for further analysis.

b) according to the Pearson Chi-square Test no differences existed in the accuracy scores of age groups fifty to fifty-nine, sixty to sixty-nine and seventy to seventy-nine years, (P=0,4197). These groups were, therefore, combined for further analysis.

c) Yates' Corrected Chi-square test revealed a highly significant difference between the performance of the combined age groups i.e. twenty to forty-nine year olds and fifty to seventy-nine year olds at the P <0,001 level of significance (P=0,0000) (see Table 8.4).

TABLE 8.4 COMPARISON OF ACCURACY SCORES FOR AGE GROUPS 20-49 AND 50-79 YEARS - PC

AGE	20-49	50-79	TOTAL
Less than 58	28,3	69,0	48,3
Greater than 58	71,7	31,0	51,7

Yates' Corrected Chi-square Value 17,905 = P <0,001

A drop off in the ability to execute the items on this test appeared at an earlier age than in the other tests, the drop off occurring at approximately fifty years of age. Thus the null hypothesis that decreased ability in the V-P-M function will not be found with advancing age, was rejected at the 0,001% level of significance.

8.4.2. Time Scores

The means and the standard deviations were calculated for the time scores for the group as a whole (Table 8.5) and for each age group (Table 8.6).

TABLE 8.5 MEAN AND SD FOR TIME SCORES - PC

n	MEAN/SECS	SD	RANGE
118	379,80	143,40	142-1115

TABLE 8.6 MEAN AND SD FOR TIME SCORES FOR EACH AGE GROUP - PC

AGE	n	MEAN/SECS	SD	RANGE
20-29	20	277,40	80,81	142-462
30-39	21	266,71	49,42	176-349
40-49	19	320,71	78,40	196-484
50-59	19	315,95	74,93	207-478
60-69	19	388,79	108,85	199-560
70-79	20	532,45	208,25	274-1115

Levine's test for equal variance established that there was a significant difference among the groups (P = 0,0000) and the Welch Test that there was a significant difference between the means obtained by the different age groups (P = 0,0000). The separate variance t was, therefore, used with the Bonferroni correction for multiple comparisons to establish where the differences were to be found (Table 8.7).

On the basis of the differences found among the various age groups it was possible to regroup the ages as previously described. A similar analysis was carried out on the re-grouped data (Table 8.8).

As can be seen from Table 8.8 the twenty to thirty-nine year old group differed from all other groups in the time taken to complete the test. The forty to fifty-nine year olds differed from the seventy to seventy-nine year olds but not from the sixty to sixty-nine year olds. The null hypothesis of no increase in speed with advancing age can therefore be rejected.

TABLE 8.7 DIFFERENCES IN TIME SCORES - PC

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,6152
20-29	40-49	0,0996
20-29	50-59	0,1307
20-29	60-69	0,0010 *
20-29	70-79	0,0000 ***
30-39	40-49	0,0157
30-39	50-59	0,0213
30-39	60-69	0,0001 **
30-39	70-79	0,0000 ***
40-49	50-59	0,8567
40-49	60-69	0,0335
40-49	70-79	0,0003 **
50-59	60-69	0,0223
50-59	70-79	0,0002 **
60-69	70-79	0,0110

* $P<0{,}05$ ** $P<0{,}01$ *** $P<0{,}0001$

TABLE 8.8 DIFFERENCES IN RE-GROUPED TIME SCORES - PC

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-59	0,0051 **
20-39	60-69	0,0002 **
20-39	70-79	0,0000 ***
40-59	60-69	0,0173
40-59	70-79	0,0002 **
60-69	70-79	0,0110

* $P<0{,}05$ ** $P<0{,}01$ ** $P<0{,}001$

Although no differences existed between the forty to fifty-nine year olds and the sixty to sixty-nine year olds as well as the sixty to sixty-nine year olds and the seventy to seventy-nine year olds, the scores were hypothesis of no increase in speed with advancing age can therefore be rejected.

TABLE 8.7 DIFFERENCES IN TIME SCORES - PC

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,6152
20-29	40-49	0,0996
20-29	50-59	0,1307
20-29	60-69	0,0010 *
20-29	70-79	0,0000 ***
30-39	40-49	0,0157
30-39	50-59	0,0213
30-39	60-69	0,0001 **
30-39	70-79	0,0000 ***
40-49	50-59	0,8567
40-49	60-69	0,0335
40-49	70-79	0,0003 **
50-59	60-69	0,0223
50-59	70-79	0,0002 **
60-69	70-79	0,0110

* P<0,05 ** P<0,01 *** P<0,0001

TABLE 8.8 DIFFERENCES IN RE-GROUPED TIME SCORES - PC

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-59	0,0051 **
20-39	60-69	0,0002 **
20-39	70-79	0,0000 ***
40-59	60-69	0,0173
40-59	70-79	0,0002 **
60-69	70-79	0,0110

* P<0,05 ** P<0,01 *** P<0,001

Although no differences existed between the forty to fifty-nine year olds and the sixty to sixty-nine year olds as well as the sixty to sixty-nine year olds and the seventy to seventy-nine year olds, the scores were kept separate for the computation of percentiles (see Appendix 2) to conform with the pattern in the other V-P-M tests.

8.4.3. Discussion Of The Results Of Testing Normal Individuals

A review of the literature revealed that perceptual constancy is a skill not often studied. There are few tests of perceptual constancy and this might account for the paucity of work on age related changes in this particular perceptual skill. The only test that made a distinction between the performance of younger subjects (twenty to forty-eight years) and older subjects (fifty to sixty-nine years) was the VOSP although Warrington and James (1991) stated that there was only a minor decrease in performance.

The bulk of the work in the field of perceptual constancy has been carried out on the ability to identify visual illusions such as the Mueller-Leyer figures. Although visual illusions are examples of defects in perceptual constancy (Hochberg 1978) they are not comparable to the perceptual constancy tasks in the Visual-Perceptual-Motor Battery described in this thesis. They are examples of misapplied rules for perceptual constancy. *"For every example of constancy ... we can find an illusion or perceptual error"* (Hochberg 1978:79).

8.4.3.1. Studies on Age Related Changes in the Ability to Perceive Visual Illusions

Studies on the effect of ageing on the interpretation of illusions have given some insight into the decrease of perceptual constancy function with age (Comalli 1965, Pollack and Zetland 1965, Eisner and Schaie 1971, Atkeson 1978, Lorden et al 1979). However, according to Birren and Schaie (1977) there is no satisfactory theory of illusions which permits a systematic classification of these faults of interpretation either in terms of their stimulus properties or the age related differences in response to them. It appears that susceptibility to some illusions increases with age, e.g. the Mueller-Leyer illusion (classified as type I) (Wapner et al

1960, Lorden et al 1979). The susceptibility to the illusions classified as type II e.g. Usnadse Effect, has been found by some workers to remain constant (Lorden et al 1979). According to Birren and Schaie (1977) there is no agreement on the effects of ageing on the ability to interpret the type II illusions.

8.4.3.2. Age Related Changes in Occupational Therapy Tests

Taylor (1968) using an adaptation of the Form Constancy test of the Developmental Test of Visual Perception (Frostig 1966) did not find a noticeable change in accuracy scores with age although she did find slowed reaction times in the older person as was also found by the candidate on the V-P-M Battery.

The sections of the RPAB (Whiting et al 1985) dealing with perceptual constancy i.e. the Size Recognition test and the Series test showed no age related changes. It seems, therefore, that the Perceptual Constancy test of the Visual-Perceptual-Motor Battery described in this thesis provides a way of evaluating age related changes in Perceptual Constancy.

8.5. RESULTS OF TESTING PATIENTS

8.5.1. Accuracy Scores

The means and the standard deviations were calculated for the age groups twenty to forty-nine and fifty to seventy-nine as these were the age groups within which no statistical differences occurred in the scores obtained by the normal group (Table 8.9). Not all of the thirty three patients making up the sample were evaluated on this test for reasons enumerated in section 8.5.4.

TABLE 8.9 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - PC

AGE	n	MEAN	SD	RANGE
20-49	11	55,73	5,0	44-62
50-79	16	47,67	7,25	35-58

8.5.2. Time Scores

The means and the standard deviations were calculated for the time scores in the age groups within which no statistical difference was found in the normal group (Table 8.10).

TABLE 8.10 MEAN AND SD FOR TIME SCORES FOR PATIENTS - PC

AGE	n	MEAN/SECS	SD	RANGE
20-39	7	435,0	173,87	216-604
40-69	19	545,5	265,31	296-1270
70-79	1	922,0	-	-

8.5.3. Comparison Between Patients and Normal Subjects

8.5.3.1. Accuracy Scores

The mean accuracy scores obtained by the patients were compared with those obtained by the normal subjects to determine differences between the two groups.

As Levine's test for variances was not significant (P=0,3570) for the age group twenty to forty-nine years, the F Test was used to test for significant differences between the two groups (Table 8.11). Levine's test was, however, significant for the age group fifty to seventy-nine years (P=0,0365), thus the Welch Test was used to determine the difference between the two groups. This analysis showed a highly significant difference between the performance of the patient group and that of the normal group (Table 8.11).

TABLE 8.11 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - PC

AGE	LEVINE'S TEST P VALUE		P VALUE
20-49	0,357	F Test	0,0022 **
50-79	0,0365	Welch Test	0,0000 ***

* P <0,05 ** P <0,01 *** P <0,001

8.5.3.2. Time Scores

As with the accuracy scores, the means of the time scores for normal individuals and patients were compared to determine whether there were any statistical differences. In the case of the twenty to thirty-nine year olds and the forty to sixty-nine year olds, Levine's test was significant, therefore, the variances could not be assumed to be equal and the Welch Test was used. No statistical differences were found between the time scores for the patients and the normal younger group, however, there was a difference between the normal group and the patients within the fifty to sixty-nine year old age band (Table 8.12). No comparisons could be made with the seventy year olds as only one patient completed the test.

TABLE 8.12 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - PC

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20-39	0,0000 **	0,0518
40-69	0,0000 ***	0,0043 **
70-79	-	-

** P <0,01 *** P < 0,001 - not calculated

8.5.4. Analysis Of Tests Executed By Patients

In analysing the percentile equivalent scores obtained by the patients the following was found:-

a) In the twenty to thirty-nine year old group, composed of seven patients, four obtained scores above the 15,9 percentile, i.e. four patients could be said to have normal function. Only three patients could be said to have problems with perceptual constancy. As in the case of the VMI test, it was interesting to compare the scores obtained by these three patients with those of the normal seventy year olds. Two of the three patients obtained scores above the 30th percentile i.e. within normal limits when making this comparison. The question to be asked, therefore, was whether this dysfunction affected their daily life tasks.

b) Of the ten patients in the forty to fifty-nine year old group, four scored above the 15th percentile, i.e. within normal limits, five at the 11th percentile and lower and one patient was discharged prior to completing the test.

c) Ten patients were tested in the sixty to sixty-nine year old group and of these, six patients scored above the 15th percentile while four scored at the 7th percentile and lower.

d) Of the six patients in the older group i.e. seventy to seventy-nine year olds one scored at the 7th percentile, the testing of four patients was discontinued as they found the test too difficult and one was discharged prior to completing the test. The patients of this age group performed extremely poorly on the test, four out of a possible five not being able to complete the test at all.

Out of the total of twenty-seven patients, thirteen obtained scores below that of the normal group.

e) Seventeen patients obtained scores below that of the normal group on the time scores, although four of these patients scored within the normal range on the accuracy scores, achieving accuracy at the expense of time.

8.5.5. Problems That May Be Diagnosed From Performance On The Perceptual Constancy Test

Problems with the ability to perceive constancy of size may be determined in items nine to fourteen. In this section the majority of individuals scored poorly. The first section of the test in which constancy of objects and animals was evaluated (items one to eight) could also be used to determine the presence of an object agnosia. However, none of the patients evaluated in this thesis could be considered to have this type of problem.

Other aspects of performance such as left neglect, visual inattention and perseveration can also be diagnosed, the candidate felt however that other tests in the V-P-M Battery are more suitable for the demonstration of these problems (see Chapter Twenty-One).

8.6. ITEM ANALYSIS

8.6.1. Difficulty Index

"The difficulty index of an item is the proportion of individuals who answered the item correctly" (Downie and Heath, 1974). The difficulty index was calculated using the responses from the normal group of individuals in the age groups within which no statistical differences were found in the accuracy scores i.e. twenty to forty-nine year olds and fifty to seventy-nine year olds (Table 8.13). Some variation occurred within the age groups, but as can be seen from Table 8.13 the shape constancy items were the most difficult. The most logical way to regroup the items would be to retain the present groups and re-order the items within each group. The shape constancy items should be presented last. The suggested order is as follows:-Item (three, six, eight, four, five, two, one, seven) (eighteen, sixteen, twenty,

seventeen, nineteen, fifteen) (eleven, thirteen, ten, nine, twelve, fourteen).

TABLE 8.13 DIFFICULTY INDEX - PC

ITEM DESCRIPTION	ITEM .	DIFFICULTY p 20-49 YRS	DIFFICULTY p 50-79 YRS
Horse	1	0,57	0,52
Dogs	2	0,85	0,70
Carrots	3	0,90	0,93
Giraffe	4	0,93	0,91
Couch	5	0,92	0,96
Comb	6	0,97	0,93
Ducks	7	0,54	0,46
Frogs	8	0,98	0,89
Triangle	9	0,21	0,11
Rectangle	10	0,23	0,23
Circle	11	0,84	0,63
Diamond	12	0,07	0,02
Oval	13	0,48	0,11
Square	14	0,44	0,02
B	15	0,59	0,38
P	16	0,97	0,88
F	17	0,92	0,82
M	18	0,92	0,96
J	19	0,72	0,68
U	20	0,95	0,89

TABLE 8.14 DIFFICULTY INDEX - PATIENT GROUP - PC

ITEM	DIFFICULTY p 20-49 YRS	DIFFICULTY p 50-79 YRS
1	,73	,40
2	,91	,40
3	,91	,87
4	1,0	,60 **
5	,91	,80
6	,91	,73 **
7	,55	,20 *
8	,91	,53 **
9	,27	,13
10	,18	,27
11	,55 **	,40
12	,27 **	,07 *
13	,09 **	,07
14	,09 *	,07
15	,55	,20
16	,64 ***	,60 **
17	,82	,40 *
18	1,0	,67 ***
19	,64	,33 **
20	,82	,67 *

* P <0,05 ** P < 0,01 *** P <0,001

The difficulty index for the responses obtained from the patient group was also calculated (Table 8.14). Fisher's Exact test was used to calculate whether there were statistical differences between the normal and patient group in the responses obtained for each item. This was done to help to determine the final test form. It can be seen from Table 8.14 that only items four, six, seven, eight, eleven, twelve, thirteen, fourteen, sixteen, seventeen, eighteen, nineteen, and twenty discriminate between normal performance and patient performance.

The difficulty levels for the shape constancy part of the Perceptual Constancy test indicated that, for the patients this part of the test was also the more difficult. In the twenty to forty-nine year age group four of the six items demonstrated a difference between patients and the normal group. But in the older age group there was a difference on only one item thus indicating that the decrease in function in the patients was more a factor of age than of dysfunction due to C.V.A. The numbers and objects sections of the test seemed to be more successful in discriminating between patients and normal subjects in the older age group.

Although the difficulty index was slightly different for the patient group, the order of items suggested from studying the responses of the normal group would seem to be suitable.

8.6.2. Item Discrimination

Item discrimination is the extent to which the item discriminates between those who can do well on each test item and those who can not. Two ways of determining item discrimination were used.

8.6.2.1. Testing the Significance of the Differences between Two Proportions

The responses obtained by the normal individuals whose overall score was in the upper 27% were compared with those in the lower 27%. Table 8.15 gives the result of Fisher's Exact test in making this comparison.

A cut point was established for each item so that the results could be calculated. This was necessary as each item had a score of more than one point. The cut point was established by examining the range of scores for each item and the number of individuals obtaining that score.

All items that showed a significant difference between the upper and lower 27% should be retained as they discriminated between those who could do the test and those who could not do the test. Only items two, seven, eight, nine, ten, eleven, thirteen, fourteen, and fifteen discriminated in this way, of these, only items nine, eleven and fifteen were discriminatory for all age groups.

TABLE 8.15 ITEM DISCRIMINATION - PC

ITEM	29-49 YRS	50-79 YRS	MAX SCORE	20-49 YRS	50-79 YRS
	Cut point	Cut point			
1	2	2	3	0,0656	0,0756
2	3	3	4	0,0169	0,0059 *
3	1	1	2	1,00	1,00
4	2	2	3	1,00	0,2258
5	1	1	2	0,0996	1,00
6	1	1	2	-	0,1012
7	2	2	3	0,4661	0,0038 **
8	2	2	3	-	0,0177 *
9	1	0	3	0,0209 *	0,0011 **
10	1	1	3	0,1086	0,0091 **
11	2	2	3	0,0421 *	0,0000 ***
12	2	1	4	0,1281	1,000
13	1	1	3	0,0005 **	0,2852
14	0	0	2	0,0092 **	-
15	3	3	4	0,0007 ***	0,0091 **
16	3	3	4	1,00	0,0829
17	4	4	5	0,0996	0,0829
18	4	3	5	0,3295	0,4839
19	5	5	6	0,2451	0,0155
20	5	5	6	0,4828	0,0434

* P <0,05 ** P <0,01 *** P <0,001

Individual items on the test, therefore, did not posses a high level of discrimination when testing the difference between two proportions.

8.6.2.2. *Correlational Approach*

This approach showed the relationship of the individual item responses to the total test score (see Table 8.16). Using the correlational approach, it appeared that only item three should be dropped from the final test form if it were to be used for all age groups. However, Flanagan's method was the less accurate method of evaluating those items which discriminated between those who could do well in the test and those who could not do well as the r value is estimated.

In view of the studies on item difficulty and item discrimination it seemed that items one and three should be dropped from the test, as they did not discriminate between normal and patient's performance and they did not discriminate within the normal group between those who could do the test and those who could not.

TABLE 8.16 BI-SERIAL CO-EFFICIENT OF CORRELATION BETWEEN ITEM AND TOTAL SCORE PC

	20 - 49 YRS			50 - 79 YRS		
ITEM	UPPER 27%	LOWER 27%	r VALUE	UPPER 27%	LOWER 27%	r VALUE
1	0,67	0,27	0,40***	0,69	0,31	0,40*
2	1,00	0,60	0,65***	0,94	0,44	0,60***
3	0,93	0,87	0,16	1,00	0,94	0,30
4	0,93	0,87	0,16	1,00	0,82	0,52***
5	1,00	0,73	0,61***	1,00	0,94	0,31
6	-		-	1,00	0,75	0,61***
7	0,60	0,40	0,20	0,75	0,19	0,56***
8	-	-	-	1,00	0,63	0,64***
9	0,87	0,40	0,52***	0,81	0,19	0,61***
10	0,87	0,53	0,44***	0,88	0,38	0,57***
11	1,00	0,67	0,63***	0,94	0,13	0,80***
12	0,53	0,20	0,35**	0,38	0,31	0,08
13	0,93	0,27	0,68***	0,56	0,31	0,27
14	0,73	0,20	0,53***	-	-	-
15	0,87	0,20	0,68***	0,63	0,13	0,54***
16	0,93	0,92	0,00	0,94	0,63	0,45**
17	1,00	0,73	0,61***	0,94	0,63	0,45**
18	0,93	0,73	0,33**	1,00	0,88	0,51**
19	0,80	0,53	0,32*	0,94	0,50	0,59***
20	1,00	0,87	0,44***	1,00	0,69	0,63***

* P <0,05 ** P <0,01 *** P <0,001

8.7. TEST RELIABILITY

8.7.1. Split-Half Reliability

This determines the internal consistency of a test with regard to content sampling. The resulting co-efficient of correlation was r = 0,56 for age groups twenty to forty-nine years and r = 0,72 for age groups fifty to seventy-nine years with a mean of 0,64. It seemed, therefore, that the reliability of the Perceptual Constancy test was low when the results of the younger age group were considered. The fact that one section of the test was more difficult than the other two sections might contribute to this as well as the low degree of variance among the items. In spite of this there was a significant correlation between the two halves of the test (Table 8.17).

8.7.2. Kuder-Richardson Formula 20

In this method the pass-fail score for each item was compared with the total score. This procedure determined the homogeneity of the test items or the inter-item consistency of the test.

Using this formula, the resulting co-efficient of internal constancy was r = 0,86 for age groups twenty to forty-nine and r = 0,95 for age groups fifty to seventy-nine years with a mean of 0,91. This indicated a high degree of correlation with the items and the total test score.

Thus the co-efficient of correlation ranged between r = 0,56 and r = 0,95. All co-efficients were found to be statistically significant at the P <0,001 level of significance. Co-efficients of r = 0,80 or above are traditionally accepted as evidence of adequate test reliability. Using this criterion, the Perceptual Constancy test showed a high degree of reliability when using the Kuder-Richardson Formula but not when the split-half reliability was used (see Table 8.17). However, many tests have been published and are used with reliabilities below r = 0,80 e.g. the Bender Visual Motor Gestalt Test (Lezak 1983). It is hoped that removing some items will improve the reliability of the test.

TABLE 8.17 RELIABILITY CO-EFFICIENTS - PC

AGE	SPLIT-HALF	KUDER-RICHARDSON
20-49	0,56 ***	0,86 ***
50-79	0,72 ***	0,95 ***

*** P <0,001

8.7.3. Inter-Rater Reliability

Friedman's two-way analysis of variance showed that there was no statistical difference between either the accuracy (P = 0,8825) or time (P = 0,3114) scores obtained by the three examiners on each of six individuals tested.

8.8. TEST VALIDITY

To determine the validity of the Perceptual Constancy test, two questions should be answered. Did the test measure perceptual constancy and how well did it measure this skill? To answer these questions two types of validity were discussed i.e. content and construct validity.

8.8.1. Content Validity

This was provided by the fifteen occupational therapists (population 1) who rated the test on a 0-5 scale. The Perceptual Constancy test (see Table 8.18) was shown to be a very much better than average method of evaluating this perceptual skill. However, according to the therapists it was also an average method of evaluating figure-ground perception.

The content evaluation demonstrated that the Perceptual Constancy test was a valid measure of this skill.

8.8.2. Construct Validity

Three aspects of construct validity were considered, i.e. age differentiation, correlation with others tests in the battery and the results of testing patients.

8.8.2.1. Age Differentiation

Discussion already presented indicated that some perceptual skills show a progressive decrease with age. In the case of the Perceptual Constancy test of the Visual-Perceptual-Motor battery, the "drop off" in accuracy occurred at about fifty years of age, a statistical difference being found in the accuracy scores of the twenty to forty-nine year olds when compared to those of the fifty to seventy-nine year olds. The Perceptual Constancy test would, therefore, seem to be a valid method of differentiating age related changes in the performance of this type of task.

TABLE 8.18 CONTENT VALIDITY - PC

PERCEPTUAL SKILL	MEAN RATING
Form as mass	2,2
Overall contour	1,6
Sub-Units of Form	1,1
Perceptual Constancy	3,9 ***
Figure-Ground	2,5 **
Directionality	0,6
Space Relations	0,4
Position in Space	0,7
Colour	0,0
Praxis	not used

** better than average assessment. *** a very much better than average assessment.

8.8.2.2. Correlation with other tests in the Battery

The highest correlation of the Perceptual Constancy test with other tests in the battery was with the Design Construction (r = 0,56) and the Figure-Ground tests (r = 0,53). These correlations might be on the basis that the Design Construction test required some ability to distinguish between sizes of shapes, therefore, tapping size constancy.

The therapists rated the Perceptual Constancy test as an average test of the assessment of figure-ground perception and, therefore, it was apparent that a degree of figure-ground ability was necessary to execute the Perceptual Constancy test.

A factor analysis was carried out on the tests in the battery to determine whether the tests had factors related to one another. The tests that appeared to measure the same factor (Factor II) as the Perceptual Constancy test were:- the Dot-Line, Figure Ground, Space Relations, Arrows, Stick Construction II and Design Construction tests. (See Table 8.19).

It seemed that the common skill in Factor II was one that required the individual to pick out a form from its background thus confirming the content identified by the therapists. However, as no marker variables were included, the results could not be readily used for the purpose of construct validity.

TABLE 8.19 SORTED ROTATED FACTOR LOADINGS - PC

TEST	FACTOR II
Dot-Line Test	0,52
Figure-Ground	0,54
Space Relations	0,79
Arrows	0,75
Stick Concentration II	0,60
Designs Construction	0,57
Perceptual Constancy	0,52

8.8.2.3. Results of Testing Patients

The test was shown to be an effective measure of testing perceptual constancy in patients as there was a significant difference between the scores obtained by the normal group and the patient group. (Twenty to forty-nine years P <0,01, fifty to seventy-nine years P <0,001).

From the results it appeared that the Perceptual Constancy test was a valid means of differentiating age related changes as well as problems in patients who had suffered a cerebro-vascular accident. It did not correlate highly with other tests in the battery and would therefore seem to have a unique place within the test battery although it appeared to contain a factor related to figure-ground and visual discriminative function.

8.9. CONCLUSIONS

The Perceptual Constancy test should be retained within the test battery as there are very few tests of this perceptual modality. In summary this test:-

- demonstrated differences between the younger and older groups;
- had both easy and difficult items, therefore grading of items is possible;
- had good inter-rater reliability;
- demonstrated a difference in performance between patients with a right C.V.A. and normal subjects although individual items on the test did not always discriminate between these two groups. Removing all items that do not discriminate between the normal subjects and the patients would leave the test with all the items with the exception of items one, two and three.
- had reliability co-efficients which were adequate (r = 0,56 to r = 0,95);
- was shown to be a valid measure of perceptual constancy.

One disadvantage was that the elderly group of patients performed poorly on the test, in fact, four out of five patients were unable to complete the test. However, if these patients had been previously screened by means of the Form Matching test, testing would probably have demonstrated a clearer picture of actual performance levels.

The results of the psychometric evaluation of all the tests of the V-P-M Battery will be compared in Chapter Twenty-One. Their comparative usefulness and, therefore, the need to retain such tests in the battery will then be discussed.

CHAPTER 9

THE FIGURE-GROUND TEST

9.1. INTRODUCTION

Figure-ground perception has been defined as the ability to select out of the mass of incoming stimuli only those that are meaningful and require attention. The selected stimuli form the figure in the perceptual field while the majority of stimuli form the dimly perceived background. This skill has also been called field dependence/field independence which is also said to be a dimension of cognition that includes not only perceptual ability but personality differences and social function (Witkin et al, 1967).

According to Kolb and Whishaw (1985, 1990) impairment of the ability to separate figure from ground is a common feature of visuo-perceptual defects. Benton (1979) stated that defective figure-ground perception was *"a cardinal feature of higher-level visuo-perceptual defects"* (page 194) needing the skills of analysis and synthesis along with visual discriminative ability.

Luria (1980) saw figure-ground ability as part of the ability to perceive objects and pictures and Lezak (1983) called the skill one of visual recognition.

In spite of the fact that Benton (1979) and Kolb and Whishaw (1985) stated that impairment of figure-ground ability was a common visual-perceptual disorder, the literature on this subject is scant. This might be due to the fact that it has also been classified as part of a spatial relations syndrome (Siev and Freishtat 1976, Siev et al 1986).

For the Figure-Ground test of the V-P-M battery see Chapter Four (4.6.5.) and Appendix 1 Section 2.5.

9.2. REVIEW OF SIMILAR TESTS

Previously published tests for figure-ground ability are of two types i.e. visual masking such as cross natching or overlapping figures and embedded figure tasks. The former being easier than the latter according to Luria (1980).

9.2.1. Visual Masking Tasks

The concept of overlapping figures (Figure 9.1) developed by Poppelreuter in 1917 has been extensively used in other tests, such as in Luria's Neuropsychological Investigation (Christensen 1974). This investigation was standardised for adults by Golden et al (1980). According to Lezak (1983), Ghent employed similar overlapping figures to indicate perceptual function in children and these figures have been used in adult studies by De Renzi and Spinnler (1966) to investigate the functions of the right and left sides of the brain.

Other workers in the field of perception working with children drew up tests which have also been used with adults. The Developmental Test of Visual Perception (Frostig 1966 and Hammill, Pearson and Voress 1993) used the visual masking task of overlapping figures as a test of figure-ground ability. The 1966 test was used as part of a perceptual assessment battery by Harlowe and Van Deusen (1984) and suggested by Siev and Freishtat (1976) and Siev et al (1986) as being suitable for evaluating adult function. No normative data are available for adults on this test.

Ayres (1972 and 1980a) used overlapping figures in the first half of her figure-ground test. This test was used by Kaplan and Hier (1982) and Peterson and Wikoff (1983) to evaluate adult function.

Peterson and Wikoff (1983) tested one hundred normal males between the ages of nineteen and fifty-seven years on the Ayres' test. They used the Spearman-Rho inter-correlation co-efficient to determine relationships between the test results and age, education and socio-economic status, but found no relationship among these variables. They stated that the test was therefore *"... not confounded by other variables that often compromise data"* (pg. 559). In addition, only one subject obtained a full score on this test, demonstrating the absence of a ceiling effect. According to them this made the Ayres' Figure-Ground test a good test for discriminating this skill in adults.

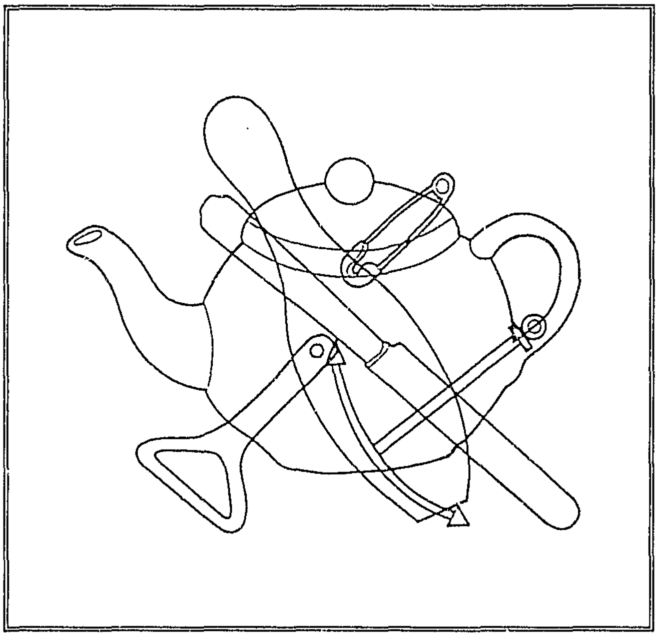


FIGURE 9.1 OVERLAPPING FIGURES (Rivermead Perceptual Assessment Battery Reduced)

Farver and Farver (1982) used the Poppelreuter superimposed figures to determine the ability of normal adults (n = 68) in their test. A statistically significant age effect was noticed ($P < 0.05$).

The RPAB (Whiting et al 1985) also used the concept of the Poppelreuter overlapping figures in the figure-ground component of this test battery (see Figure 9.1). The Rivermead test showed no age differentials (Cockburn et al 1982). The COTNAB (Tyerman et al 1986) offered a more comprehensive test of overlapping figures consisting of five items. Tyerman et al (1986) however, labelled their test a test of form constancy. The COTNAB has normative data for ages sixteen to thirty, thirty-one to forty-nine and fifty to sixty-five.

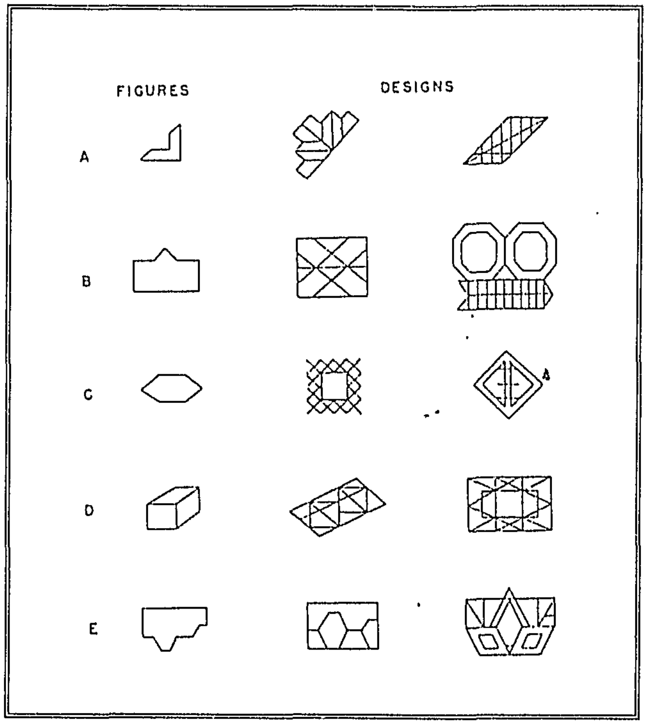


FIGURE 9.2 HIDDEN FIGURES

The LOTCA (Itzkovitch et al 1990) is similar to the Rivermead Test but uses two cards. The subject, however, is required to name the objects seen. This test did not reflect age related changes.

No tests in which cross hatchings were used to mask the figure were found.

9.2.2. Embedded Figure Tests

Gottschaldt was the first to produce a hidden figures or embedded figure test (Fig 9.2) which was used by Witkin (1950) and Talland (1965). According to Peterson and Wikoff (1983) the Embedded Figures Test of Witkin (1950) is the most widely cited test for figure-ground perception. This test, however, is dependent on motor ability (the subject is required to outline the figure) and memory. The subject is shown a complex figure followed by a simple figure. He is then expected to outline the figure as it appears in the complex figure.

Embedded figures have subsequently been used in several studies and in the formation of other tests, such as Luria's neuropsychological battery (Christensen 1975). Ayres (1972 and 1980a) also used embedded figures in the second half of her figure-ground test. Kaplan and Hier (1982) found a significant age related effect on the embedded figure component of the Ayres' test between subjects younger than sixty-eight years and those older than sixty-eight years (P<0,01). They compared the performance of thirty-four patients with right hemisphere lesions with that of sixteen control subjects. Lezak (1983) stated that the items in the Southern California Figure-Ground test (1980a) were suitable for adults as they were *"challenging at all ages"* (pg. 364). Peterson and Wikoff (1983) obtained data on the performance of one hundred males between the ages of nineteen years and fifty-seven years on the Ayers' test. The mean and the standard deviation of the scores were provided.

9.3. Reasons For Compiling A New Test

As can be seen from the review of the tests published for figure-ground ability, the RPAB, the LOTCA and the COTNAB contain a standardised test for the evaluation of figure-ground ability for adults. The RPAB consists of one test plate of overlapping *"outline articles"*, the patient is required to identify articles presented on cards that are present or not present in the test plate. This test was not available at the time that the tests for this thesis were compiled (Figure 9.1) nor was the LOTCA or the COTNAB.

The other tests mentioned were either not available, were felt to be too complicated (the Gottschaldt figures) or did not have normative data.

The candidate decided to compile a test that was more difficult than the overlapping figures tests but easier than the embedded figures test, that is a test with hatching over the figures was constructed as no similar tests seemed to exist.

The Figure-Ground test of the V-P-M Battery consists of a multiple choice section of eight pictures and a hatched over section containing five of the pictures. The subject is required to indicate on the multiple choice plate the figures that are also in the hatched over section.

9.4. RESULTS OF TESTING NORMAL INDIVIDUALS

9.4.1. Accuracy Scores

The means and the standard deviations were calculated for the normal group as a whole (Table 9.1) and for each group individually (Table 9.2).

TABLE 9.1 MEAN AND SD OF ACCURACY SCORES - F-G

n	MAX SCORE	MEAN	SD	RANGE
118	80	77,57	3,25	63-80

TABLE 9.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - F-G

AGE	n	MEAN	SD	RANGE
20-29	20	[illegible],9	1,52	74-80
30-39	21	78,86	1,39	74-80
40-49	19	78,63	1,67	73-80
50-59	19	77,05	4,54	63-80
60-69	19	76,74	3,74	64-80
70-79	20	75,15	3,68	68-80

The Pearson Chi-square Test was used to determine if any differences existed in the scores obtained by the different age groups. A cut point of seventy-eight was used, with the percentage obtaining a score greater than and less than the cut point being used to calculate the Chi-square value (Table 9.3).

TABLE 9.3 COMPARISON OF ACCURACY SCORES - F-G

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 78	25,0	33,0	26,3	47,4	52,6	85,0	44,9
Greater than 78	75,0	66,7	73,7	52,6	47,4	15,0	55,1

Pearson Chi-square Value = 20,493 P = <0,01

As a significant difference (P <0,01) was found to exist in the scores, further analysis took place to determine where the differences were to be found. The results of this analysis are as follows:-

a) The Pearson Chi-square Test showed that there was no statistical difference among the age groups twenty to twenty-nine, thirty to thirty-nine and forty to forty-nine (P = 0,8462). These age groups were thus combined for further analysis

b) Fisher's Exact test showed that there was a difference between the combined scores for age groups twenty to forty-nine and fifty to sixty-nine P = 0.0343 (see Table 9.4). However, this difference is only at the 5% level of significance.

c) Fisher's Exact test showed no difference between the scores obtained for the fifty to fifty-nine year olds and the sixty to sixty-nine year olds (P = 0,5000). These ages were therefore grouped together for further analysis.

d) Fisher's Exact test showed a statistical difference between the new group of fifty to sixty-nine year olds and the seventy to seventy-nine year olds, (P <0,01) (Table 9.5).

TABLE 9.4 COMPARISON OF ACCURACY SCORES FOR AGE GROUPS 20 - 49 AND 50 - 69 YEARS - F-G

AGE	20-49	50-59	TOTAL
Less than 78	28,3	50,0	36,7
Greater than 78	71,7	50,0	63,3

Fisher's Exact Test = 0,0343 P <0,05

TABLE 9.5 COMPARISON OF ACCURACY SCORES FOR AGE GROUPS 50 - 69 YEARS AND 70 - 79 YEARS - F-G

AGE	50-69	70-79	TOTAL
Less than 78	50,0	85,0	62,1
Greater than 78	50,0	15,0	37,9

Fisher's Exact Test = 0,0110 P = <0,01

Thus differences were found in the performance of the age groups twenty to forty-nine, fifty to sixty-nine and seventy to seventy-nine. This was different from the other tests already discussed where the drop off occurred only at one point in the age continuum.

Therefore the null hypothesis that decreased ability in V-P-M function will not be found with advancing age, was rejected in the case of the figure-ground test.

Percentiles for the accuracy scores are to be found in Appendix 2.

9.4.2. Time Scores

The means and the standard deviations were calculated for the time scores for the group as a whole (Table 9.6) and for each age group (Table 9.7).

TABLE 9.6 MEAN AND SD OF TIME SCORES - F-G

n	MEAN/SECS	SD	RANGE
118	276,07	133,61	125-940

TABLE 9.7 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - F-G

AGE	n	MEAN/SECS	SD	RANGE
20-29	20	191,05	29,70	125-325
30-39	21	197,91	42,05	139-293
40-49	19	239,16	70,70	142-470
50-59	19	256,68	82,64	147-482
60-69	19	307,63	120,85	175-591
70-79	20	466,70	163,71	184-940

Levine's test established that there was a significant difference in the variance between the groups (P = 0,0000) and the Welch Test that there was a significant difference in the mean time scores obtained by each age groups (P = 0,0000), therefore the separate variance t with the Bonferroni correction for multiple comparisons was used to determine where these differences were (Table 9.8).

On the basis of the differences found in the time scores obtained by each age group, it was possible to re-group the ages as previously described. A similar analysis to that described above was carried out on the re-grouped data (Table 9.9).

Although there were no differences between the scores obtained by the forty to fifty-nine year olds and the sixty to sixty-nine year olds, percentiles were calculated for the age groups as previously defined, (see Appendix 2).

The progressive increase in the time taken and the fact that the older age groups took a significantly longer time to complete the test, enabled the null hypothesis of no significant difference in the time scores obtained by the different ages, to be rejected.

TABLE 9.8 DIFFERENCES IN TIME SCORES - F-G

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,5488
20-29	40-49	0,0113
20-29	50-59	0,0035
20-29	60-69	0,0006 **
20-29	70-79	0,0000 ***
30-39	40-49	0,0350
30-39	50-59	0,0097
30-39	60-69	0,0011 *
30-39	70-79	0,0000 ***
40-49	50-59	0,4870
40-49	60-69	0,0416
40-49	70-79	0,0000 ***
50-59	60-69	0,1392
50-59	70-79	0,0000 ***
60-69	70-79	0,0014 *

* P <0,05 ** P <0,01 *** P<0,001

TABLE 9.9 DIFFERENCES IN RE-GROUPED TIME SCORES - F-G

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-59	0,0003 **
20-39	60-69	0,0007 **
20-39	70-79	0,0000 ***
40-59	60-69	0,0603
40-59	70-79	0,0000 ***
60-69	70-79	0,0014 **

* P <0,05 ** P <0,01 *** P <0,001

9.4.3. Discussion Of The Results Of Testing Normal Individuals

A review of the literature yielded little information on age related changes in figure-ground ability. However, of the three studies available, two (Farver and Farver 1982 and Kaplan and Hier 1982) revealed significant age effects when using the Poppelreuter figures and the Ayres (1980) embedded figures. The overlapping figures of the Figure-Ground test of the Rivermead Perceptual Assessment Battery , Whiting et al (1985) did not prove to be sensitive to age related changes, probably due to the simplistic nature of the test.

It appeared, therefore, that the Figure-Ground test of the V-P-M Battery compiled for this thesis provided a standardised means of evaluating age related changes in perceptual function and revealed unusual changes at both the forty-nine year and the sixty-nine year levels.

9.5. RESULTS OF TESTING PATIENTS

9.5.1. Accuracy Scores

The means and the standard deviations were calculated for the age groups twenty to forty-nine and fifty to sixty-nine (Table 9.10) as these were the age groups within which no statistical difference occurred in the scores obtained by the normal group. As the score of only one patient was available, no calculations were carried out for the seventy to seventy-nine year old group. Four of the seventy year old patients were untestable and one was discharged prior to the execution of this test. Two patients in the age group fifty to sixty-nine years were also untestable.

TABLE 9.10 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - F-G

AGE	n	MEAN	SD	RANGE
20-49	11	74,36	4,30	63-79
50-69	14	69,00	11,07	50-80
70-79	1	-	-	58

9.5.2. Time Scores

The means and the standard deviations were calculated for the time scores for the age groups within which no statistical difference was found in the scores obtained by the normal group i.e. twenty to thirty-nine year olds and forty to sixty-nine year olds. Again no calculations were possible for the seventy to seventy-nine year olds as there was only one patient in this age group (Table 9.11).

TABLE 9.11 MEAN AND SD FOR TIME SCORES FOR PATIENTS - F-G

AGE	n	MEAN/SECS	SD	RANGE
20-39	7	518,71	256,20	223-941
40-69	18	587,50	325,26	212-1332
70-79	1	-	-	833

9.5.3. Comparison Between Patients And Normal Subjects

9.5.3.1. Accuracy Scores

The mean accuracy scores obtained by the patients were compared to those obtained by the normal subjects.

As Levine's test for equal variances was significant (twenty to forty-nine years $P = 0,0022$; fifty to sixty-nine years $P = 0,0000$) the Welch Test was used to determine whether there were significant differences in the scores obtained by the two groups (see Table 9.12).

TABLE 9.12 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - F-G

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20-49	0,0022	0,0070 **
50-69	0,0000	0,0223 *
70-79	-	-

* P <0,0 5 ** P <0,01

It can be seen that the difference in scores obtained by the normal and the patient groups was significant.

9.5.3.2. Time Scores

As with the accuracy scores, the mean time scores for the normal group and for the patient group were compared to determine whether there were any statistical differences.

As Levine's test for equal variances was significant (P = 0,0000) in both the age groups i.e. twenty to thirty-nine years and forty to sixty-nine years, the variances could not be assumed to be equal and the Welch Test was used (Table 9.13).

TABLE 9.13 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - F-G

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20-39	0,0000	0,0158 *
40-69	0,0000	0,0006 ***
70-79	-	-

* P < 0,05 ** P < 0,001

As can be seen from Table 9.13 there was a greater difference in the scores obtained by the older age groups than that obtained by the younger age group.

9.5.4. Analysis Of Tests Executed By Patients

An analysis of the scores obtained by the patients showed the following:-

a) In the twenty to thirty-nine year age group, of the seven patients tested, only one patient performed above the 15,9 percentile. Thus six patients could be said to have defective figure-ground perception.

b) Of the ten patients tested in the forty to fifty-nine year old group, only two patients performed within normal range, thus eight patients achieved scores which compared to the tenth percentile or lower and were considered defective.

c) Of the ten patients in the sixty year old group, two had severe problems and the testing was discontinued after the first few items on which no score was obtained. Six of the patients achieved scores within normal range. This was the only age group in which two patients obtained the maximum score.

d) In the age group seventy to seventy-nine years, four of the six patients were untestable, one was discharged prior to completing the test and one obtained a score well below normal limits. The patients in this age group found this test extremely difficult.

9.5.5. Problems That May Be Diagnosed From Performance On The Figure-Ground Test

a) Left neglect - this was obvious when the patient pointed to only those figures, letters, numbers and words that were on the right of the multiple choice sheet (patient nine).

b) Visual inattention - this was seen when a patient was unable to focus his attention on all parts of the stimulus and thus neglected to point out objects. This neglect occurred randomly and was not restricted to one part of the visual field (patient number eighteen and twenty-one).

c) Concentration - patient number twenty-three commented that concentration was needed for this test.

d) Short term memory (perceptual memory) - Patient number fifteen and twenty-one had difficulty in remembering which cues they had already pointed out, therefore repeating the items.

e) Perseveration - Patient number thirty-three pointed to each item many times before being able to move on to the next item.

f) As the test was constructed to evaluate figure-ground ability, low scores on the test indicated a disturbance of this aspect of perception.

9.6. ITEM ANALYSIS

9.6.1. Difficulty Index

The item difficulty levels can be seen in Table 9.14.

From Table 9.14 it is apparent that the items dealing with words (i.e. items thirteen to sixteen) were the easiest in the test and should be placed first followed by the objects/designs and then the individual letters and numbers.

The difficulty index for the responses obtained from the patient group was also calculated and compared with the responses obtained in the normal group (see Table 9.15).

The difficulty index for the seventy to seventy-nine year old patients could not be calculated as only one patient completed the test. As can be seen from Table 9.15 only items one and two demonstrated no significant difference in the results obtained by the patient group and the normal group and therefore on this basis can be dropped from the test.

TABLE 9.14 DIFFICULTY INDEX - F-G

ITEM	DIFFICULTY p 20-49 YRS	DIFFICULTY p 50-69 YRS	DIFFICLUTY p 70-79 YRS
1 objects	0,97	0,94	1,00
2 designs	0,98	0,83	0,90
3 designs	0,87	0,81	0,60
4 designs	0,90	0,78	0,75
5 designs	0,92	0,83	0,60
6 designs	0,82	0,75	0,70
7 designs	0,93	0,69	0,55
8 designs	0,97	0,81	0,80
9 letters	1,00	0,92	0,70
10 letters	0,90	0,83	0,50
11 numbers	0,80	0,53	0,30
12 numbers	0,98	0,97	1,00
13 words	0,97	0,94	0,90
14 words	0,97	1,00	0,90
15 words	0,97	1,00	0,80
16 words	0,98	0,97	0,85

9.6.2. Item Discrimination

This is the extent to which the item discriminates between those individuals who can do well on the test and those who can not. Two ways of determining item discrimination were used.

TABLE 9.15 DIFFICULTY INDEX - PATIENT GROUP - F-G

Item No	DIFFICULTY p 20-49 YRS	DIFFICULTY p 50-69 YRS
1	0,91	0,79
2	0,91	0,57
3	0,55 *	0,71
4	0,55 ***	0,64
5	0,73 ***	0,50 *
6	0,73 ***	0,36
7	0,64 **	0,50
8	0,64 ***	0,57
9	0,91 *	0,71
10	0,82	0,29 ***
11	0,55	0,14 *
12	0,82 **	0,79 *
13	0,91	0,71 *
14	0,64 **	0,71 **
15	0,91	0,71 **
16	0,91	0,57 ***

* P < 0,05 ** P < 0,01 *** P < 0,001

9.6.2.1. Testing the Significance of the Difference between Two Proportions

The responses obtained by those normal individuals whose overall scores were in the upper 27% were compared with those whose scores were in the lower 27%. Table 9.16 gives the result of Fisher's Exact test in making this comparison for the ages twenty to forty-nine years, fifty to sixty-nine years and seventy to seventy-nine years. A cut point of four was used for this purpose, the maximum score per item being five. Only half of the items i.e. items three, four, six, seven, eight, nine, eleven and fifteen discriminated between those who could do the test and those who could not.

TABLE 9.16 ITEM DISCRIMINATION - F-G

ITEM	20-49 YRS	50-69 YRS	70-79 YRS
1	-	-	-
2	-	-	0,4444
3	0,1012	-	0,0476 *
4	-	0,4706	0,0476 *
5	0,1012	0,4706	0,4444
6	0,1012	0,0294 *	0,0476 *
7	0,2258	0,0090 **	0,0476 *
8	-	0,0470	0,0476 *
9	-	-	0,0476 *
10	0,2258	0,2059	0,5238
11	0,0024 **	0,0090 **	0,2063
12	-	-	-
13	-	-	0,4444
14	-	-	1,000
15	-	-	0,0476 *
16	-	-	-

* P<0,05 ** P <0,01 ** P <0,001

It was apparent that items thirteen to sixteen (i.e. the words) had a very low discriminative level and consideration should be given to dropping them from the final test with the possible retention of item fifteen.

9.6.2.2. Correlational Approach

This approach showed the relationship of the individual responses to the total test score (Table 9.17).

Using the correlational approach it appeared that all items had statistically significant correlation values when all the age groups were considered, except for items one, twelve and sixteen. When analysing the scores obtained by the three different age groups, it again appeared that items thirteen to sixteen, as well as, item one could be excluded on the basis of their poor ability to discriminate between those who could do well on the test and those who did poorly.

TABLE 9.17 BI-SERIAL CO-EFFICIENT OF CORRELATION BETWEEN ITEM AND TOTAL SCORE - F-G

	20-49 YRS			50-69 YRS		
Item	Upper 27%	Lower 27%	r Value	Upper 27%	Lower 27%	r Value
1	-	-	-	-	-	-
2	-	-	-	-	-	-
3	1,0	0,75	0,60***	-	-	-
4	1,0	0,75	,060***	1,0	0,78	0,53***
5	-	-	-	1,0	0,78	0,53***
6	1,0	0,75	0,60***	1,0	0,44	0,74***
7	1,0	0,81	0,52***	1,0	0,33	0,80***
8	-	-	-	1,0	0,78	0,53***
9	-	-	-	-	-	-
10	1,0	0,81	0,52***	1,0	0,67	0,62***
11	1,0	0,50	0,72***	1,0	0,33	0,80***
12	-	-	-	-	-	-
13	-	-	-	-	-	-
14	-	-	-	-	-	-
15	-	-	-	-	-	-
16	-	-	-	-	-	-

* P <0,05   ** P <0,01   *** P <0,001

TABLE 9.17 (Continued)

Item	70-79 YRS		
	Upper 27%	Low 27%	r Value
1	-	-	-
2	1,0	0,60	0,70 ***
3	1,0	0,20	0,85***
4	1,0	0,20	0,85***
5	1,0	0,600	0,70**
6	1,0	0,20	0,85***
7	1,0	0,20	0,85***
8	1,0	0,20	0,85***
9	1,0	0,20	0,85***
10	0,80	0,40	0,43
11	0,80	0,20	0,60*
12	-	-	-
13	1,0	0,60	0,70**
14	1,0	0,80	0,50
15	1,0	0,20	0,85***
16	-	-	-

* P <0,05 ** P <0,01 *** P <0,001

9.7. TEST RELIABILITY

9.7.1. Split-Half Reliability

This determined the internal consistency of a test with regards to content sampling. The resultant co-efficient of correlation between the two halves of the test was r = 0,44 for the age group twenty to forty-nine, r = 0,92 for the age group fifty to sixty-nine years, r = 0,89 for the age groups seventy to seventy-nine years with a mean of 0,75.

TABLE 9.17 (Continued)

Item	70-79 YRS		
	Upper 27%	Low 27%	r Value
1	-	-	-
2	1,0	0,60	0,70 ***
3	1,0	0,20	0,85***
4	1,0	0,20	0,85***
5	1,0	0,600	0,70**
6	1,0	0,20	0,85***
7	1,0	0,20	0,85***
8	1,0	0,20	0,85***
9	1,0	0,20	0,85***
10	0,80	0,40	0,43
11	0,80	0,20	0,60*
12	-	-	-
13	1,0	0,60	0,70**
14	1,0	0,80	0,50
15	1,0	0,20	0,85***
16	-	-	-

* P <0,05 ** P <0,01 *** P <0,001

9.7. TEST RELIABILITY

9.7.1. Split-Half Reliability

This determined the internal consistency of a test with regards to content sampling. The resultant co-efficient of correlation between the two halves of the test was r = 0,44 for the age group twenty to forty-nine, r = 0,92 for the age group fifty to sixty-nine years, r = 0,89 for the age groups seventy to seventy-nine years with a mean of 0,75.

As in several of the other tests, the split-half reliability for the younger age group was considerably lower than for the older age groups (see Table 9.18).

9.7.2. Kuder-Richardson Formula 20

In this method, the pass-fail score for each item was compared with the total test score. The resulting co-efficient of correlation was r = 0,61 for the age group twenty to forty-nine years, r = 0,96 for the age group fifty to sixty-nine years and r = 0,87 for the seventy to seventy-nine year age group (Table 9.18).

All reliability co-efficients were found to be statistically significant. However, as co-efficients of 0,80 or above are traditionally accepted as evidence of adequate test reliability, it seemed that the figure-ground test was most reliable when testing the older age groups.

TABLE 9.18 RELIABILITY CO-EFFICIENTS - F-G

AGE	SPLIT-HALF	KUDER-RICHARDSON
20-49	0,44 **	0,61 ***
50-69	0,92 ***	0,96 ***
70-79	0,89 ***	0,87 ***

** P < 0,01 *** P <0,001

9.7.3. Inter-Rater Reliability

The Friedman two-way analysis of variance showed that there was no statistical difference between either the accuracy (P = 0,4169) or time (P = 0,5134) scores obtained by the three examiners on each of six individuals tested.

9.8. TEST VALIDITY

The following types of validity were considered:-

9.8.1. Content Validity

Content validity was provided by the fifteen occupational therapists (population 1) who rated the test on a zero to five scale according to how well they found the test measured the ten V-P-M skills (Table 9.19). The therapists rated this test as a very much better than average assessment of the perceptual skill of figure-ground. It seemed, therefore, from the results of the therapists ratings, that other perceptual skills did not feature in this particular test.

TABLE 9.19 CONTENT VALIDITY - F-G.

PERCEPTUAL SKILL	MEAN RATING
Form as mass	1,3
Overall contour	0,6
Sub-Units of Form	0,4
Perceptual constancy	1,3
Figure ground	4,0 ***
Directionality	0,5
Space relations	1,1
Position in space	0,7
Colour	0,0
Praxis	not required

*** a very much better than average assessment

9.8.2. Construct Validity

9.8.2.1. Age Differentiation

In the case of the test of figure-ground ability, the test discriminated between the performance of three categories of subjects i.e. twenty to forty-nine year olds, fifty to sixty-nine year olds and seventy to seventy-nine year olds. The Figure-Ground test would, therefore, appear to be a valid method of differentiating age related changes.

9.8.2.2. Correlation with other tests in the Battery

The highest correlation of the Figure-Ground test with other tests in the battery was with the Dot-Line test (r = 0,62). It also correlated with the Design Construction test (r = 0,59) and the VMI (r = 0,51). It was interesting to note that other tests that appeared to need figure-ground ability i.e. the Space Relations test (r = 0,48) and the Position in Space test (r = 0,42) had fairly low correlations with the Figure-Ground test.

A factor analysis was carried out on the tests in the V-P-M Battery to determine whether they contained factors related to each other. The Figure-Ground test loaded for both Factor I and Factor II (see Table 9.20). Thus the test appeared to belong to the cluster that may be said to indicate different forms of visual discrimination (i.e. Factor II) as well as different spatial concepts (Factor I). As with the other tests no marker variables were included, therefore the results can not be readily used for the purpose of construct validity. It was interesting to note that Siev et al (1986) classified figure-ground perception as part of a spatial syndrome.

9.8.2.3. Results of Testing Patients

The test was found to be sensitive to decreased function in patients at both the younger (i.e. twenty to forty-nine) and the older age groups (i.e. fifty to sixty-nine). It would seem therefore to be a valid tool to evaluate perceptual function in the adult C.V.A. patient.

From the foregoing it would appear that the Figure-Ground test is a valid means of differentiating age related changes as well as changes in V-P-M function in adults who have suffered a cerebro-vascular accident.

TABLE 9.20 SORTED ROTATED FACTOR LOADINGS - F-G

TEST	FACTOR I	FACTOR II
Position in Space	0,80	0,0
Block Construction II	0,74	0,0
Part-Whole	0,73	0,0
VMI	0,63	0,41
Dot-Line	0,61	0,52
Figure-Ground	0,59	0,54
Space Relations	0,0	0,78
Arrows	0,0	0,75
Stick Construction II	0,0	0,60
Design Construction	0,51	0,57
Perceptual Constancy	0,37	0,52

9.9. CONCLUSIONS

In summary the following has been found:-

- visual masking was used in this Figure-Ground test as it was felt that this type of task was more appropriate to evaluate the skill of visual figure ground perception than embedded figures. It is believed that embedded figures require the use of the skill of space visualisation;
- the test was easy to administer and to score with the inter-rater reliability being good;
- the test was very sensitive to changes in function in the normal individuals and demonstrated changes in ability at a younger age than the other tests in the battery;
- as in all the other tests, the older person took longer to complete the test;

- differences between the performance of the normal group and that of the patient group were demonstrated, therefore the test was sensitive to perceptual dysfunction as a result of a right sided C.V.A.;

- item analysis has shown that the test could be improved by dropping items one, two and thirteen to sixteen;

- reliability co-efficients, although low for the younger age group were well within acceptable limits for the older group. However, in spite of the co-efficients being low for ages twenty to forty-nine years they were still statistically significant;

- it was also shown that the test was a valid test of figure-ground perception. The fact that the test loads for both Factor I and II seemed to indicate that it made a unique contribution to both these factors, especially as the factor loadings were just within the cut off point of 0,50. Correlations with other tests in the battery were not very high, the highest being with the Dot-Line test r = 0,62, the Design Construction test r = 0,59, and the VMI r = 0,51.

The fact that the other tests in the battery that appeared to need some visual discriminative ability did not correlate highly with this test suggested that this test measured a discrete skill. Because of this and the robust nature of the items to be kept in the test, it is suggested that the Figure-Ground test be retained in the V-P-M Battery.

CHAPTER 10

TESTS OF DIRECTIONALITY - THE ARROWS AND THE DOT-LINE TESTS

10.1. INTRODUCTION

Directionality has been defined as "the ability to perceive the directional location of an object in relation to the line of gravity and to transpose this concept into one dimensional form on paper". Directionality is thus a space concept which belongs to the subjective group of space or grasping space. Directionality in this context has a line/pointing connotation and developmentally appears before that of spatial relations. It deals with the skills of up, down, towards and away from.

As far as can be ascertained there are no tests specifically designed to evaluate directionality although this skill is discussed in the literature dealing with the treatment of perceptual problems in children, (Kephart 1960, Banus and Wills 1977, Frostig and Maslow 1973). The literature on adult perceptual problems mentions *"sense of direction"* (Benton 1979, Lezak 1983).

The Arrows test described in Chapter Four (4.6.6.) and Appendix 1 Section 2.6 is a motor free test. The Dot-Line test is to be found in Section 4.6.7. of Chapter Four and Appendix 1 Section 2.7 is a motor enhanced test. Both tests were compiled to evaluate directionality.

10.2. REVIEW OF SIMILAR TESTS

10.2.1. Motor Free Tests

In the standardised Road-Map Test of Directional Sense (Money 1976) the subject is required to tell the examiner the direction taken as the examiner draws a dotted pathway through a map of roadways (Fig. 10.1). Norms up to 18 years only are available for this test.

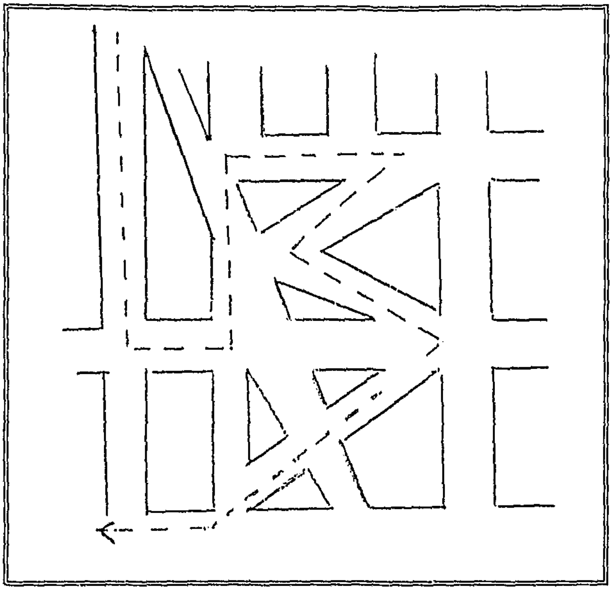


FIGURE 10.1 ROAD MAP TEST OF DIRECTIONAL SENSE (Money 1976)

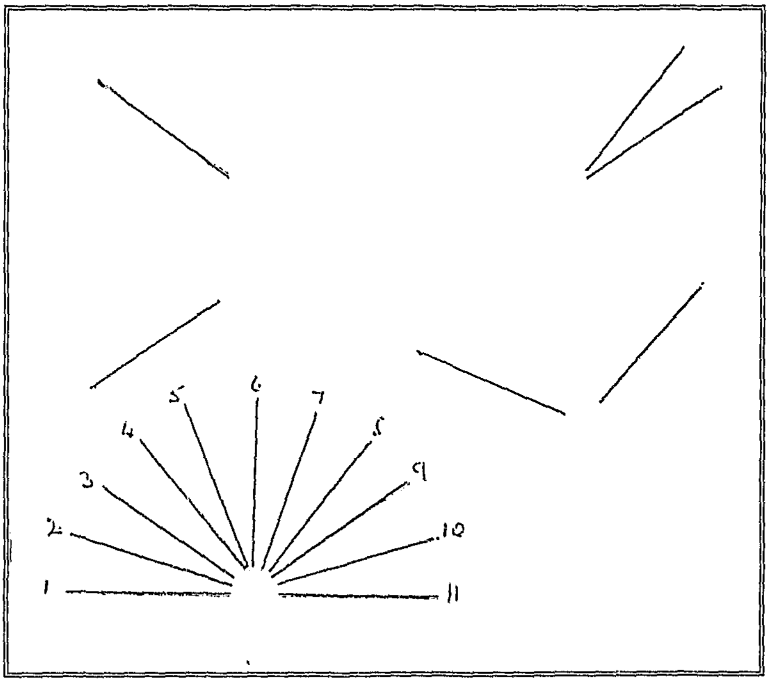


FIGURE 10.2 DIRECTIONAL ORIENTATION TEST (Benton 1979)

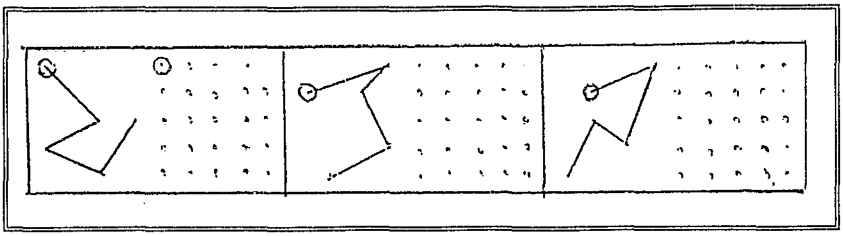


FIGURE 10.3 THE MACQUARRIE TEST OF MECHANICAL ABILITY - COPYING SUB-TEST (MacQuarrie 1953)

Benton (1979) described a test which required the subject to identify the directional orientation of lines. Single and double line stimuli were shown, followed by the presentation of a response card on which the patient identified the line or lines that he had seen. (Fig. 10.2).

10.2.2. Motor Tests Of Directional Sense

There are several tests requiring the subject to draw lines between dots. These are part of tests designed to evaluate perceptual function in children i.e. the Spatial Relations component of the Developmental Test of Visual Perception (Frostig 1966), Dennis Visual Perception Scale (Dennis and Dennis 1969), the Design Copying component of the Southern California Sensory Integration Test (Ayres 1980a).

The Fret Repetition test (Grant 1969) is a dot-to-dot test designed to evaluate perceptual analysis in illiterate workers only. This and the "copying" sub-test of the MacQuarrie Test for Mechanical Ability (MacQuarrie 1953) (Fig. 10.3) appear to be the only dot-to-dot tests with normative data for adults. Both these tests are not generally used with brain damaged subjects.

Goody and Reinhold formulated a task which required the subject to draw arrows pointing in different directions (Lezak 1983).

Many of the stick construction tests discussed in Chapter Thirteen can also be used to evaluate the concept of directionality.

10.3. REASONS FOR COMPILING NEW TESTS

10.3.1. The Arrows Test

The standardised Road Map Test (Money 1976) demanded a verbal response and was eliminated from the V-P-M Battery for this reason as attempts were made to limit the verbal components as much as possible both in instructions and in response. The Benton Test (1979) in which the subject is expected to identify the directional orientation of a line was eliminated on account of the memory component. Thus a motor free test was required to evaluate directional sense and the Arrows test was compiled accordingly.

For the Arrows test the subject has to choose between two arrow designs, the one that matches a design on the multiple choice sheet.

10.3.2. The Dot-Line Test

Although several dot-to-dot tests exist, they have been constructed for children and were thought to be too simple for adults. Accordingly a new drawing test was constructed that included designs requiring the subject to estimate the direction and placement of his lines. For this test the subject has to copy a design in which lines have been drawn between dots.

10.4. RESULTS OF TESTING NORMAL INDIVIDUALS

10.4.1. Accuracy Scores -The Arrows Test

The means and standard deviations were calculated for the normal group as a whole (Table 10.1) and for each age group individually (Table 10.2).

TABLE 10.1 MEAN AND SD OF ACCURACY SCORES - AR

n	MAX SCORE	MEAN	SD	RANGE
118	24	23,08	1,28	15-24

TABLE 10.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - AR

AGE	n	MEAN	SD	RANGE
20-29	20	23,2	1,11	20-24
30-39	21	23,38	0,97	20-24
40-49	19	23,11	0,88	22-24
50-59	19	23,42	0,84	21-24
60-69	19	23,05	0,85	22-24
70-79	20	22,30	2,20	15-24

Although all means were very close, the Pearson Chi-square test was used to determine the existence of any differences in the mean scores, a cut-point of twenty-two being used. No significant differences were found amongst the different age groups in the performance of the normal individuals on the Arrows test (Table 10.3).

Percentiles, however, were calculated for the age groups previously mentioned (see Appendix 2 Section 2.6).

10.4.2. Accuracy Scores - The Dot-Line Test

The means and standard deviations were calculated for the normal group as a whole (Table 10.4) and for each group individually (Table 10.5).

TABLE 10.3 COMPARISON OF ACCURACY SCORES - AR

Age	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 22	25,0	9,5	31,6	10,5	31,6	35,0	23,7
Greater than 22	75,0	90,5	68,4	89,5	68,4	65,0	76,3

Pearson Chi-square Value = 6,887 P = 0,2292

TABLE 10.4 MEAN AND SD OF ACCURACY SCORES - D-L

n	MAX SCORE	MEAN	SD	RANGE
118	16	13,95	2,14	6-16

TABLE 10.5 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - D-L

AGE	SD	MEAN	SD	RANGE
20-29	20	14,3	1,63	10-16
30-39	21	14,81	0,81	13-16
40-49	19	14,79	1,69	11-16
50-59	19	14,37	1,46	11-16
60-69	19	13,68	2,38	6-16
70-79	20	11,75	2,80	7-16

There was a noticeable drop off in the mean of the accuracy scores from about sixty years of age. To determine whether these differences were significant the Pearson Chi-square test was used with a cut-point of fourteen (Table 10.6).

TABLE 10.6 COMPARISON OF ACCURACY SCORES - DL

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 14	45,0	33,3	31,6	42,1	57,9	80,0	48,3
Greater than 14	55,0	66,7	68,4	57,9	42,1	20,0	51,7

Pearson Chi-square Value = 13,139 P <0,05

As significant differences occurred within the age groups constituting the sample, further tests were carried out as follows:-

a) The Pearson Chi-square Test demonstrated that there were no significant differences in the accuracy scores obtained by the age groups twenty to fifty-nine years.

b) Fisher's Exact test showed that there were no significant differences in the age groups sixty to seventy-nine years.

c) The age groups twenty to fifty-nine years and sixty to seventy-nine years were then combined and compared using Fisher's Exact test. A highly significant difference was found between these age groups, P <0,01 (see Table 10.7).

The overall grouping as previously discussed was used for the calculation of percentile ranks (see Appendix 2) in keeping with other tests in the battery.

TABLE 10.7 COMPARISON OF ACCURACY SCORES FOR AGE GROUPS 20 - 59 AND 60 - 79 YEARS - D-L

AGE	20-59	60-79	TOTAL
Less than 14	38,0	69,2	48,3
Greater than 14	62,0	30,8	51,7

Fisher's Exact test P = 0,0017 P <0,01

10.4.3. Time Scores - The Arrows Test

The means and standard deviations were calculated for the time scores for the group as a whole (see Table 10.8) and for each age group (see Table 10.9).

TABLE 10.8 MEAN AND SD OF TIME SCORES - AR

n	MEAN/SECS	SD	RANGE
118	103,46	52,72	43-367

TABLE 10.9 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - AR

AGE	n	MEAN/SECS	SD	RANGE
20-29	20	70,8	18,84	43-124
30-39	21	80,24	27,91	49-145
40-49	19	83,79	22,21	44-147
50-59	19	87,21	28,19	56-148
60-69	19	124,90	26,33	73-192
70-79	20	174,25	73,24	73-367

Levine's test for equal variance established that there was a significant difference in the variances among the groups (P = 0,0000) and the Welch Test demonstrated that there was a significant difference in the mean of the time scores obtained by the different groups (P = 0,0000). The separate variance t was, therefore, used with the Bonferroni correction for multiple comparisons, to establish where the differences were to be found (Table 10.10).

TABLE 10.10 DIFFERENCES IN TIME SCORES - AR

AGE		SEPARATE VARIANCE t P VALUE
20 - 29	30 - 39	0,2108
20 - 29	40 - 49	0,0573
20 - 29	50 - 59	0,0415
20 - 29	60 - 69	0,0000 ***
20 - 29	70 - 79	0,0000 ***
30 - 39	40 - 49	0,6572
30 - 39	50 - 59	0,4374
30 - 39	60 - 69	0,0001 **
30 - 39	70 - 79	0,0000 ***
40 - 49	50 - 59	0,6803
40 - 49	60 - 69	0,0002 **
40 - 49	70 - 79	0,0000 ***
50 - 59	60 - 69	0,0011 *
50 - 59	70 - 79	0,0000 ****
60 - 69	70 - 79	0,0120

* P <0,05 ** P <0,01 *** P <0,001

The age groups among which no statistical difference occurred in the scores were then grouped together and the analysis repeated as above (Table 10.11).

The analysis revealed that, there were no differences in the scores obtained by the twenty to thirty-nine year olds and the forty to fifty-nine year olds as well as the sixty to sixty-nine and seventy to seventy-nine year olds. However, these groups were kept separate for the compilation of percentiles (see Appendix 2).

TABLE 10.11 DIFFERENCES IN RE-GROUPED TIME SCORES - AR

AGE		SEPARATE VARIANCE T P VALUE
20 - 39	40 - 59	0,0791
20 - 39	60 - 69	0,0000 ***
20 - 39	70 - 79	0,0000 ***
40 - 59	60 - 69	0,0002 **
40 - 49	70 - 79	0,0000 ***
60 - 69	70 - 79	0,0120

** P <0,01 *** P <0,001

10.4.4. Time Scores - Dot-Line Test

The means and standard deviations were calculated for the time scores for the group as a whole (Table 10.12) and for each age group (Table 10.13).

TABLE 10.12 MEAN AND SD OF TIME SCORES - D-L

n	MEAN/SEC	SD	RANGE
118	617,48	304,89	287-2124

Levine's test for equal variance established that there was a significant difference in the variances among the different age groups (P = 0,0000) and the Welch test that there was a significant difference in the mean of the time scores (P = 0,0000). The separate variance t was, therefore, used with the Bonferroni correction for multiple comparisons to establish where the differences were to be found (Table 10.14).

TABLE 10.11 DIFFERENCES IN RE-GROUPED TIME SCORES - AR

AGE		SEPARATE VARIANCE T P VALUE
20 - 39	40 - 59	0,0791
20 - 39	60 - 69	0,0000 ***
20 - 39	70 - 79	0,0000 ***
40 - 59	60 - 69	0,0002 **
40 49	70 - 79	0,0000 ***
60 - 69	70 - 79	0,0120

** P <0,01 *** P <0,001

10.4.4. Time Scores - Dot-Line Test

The means and standard deviations were calculated for the time scores for the group as a whole (Table 10.12) and for each age group (Table 10.13).

TABLE 10.12 MEAN AND SD OF TIME SCORES - D-L

n	MEAN/SEC	SD	RANGE
118	617,48	304,89	287-2124

Levine's test for equal variance established that there was a significant difference in the variances among the different age groups (P = 0,0000) and the Welch test that there was a significant difference in the mean of the time scores (P = 0,0000). The separate variance t was, therefore, used with the Bonferroni correction for multiple comparisons to establish where the differences were to be found (Table 10.14).

TABLE 10.13 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - D-L

AGE	n	MEAN	SD	RANGE
20-29	20	436,60	127,48	287-767
30-39	21	478,24	101,45	289-689
40-49	19	559,26	79,84	417-716
50-59	19	517,10	144,98	327-968
60-69	19	699,68	233,11	347-1289
70-79	20	1017,05	473,04	414-2424

The results indicated that the test scores could be combined as previously described. Table 10.15 gives the results of the analysis of the re-grouped scores.

The percentiles were compiled as previously explained (see Appendix 2).

10.4.5. Discussion Of The Results Of Testing Normal Individuals

The motor free test for directionality i.e. the Arrows test demonstrated no age related changes and thus was comparable to the two other tests that use a similar process of matching test cards to the answer on a multiple choice sheet i.e. the Colour Matching test and the Form Matching test. Unfortunately there were no other adult tests with which to compare these results.

TABLE 10.14 DIFFERENCES IN TIME SCORES - D-L

AGE		SEPARATE VARIANCE t P VALUE
20 - 29	30 - 39	0,2562
20 - 29	40 - 49	0,0010 *
20 - 29	50 - 59	0,0744
20 - 29	60 - 69	0,0002 **
20 - 29	70 - 79	0,0000 ***
30 - 39	40 - 49	0,0077
30 - 39	50 - 59	0,3380
30 - 39	60 - 69	0,0008 *
30 - 39	70 - 79	0,0001 ***
40 - 49	50 - 59	0,2763
40 - 49	60 - 69	0,0210
40 - 49	70 - 79	0,0004 **
50 - 59	60 - 69	0,0069
50 - 59	70 - 79	0,0002 **
60 - 69	70 - 79	0,0123

* P <0,05 ** P <0,01 *** P <0,001

In the Dot-Line test which required a motor output, the decreased ability to perform the test occurred at approximately sixty years of age, as in the test of Visual Motor Integration (VMI), the other drawing test in the battery. Just as the drawings completed by the elderly individuals on the VMI tended to look like those of a patient with a right hemisphere lesion, so did the tests of the elderly on the Dot-Line test. Neglecting part of the design (Figure 10.4) linking incorrect dots so that the lines did not go in the correct direction (Figure 10.5), spatial misplacement of dots and lines (Figures 10.6, and 10.7) could be seen.

TABLE 10.15 DIFFERENCES IN RE-GROUPED TIME SCORES - D-L

AGE		SEPARATE VARIANCE t P VALUE
20 - 39	40 - 59	0,0030 *
20 - 39	60 - 69	0,0003 **
20 - 39	70 - 79	0,0000 ***
40 - 59	60 - 69	0,0093
40 - 59	70 - 79	0,0002 **
60 - 69	70 - 79	0,0123

* P<0,05 ** P <0,01 *** P <0,0001

It may, therefore, be inferred that the Dot-Line test is an adequate measure of age related changes in perceptual function. However, the Arrows test is not. In the case of the Dot-Line test hypothesis I is rejected but in the case of the Arrows test, it is retained.

10.5. Results Of Testing Patients

10.5.1. Accuracy Scores - The Arrows Test

The mean and standard deviation were calculated for all age groups together as no statistical difference was found in the different age groups for the normal subjects (Table 10.16). Only twenty-four patients were included in the analysis as four patients were discharged prior to the completion of the test and five were found to be unable to do the test. Four in this latter category were in the seventy to seventy-nine year age group.

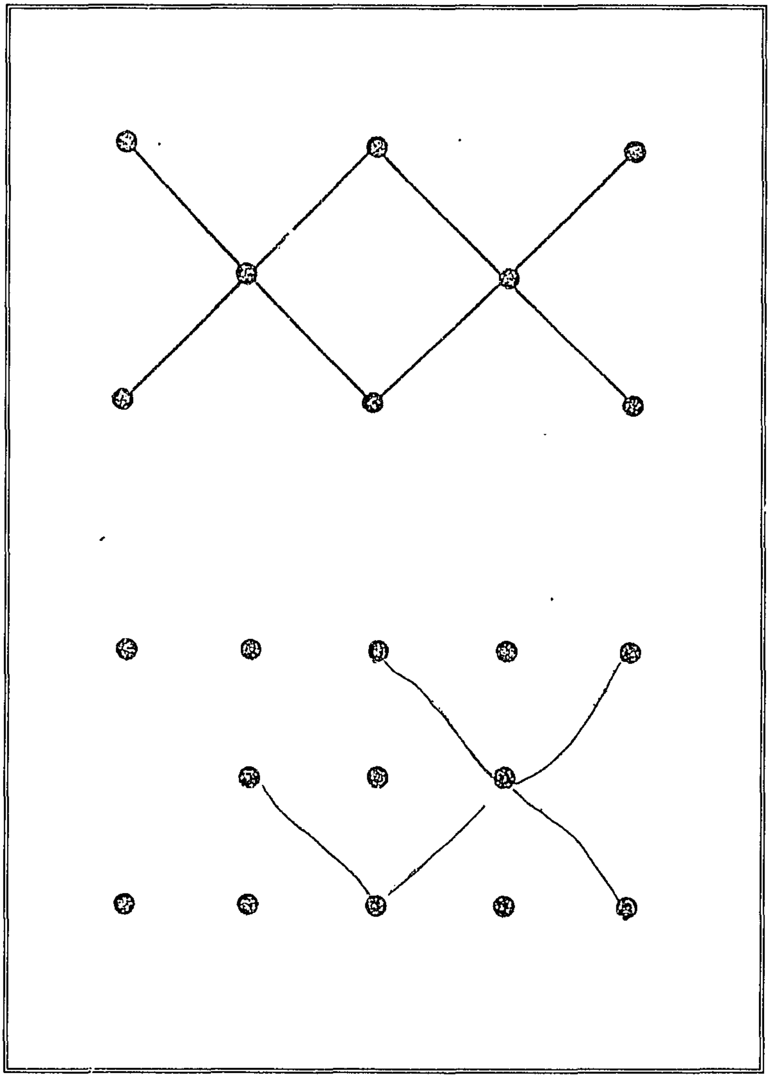


FIGURE 10.4 NEGLECTING PART OF DESIGN-NORMAL 68 YEAR OLD SUBJECT

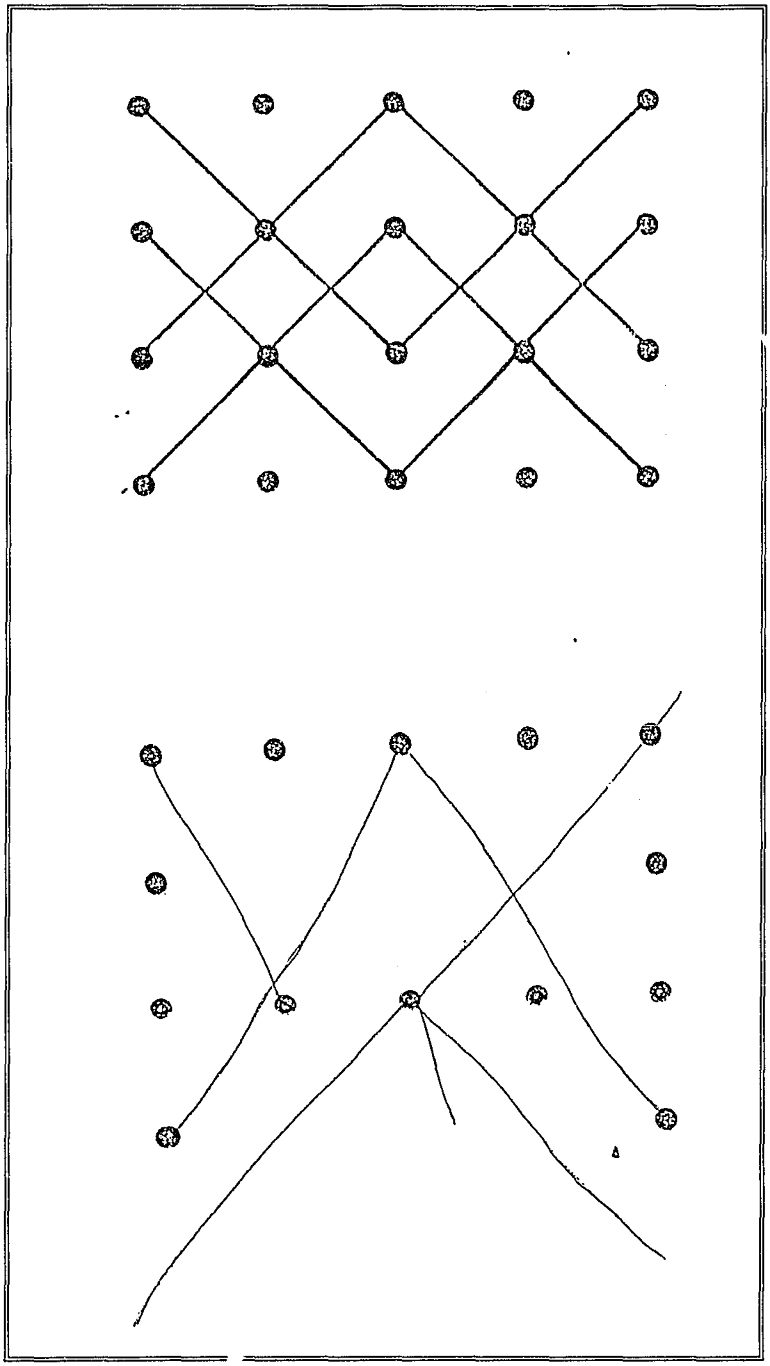


FIGURE 10.5 ıNKING INCORRECT DOTS - NORMAL 72 YEAR OLD SUF JECT

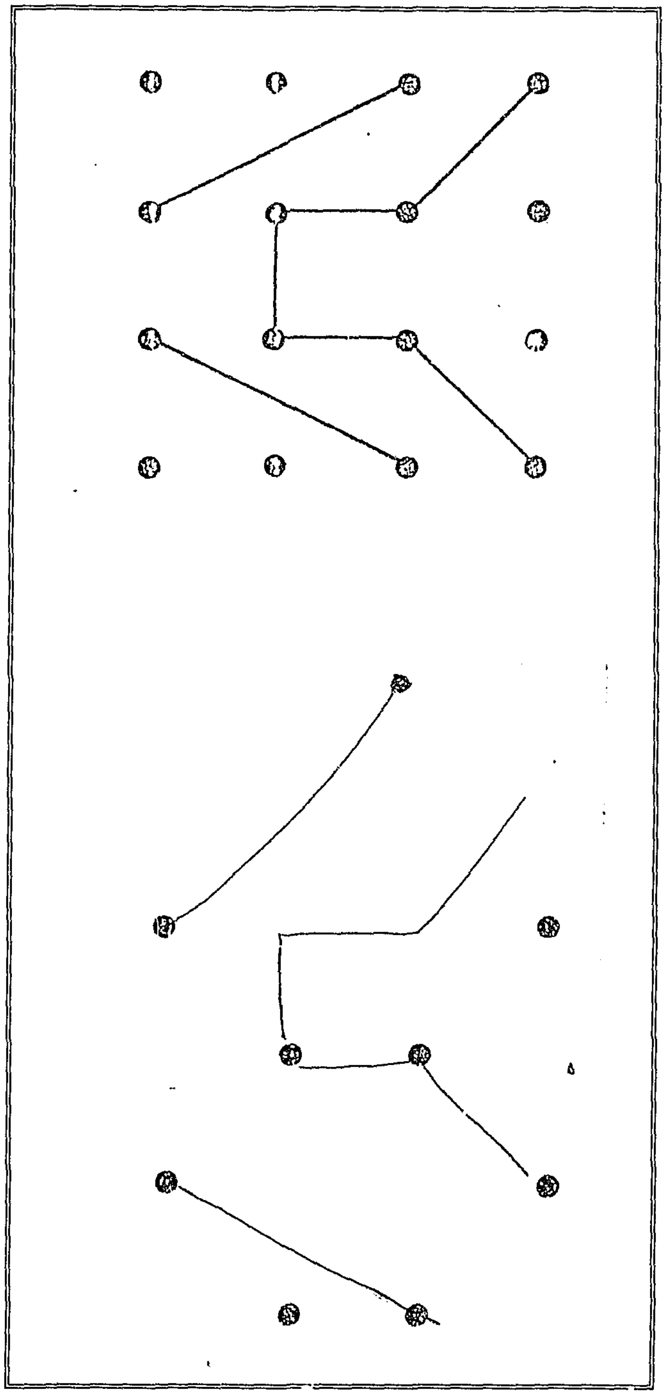


FIGURE 10.6 SPATIAL MISPLACEMENT OF LINES AND DOTS-NORMAL 72 YEAR OLD SUBJECT

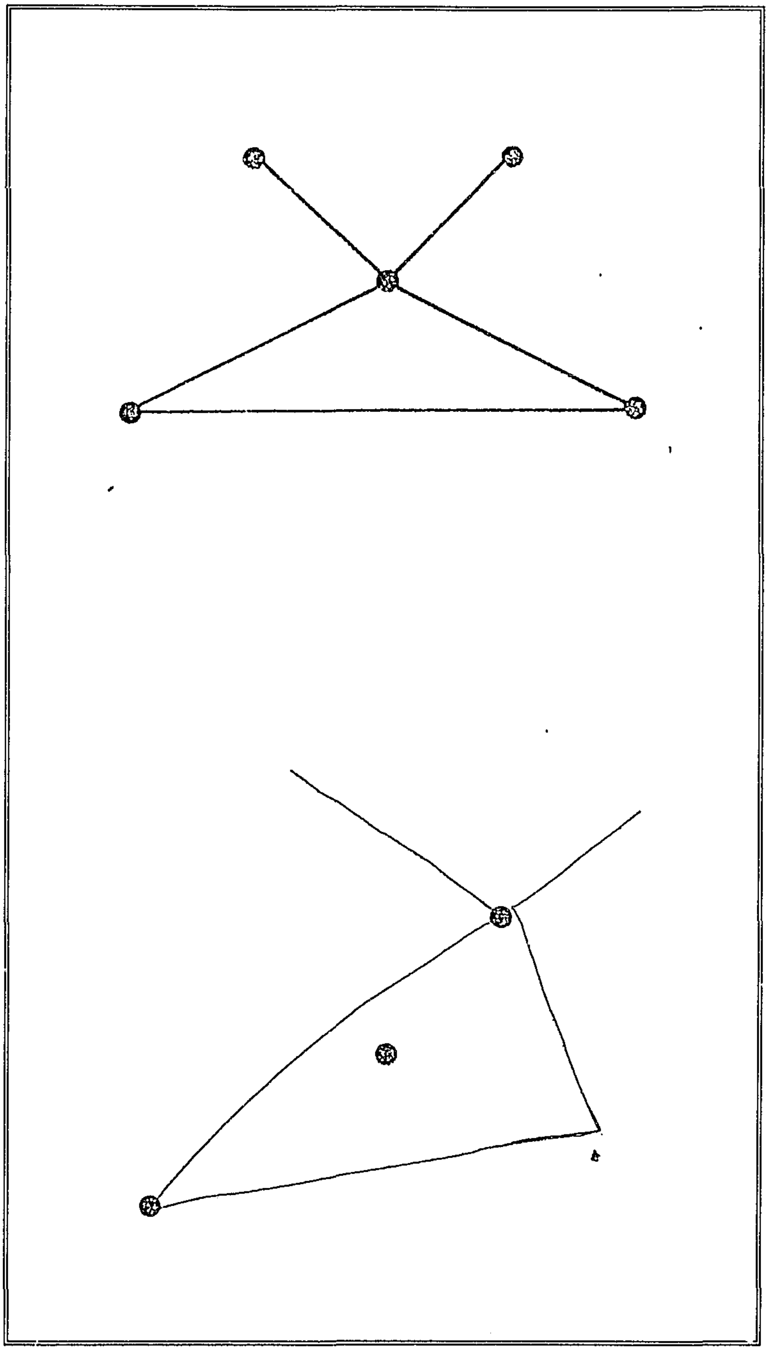


FIGURE 10.7 SPATIAL MISPLACEMENT OF LINES - NORMAL 72 YEAR OLD SUBJECT

TABLE 10.16 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - AR

AGE	n	MEAN	SD	RANGE
20-79	24	20,29	4,25	9 - 24

10.5.2. Accuracy Scores - The Dot-Line Test

The means and standard deviations of the accuracy scores were calculated for the age groups among which no statistical difference occurred in the scores obtained by the normal subjects i.e. the twenty to fifty-nine year olds and sixty to seventy-nine year olds (Table 10.17). In the age group twenty to fifty-nine years one patient was unable to do the test and in the sixty to seventy-nine year group, one patient was discharged prior to completing the test. In this group testing was discontinued in two cases due to the patients obvious inability to do the tasks and one patient was unable to be tested at all.

TABLE 10.17 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - D-L

AGE	n	MEAN/SECS	SD	RANGE
20 - 59	16	10,88	3,28	5 - 16
60 - 79	12	9,0	4,35	1 - 15

10.5.3. Time Scores - The Arrows Test

The means and standard deviations were calculated for the age groups among which no statistical difference was found in the normal group i.e. twenty to fifty-nine years and sixty to seventy-nine years (Table 10.18).

TABLE 10.18 MEAN AND SD FOR TIME SCORES FOR PATIENTS - AR

AGE	n	MEAN/SECS	SD	RANGE
20 - 59	15	228,00	145,10	86 - 563
60 - 79	9	353,22	264,8	98 - 947

10.5.4. Time Scores - The Dot-Line Test

The means and standard deviations were calculated for the age groups among which no statistical difference was found in the normal group i.e. the twenty to thirty-nine year olds and forty to sixty-nine year olds (see Table 10.19). The standard deviation was not calculated for the age group seventy to seventy-nine years as there were only two subjects in this group who completed the test.

TABLE 10.19 MEAN AND SD FOR TIME SCORES FOR PATIENTS - D-L

AGE	n	MEAN/SECS	SD	RANGE
20 - 39	7	876,71	385,74	415-1421
40 - 69	19	831,58	402,87	355-2129
70 - 79	2	1355,00	-	967-1743

10.5.5. Comparison Between Patients and Normal Subjects

10.5.5.1. Accuracy Scores - The Arrows Test

Comparisons were made between the patients and the normal subjects according to the differences found in the normal age groups i.e. twenty to seventy-nine years. As Levine's test for variances between the two groups was significant (P= 0,0000), the Welch Test was used to test for significant differences between the normal and the patient group (Table 10.20).

TABLE 10.20 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - AR

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20-79	0,0000	0,0040 **

** P<0,01

As can be seen there was a significant difference between the performance of the normal group of individuals and the patient group on the Arrows test (P<0,01).

TABLE 10.21 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - D-L

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20 - 59	0,0000 ***	0,0004 ***
60 - 79	0,0469 *	0,0153 *

* P <0,05 *** P <0,001

10.5.5.2. Accuracy Scores - Dot-Line Test

Comparisons between the normal group of individuals and the patient group were made using the statistical procedures described above but with the ages divided as per the differences in performance found in the normal group i.e. twenty to fifty-nine years and sixty to seventy-nine years (Table 10.21).

As can be seen from Table 10.21 the differences in the scores obtained by the normal and the patient groups in the younger ages, show a greater significant difference than those in the older age group. This could be due to the poor performance of the older normal individuals on the test.

10.5.5.3. Time Scores - The Arrows Test

The comparisons were made in the age groups among which no statistical difference was found in the times taken by the normal group to complete the test (Table 10.22). As can be seen there was a statistical difference in the time taken by the patients when compared to the normal group.

TABLE 10.22 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - AR

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20 - 59	0,0000	0,0015 **
60 - 79	0,0000	0,0514 *

* P <0,05 ** P <0,01

10.5.5.4. Time Scores - The Dot-Line Test

A comparison of the time scores were made as described in section 10.5.5.3 (see Table 10.23).

TABLE 10.23 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - D-L

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20 - 39	0,0000	0,0293 *
40 - 69	0,0010	0,0132 **
70 - 79 years	-	-

* P <0,05 ** P <0,01

The scores obtained by the normal individuals and those obtained by the patients in the seventy year old group were not compared, due to the small number of patients in this group.

10.5.6. Analysis Of Tests Executed By Patients

In the analysis of the percentile equivalent scores obtained by the patients the following was found:

10.5.6.1. The Arrows Test

a) Of the seven patients tested in the twenty to thirty-nine year old group, four obtained a score above the fifteenth percentile and one was discharged. Two patients appeared to have a problem with the directional skills needed in this test as they scored below the fifteenth percentile cut off for normal function.

b) Of the ten patients in the forty to fifty-nine year old group, five scored above the fifteenth percentile, one was discharged, four were thought to have problems as their scores fell below the fifteenth percentile.

c) In the sixty-nine year old group, five patients received scores that could be considered within normal range, three received scores that were below the fifteenth percentile and therefore were considered to have a directional problem. One was untestable on this test and one was discharged prior to completing the test.

d) Of the six patients in the seventy to seventy-nine year old group, four were untestable, one received a score below the fifteenth percentile and one was discharged prior to completing the test.

10.5.6.2. The Dot-Line Test

a) In the twenty to thirty-nine year old group all of the seven patients failed to obtain a score above the fifteenth percentile.

b) Of the ten patients in the forty to fifty-nine age group five patients scored below the fifteenth percentile and one was discharged prior to testing.

c) In the sixty to sixty-nine year old group, the six patients who completed the test obtained scores below the fifteenth percentile.

d) In the seventy to seventy-nine year group of six patients, only two patients were able to be tested. Of these, one performed below the fifteenth percentile. The patient that performed within normal limits was discharged prior to completing the test.

It is clear from the above that the patients found the drawing test more difficult than the motor-free Arrows test.

10.5.7. Problems That May Be Diagnosed From Performance On The Arrows Test And The Dot-Line Test

10.5.7.1. The Arrows Test

a) Left Neglect of the Visual Field

This problem may be diagnosed in a test of this nature if the patient selects exclusively responses from the right side of the response sheet and only selects the stimulus cards placed on the right side (patients fifteen and twenty-seven).

b) Directionality

The test was specifically constructed to evaluate this problem. Directional confusion with the arrows was observed in patient fifteen. This diagnosis was confirmed with his results on the VMI and the Dot-Line test (see Appendix 3). Several patients twisted the stimulus cards around in an attempt to fit the direction of the arrows to an incorrect selection on the response sheet, (patients fifteen and twenty-seven).

c) Visual Discrimination

Poor visual discriminative ability or simultanagnosia could be diagnosed from the response of patient sixteen who stated that he could not see parts of the response sheet. There was also a patient who demonstrated no pattern in his incorrect choices as they were from both sides of the response sheet (patient fifteen). This could have been due to simultanagnosia.

10.5.7.2. The Dot-Line Test

a) Directionality

The Dot-Line test was designed to evaluate this skill. A problem in directional orientation may be observed in the designs completed by patient number three (see Figure 10.8). Because the dots in the test act as a guide, it became clear to the candidate that the Dot-Line test is also of value in evaluating perceptual problems other than directionality. Items in the VMI test (see 6.5.5.2) seemed more successful at discerning directional contusions in a motor test, however, each item of the VMI should be scrutinised to provide information on directional problems as the score will not necessarily reflect the problem.

b) Spatial Relations

Frostig (1966) called her dot-to-dot test a test of spatial relations as the subject is required to evaluate the relationship of the line to the various dots and to reproduce this relationship. The same skill is needed in items one to eight of the dot-line test. Spatial relationship problems such as neglect of part of the design was observed in patient number two (Figure 10.9), however this neglect could also have been the result of poor concentration or simultanagnosia. Drawing in additional lines onto the design (patient number nineteen, Figure 10.10), could be spatial in nature or the result of impulsivity. Using the model as part of the design as seen in patient number nine (Figure 10.11) is said to be a good indication of a spatial problem (Critchley 1953). From the above, it appeared that spatial relationship problems should be verified by the patients' performance in other tests.

c) Space Visualisation

Items nine to sixteen of the Dot-Line test could evaluate space visualisation in conjunction with motor planning as the guideline grid for the patient to complete his design has a different number of dots from that of the model. Problems in this area were observed in the test completed by patient number twenty (Figure 10.12). In this case the dots were restructured to meet the patient's concept of the design. Problems with spatial judgement were seen in patient number six (Figure 10.13). Rotation of the design was observed in the test completed by patient number seven (Figure 10.14).

d) Visual Neglect

The drawing of patient nine demonstrated neglect of the left visual field (Figure 10.15). Neglecting to draw part of the design was commonly seen in the designs reproduced by the patient group. Patient number two, neglected parts of the design (Figure 10.16). This could indicate a simultanagnosia as the problem was not linked to the loss of the left visual field only and included incorrect linking of lines.

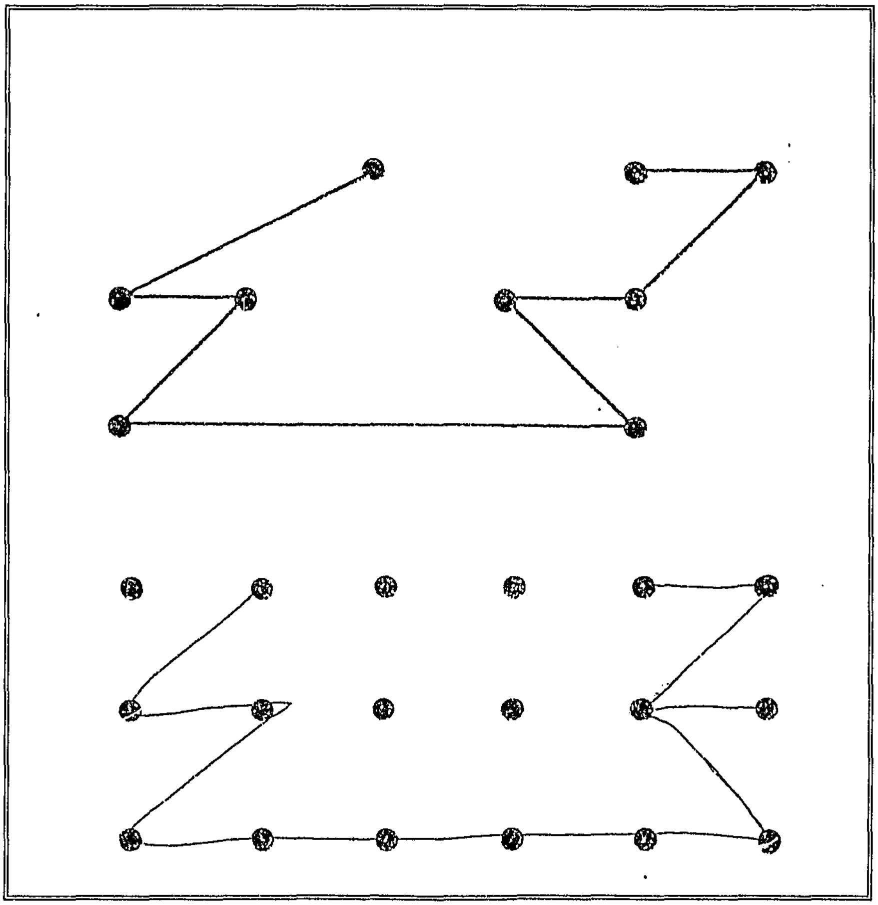


FIGURE 10.8 DIRECTIONAL CONFUSION (PATIENT NO. 3)

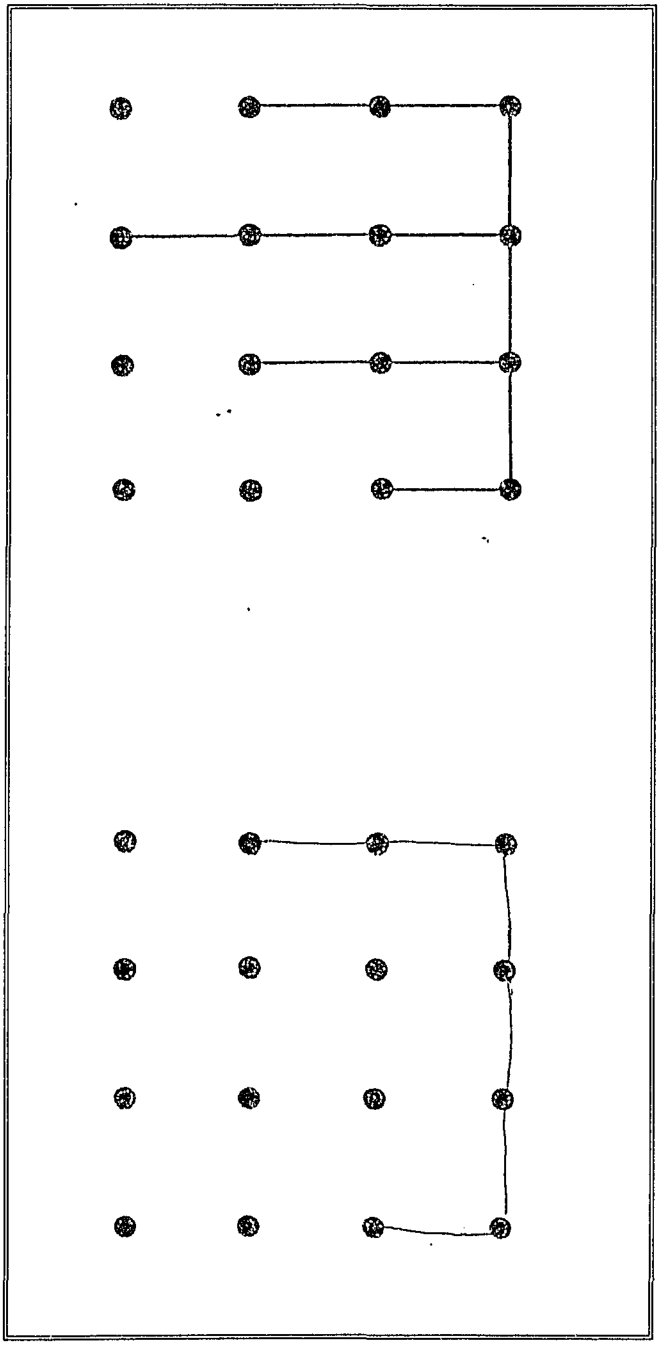


FIGURE 10.9 NEGLECT OF PART OF DESIGN (PATIENT NO. 2)

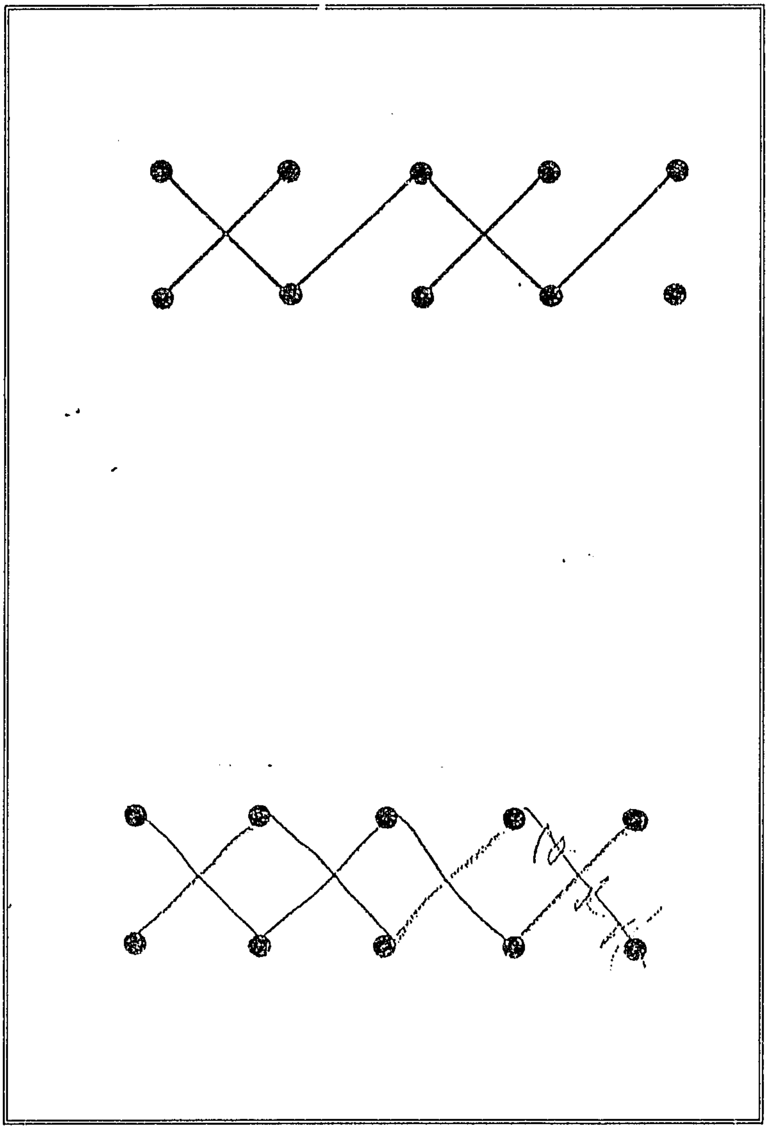


FIGURE 10.10 DRAWING IN ADDITIONAL LINES (PATIENT NO. 19)

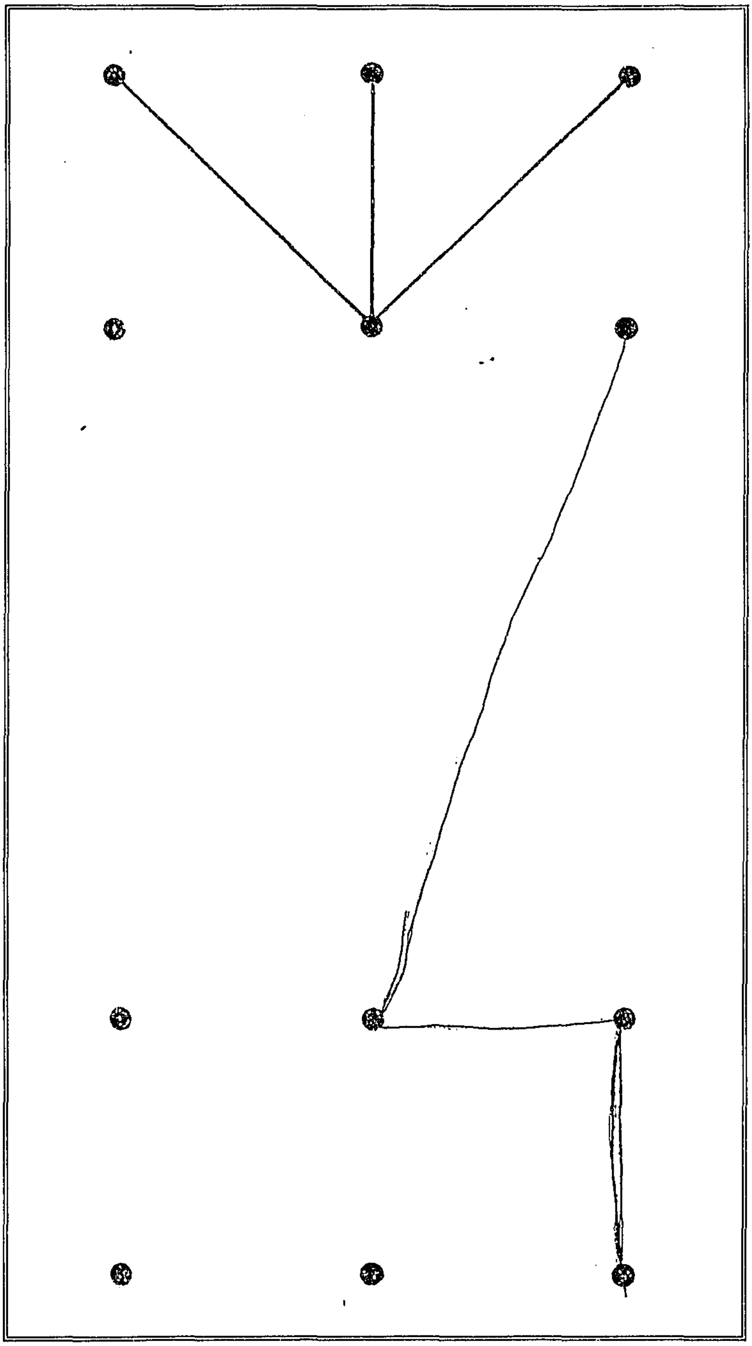


FIGURE 10.11 USING MODEL AS PART OF DESIGN (PATIENT NO. 9)

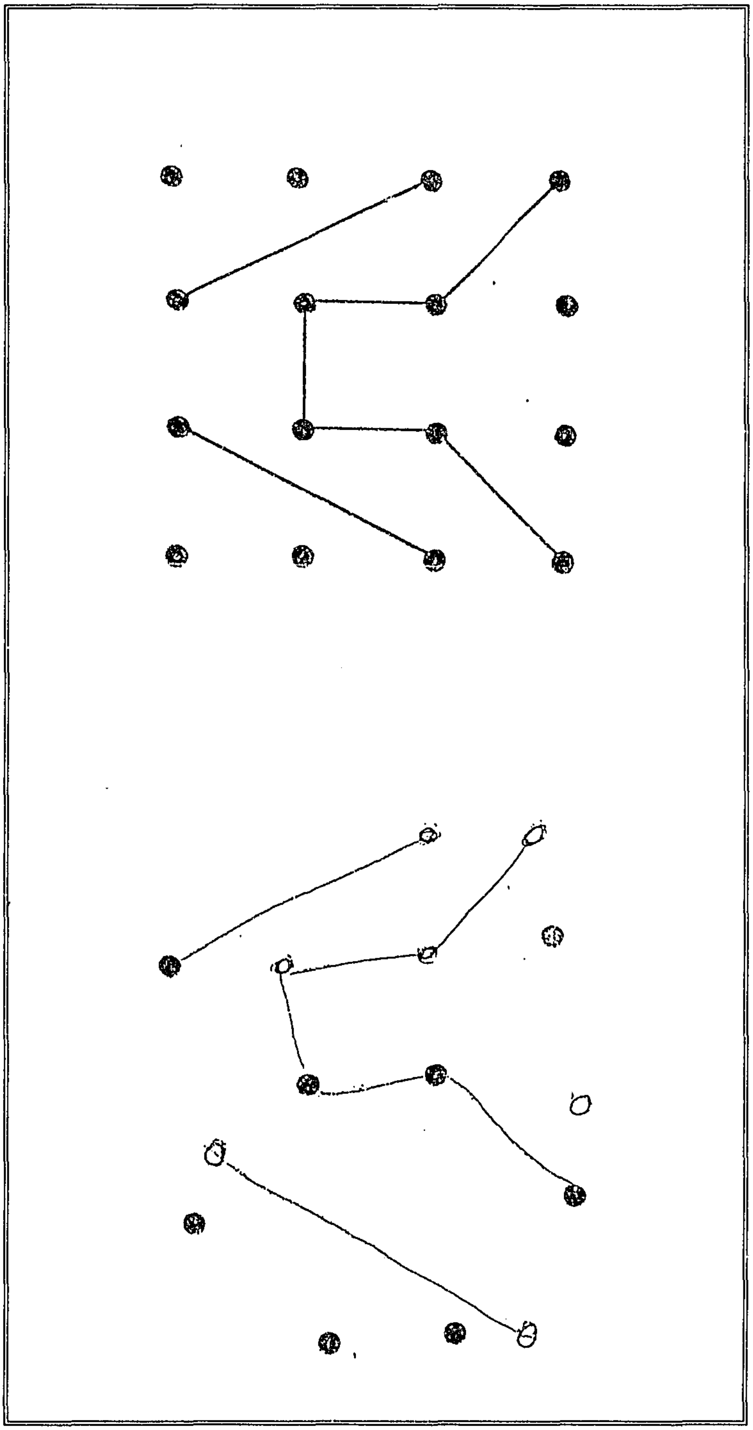


FIGURE 10.12 DOTS MISPLACED (PATIENT NO. 12)

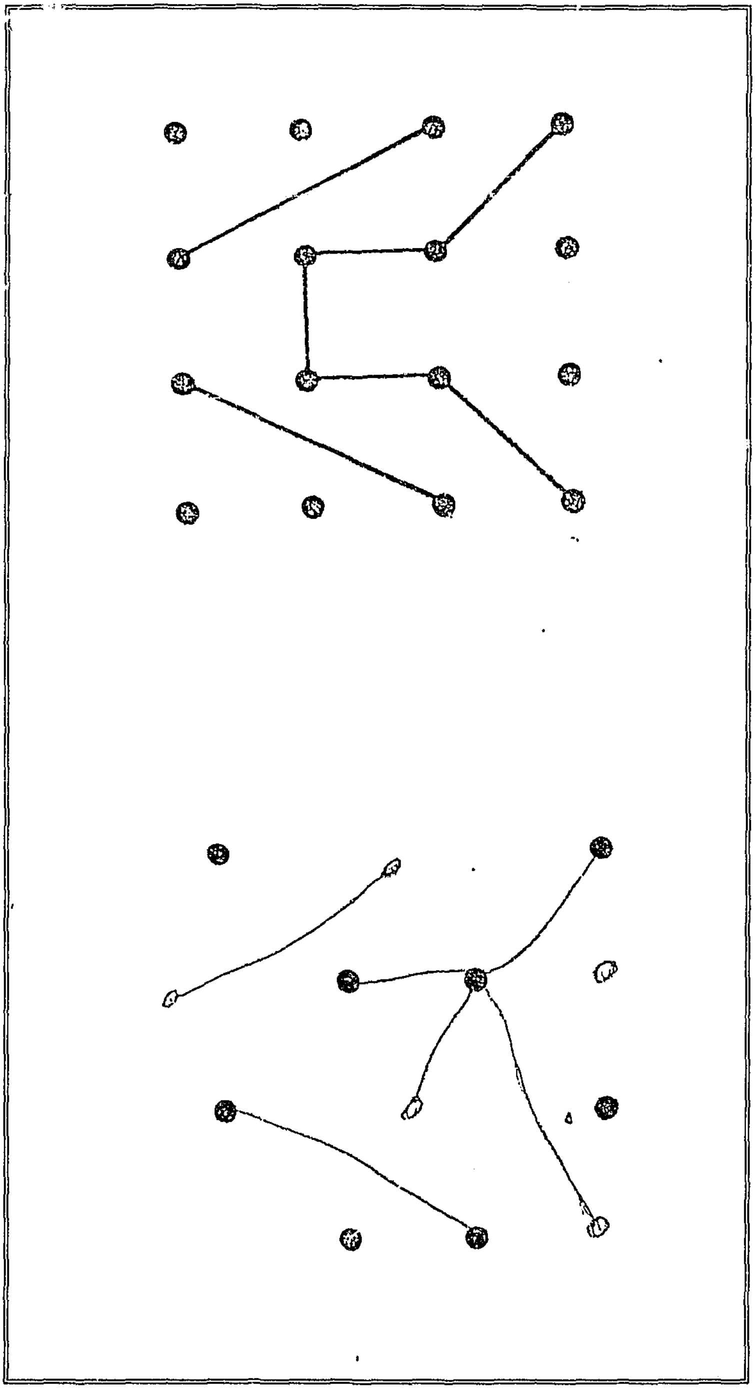


FIGURE 10.13 PROBLEM WITH SPACE JUDGEMENT (PATIENT NO. 6)

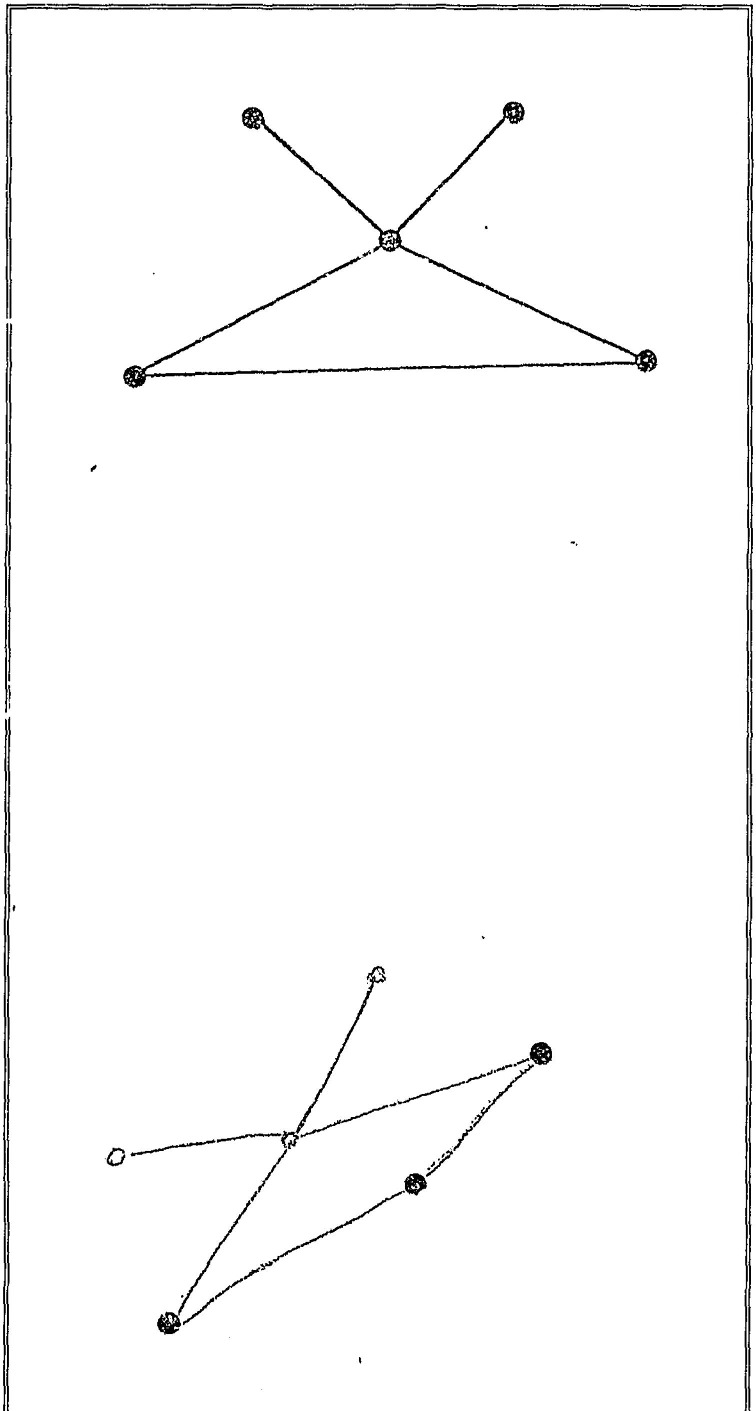


FIGURE 10.14 SPATIAL ROTATION (PATIENT NO. 7)

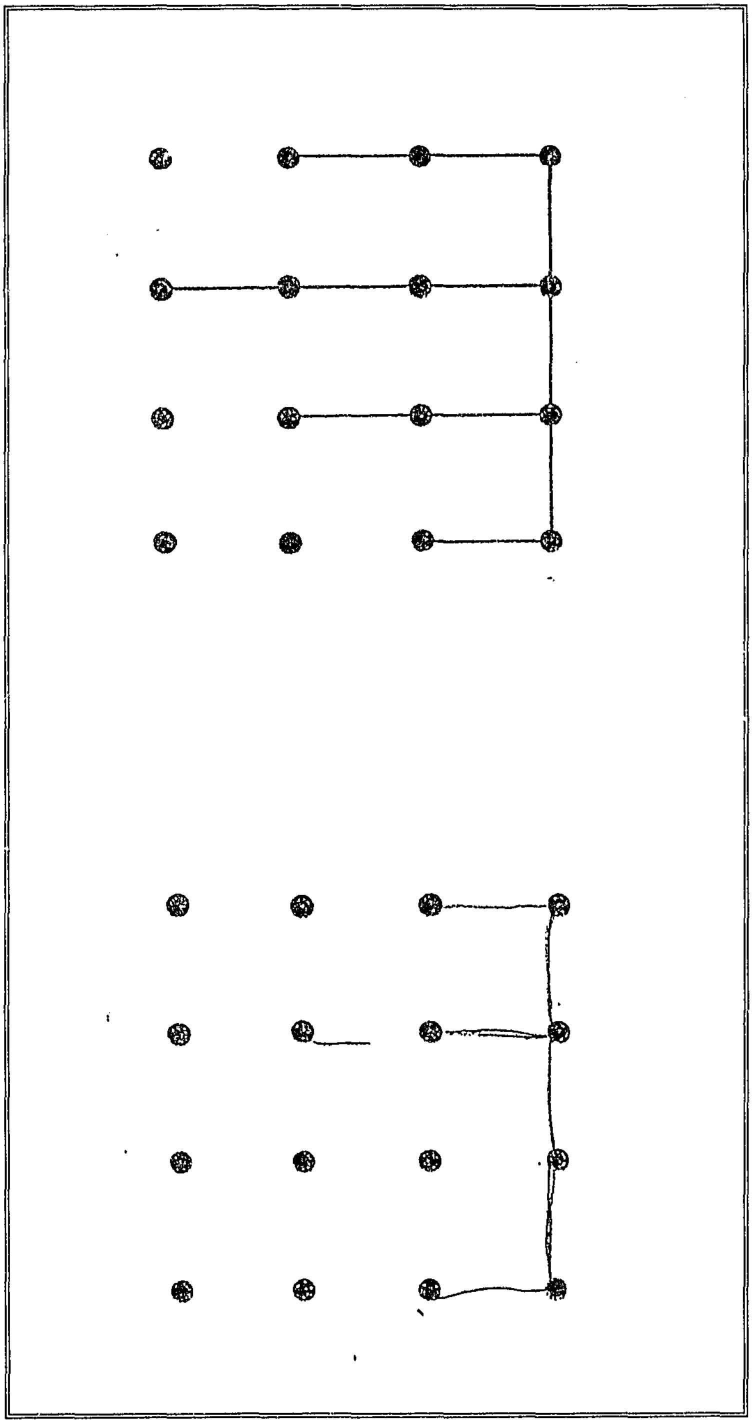


FIGURE 10.15 LEFT NEGLECT (PATIENT NO. 9)

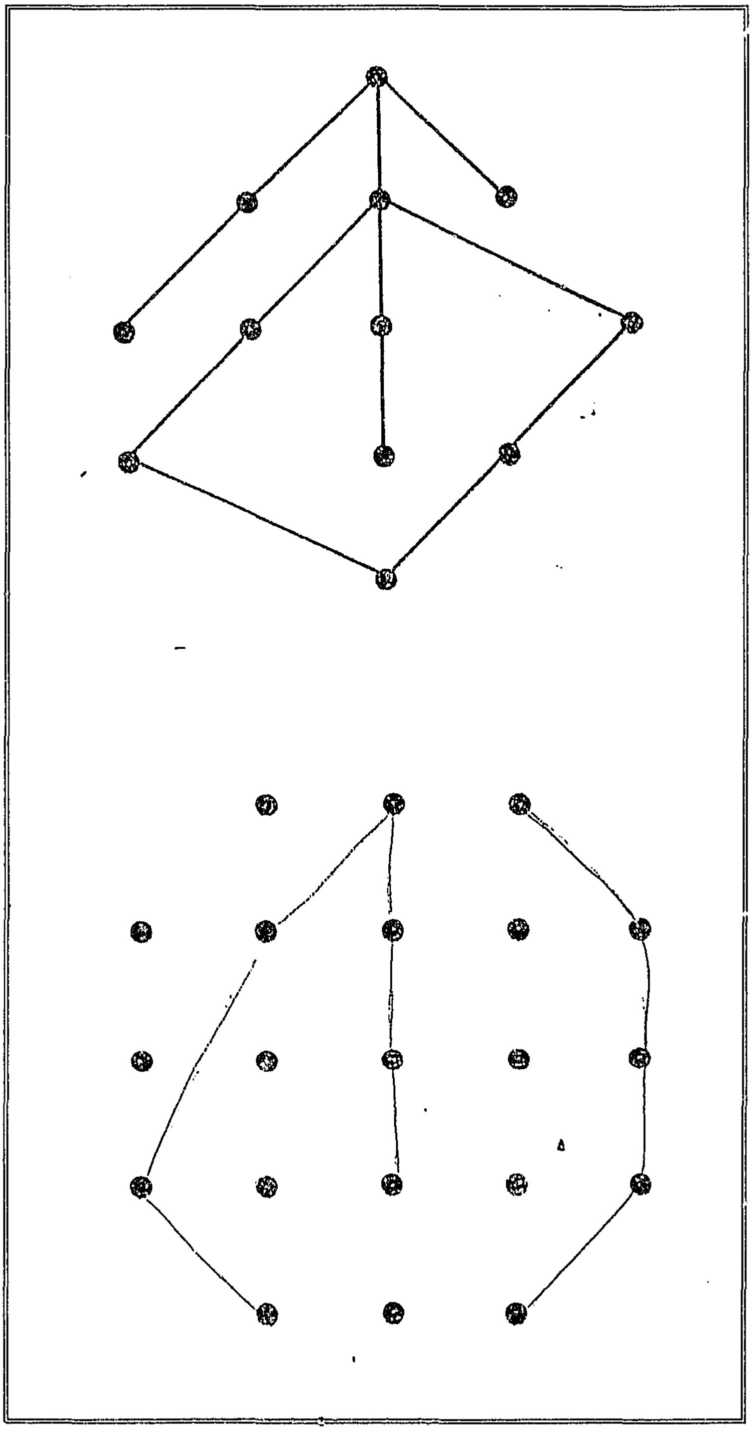


FIGURE 10.16 NEGLECT OF PART OF A DESIGN (PATIENT 2)

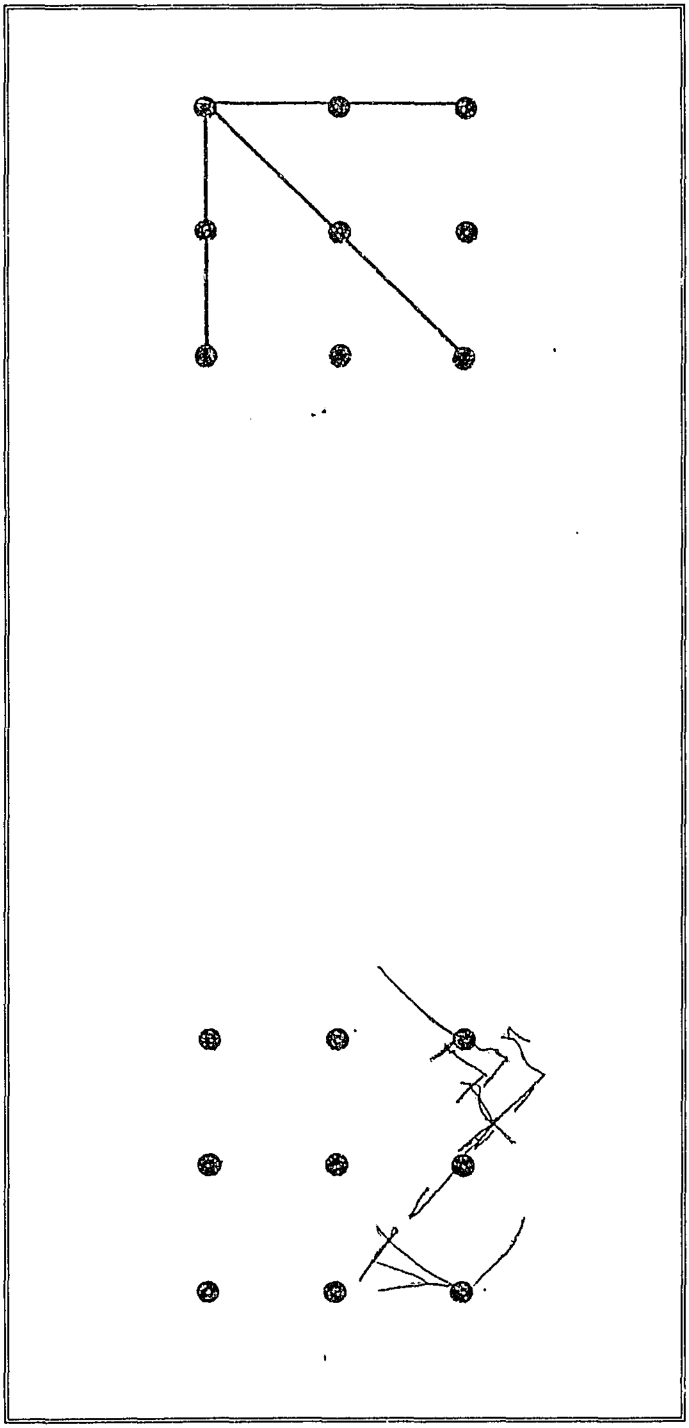


FIGURE 10.17 INABILITY TO DRAW (PATIENT NO.27)

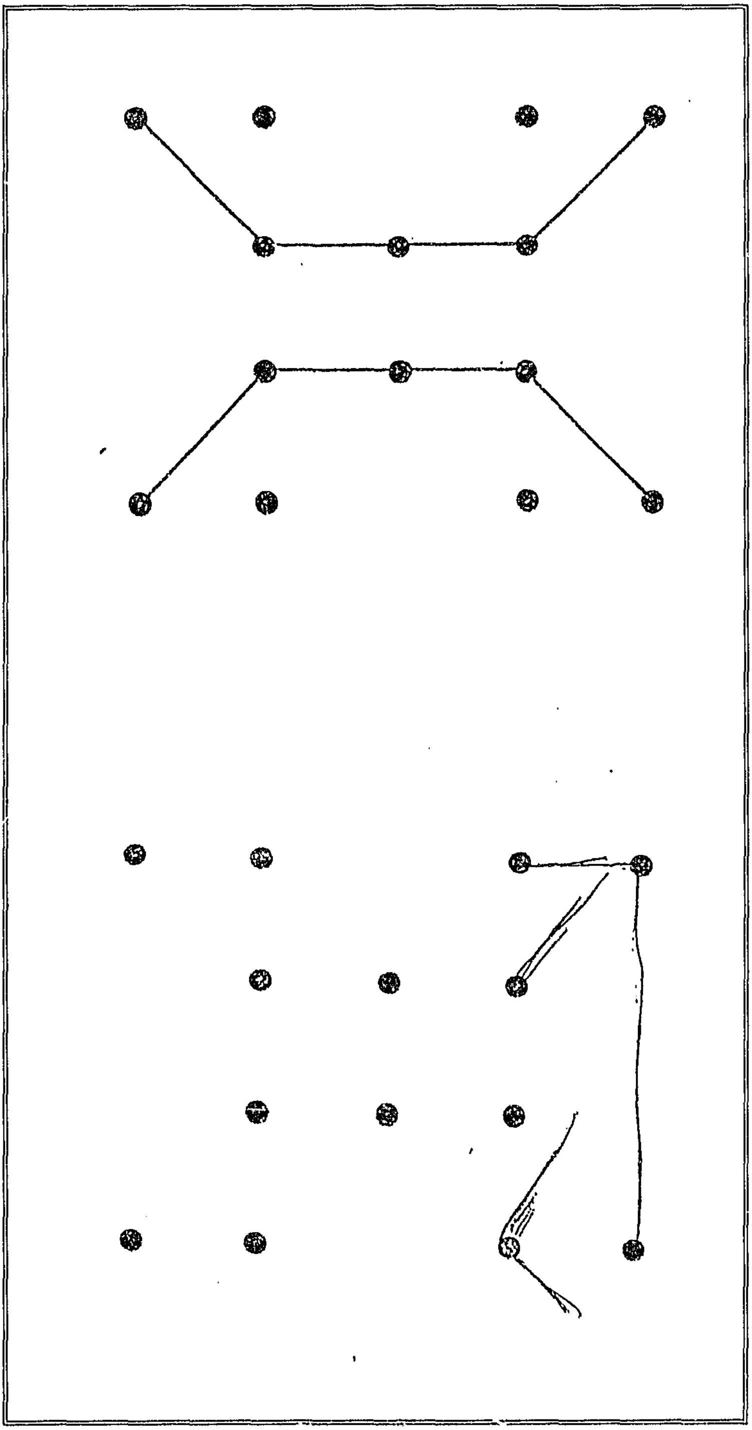


FIGURE 10.18 SPATIAL PROBLEMS AND PERSEVERATION (PATIENT NO. 9)

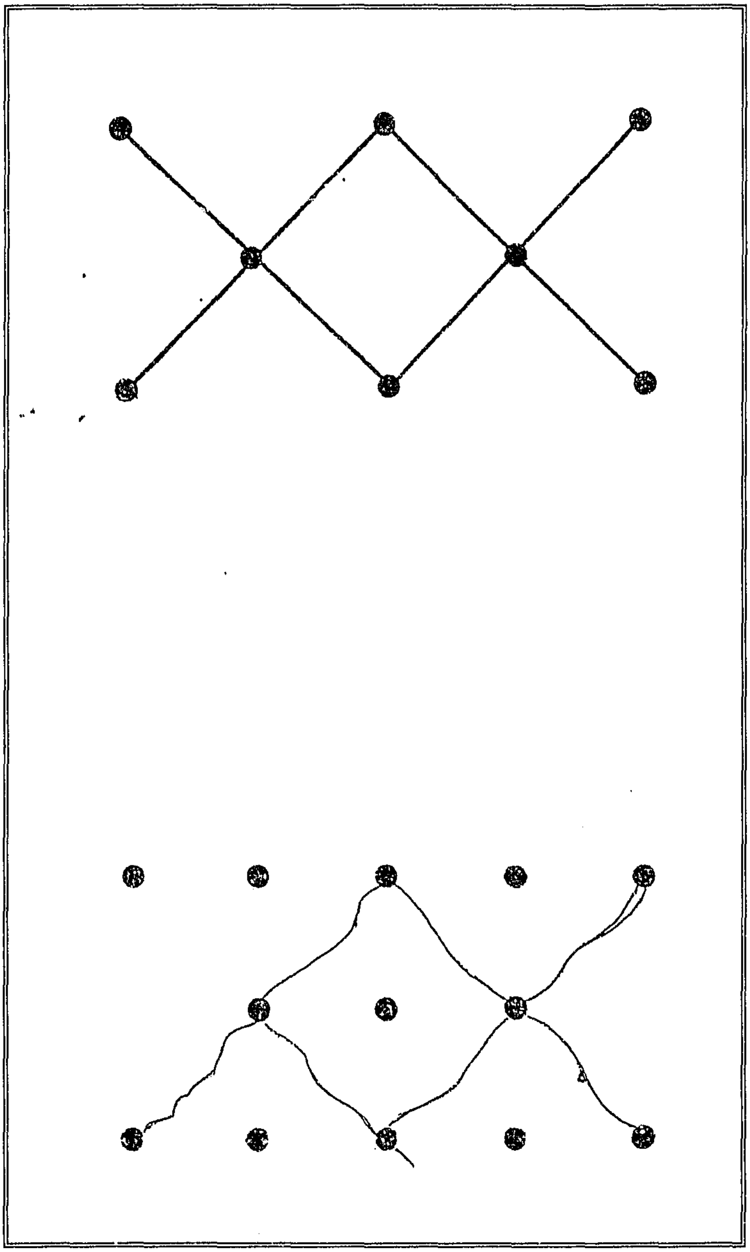


FIGURE 10.19 WAVY LINES (PATIENT NO. 19)

a) Apraxia

Some of the indications of a drawing apraxia could be seen in the drawings of the patients ranging from the almost complete inability to draw (patient number twenty-seven, Figure 10.17) to the perseveration and severe spatial problems of patient number nine (Figure 10.18). The mild problems of left spatial neglect and drawing with wavy lines seen in patient number sixteen (Figure 10.19) could also be symptoms of a mild drawing apraxia.

10.6. ITEM ANALYSIS

10.6.1. Difficulty Index

The difficulty index was calculated using the responses from the normal group of individuals in the age groups among which no statistical difference was found for the accuracy scores. In respect of the Arrows test - all age groups were combined (Table 10.24) and for the Dot-Line test, the ages between twenty to fifty-nine years and sixty to seventy-nine years were used for the calculations (see Table 10.25).

Items 1-6 of the Arrows test were not recorded in Table 10.24 as they were not administered to the normal group. These simple items were constructed for patients who found items seven to twenty-four too difficult. It was found unnecessary to administer these items to the normal group but as they were included in the overall score, each individual was credited with an additional six points. All the items of the test demonstrated a high difficulty p and were therefore fairly easy for normal subjects.

TABLE 10.24 DIFFICULTY INDEX - AR

ITEM	DIFFICULTY p	ITEM NO	DIFFICULTY p
7	0,98	16	0,97
8	0,96	17	0,96
9	0,94	18	0,96
10	0,88	19	0,77
11	0,97	20	0,97
12	0,98	21	0,95
13	0,98	22	1,00
14	0,97	23	0,95
15	0,98	24	0,95

In the Dot-Line test the difficulty index was compiled for the different age clusters. The first seven items of the Dot-Line test are extremely easy. However item eight is the most difficult item in the test. This indicated that more complicated designs would be more likely to reveal perceptual problems. The items which contain more dots in the response item than in the model (items nine to twelve) are marginally less difficult than those that contain less dots (items thirteen to sixteen).

The difficulty index for the responses obtained from the patient group was also calculated (Tables 10.26 for the Arrows test and 10.27 for the Dot-Line test). The Pearson Chi-square Test was used to calculate whether there was a statistical difference between the normal and the patient group in the responses obtained on each item. The probability levels given in the tables, therefore, refer to the probable differences between the patients and the normal individuals on each item of this test.

TABLE 10.25 DIFFICULTY INDEX - D-L

ITEM	DIFFICULTY p 20 - 59 YRS	DIFFICULTY p 60 - 79 YRS
1	1,00	1,00
2	1,00	0,97
3	1,00	1,00
4	0,98	0,92
5	1,00	0,97
6	1,00	1,00
7	1,00	0,95
8	0,72	0,34
9	0,99	0,97
10	0,77	0,50
11	0,93	0,63
12	0,89	0,79
13	0,96	0,71
14	0,70	0,61
15	0,77	0,63
16	0,86	0,68

Those items that revealed no statistical difference between the performance of patients and that of normal individuals should be eliminated from any future version of the Arrows test i.e. items nine, eleven, twelve, seventeen, nineteen and twenty-one. The first six items which were used for patients only can also be eliminated as the difficulty p for these items was 1,00. Only one patient needed to be tested with these items (patient number two). The Arrows test could, therefore, be reduced to twelve items.

TABLE 10.26 DIFFICULTY INDEX - PATIENT GROUP - AR

ITEM	DIFFICULTY p 20 - 79 YRS	ITEM	DIFFICULTY p 20 - 79 YRS
7	0,75 ***	16	0,83 *
8	0,67 ***	17	0,88
9	0,88	18	0,67 ***
10	0,67 **	19	0,67
11	0,92	20	0,71 ***
12	1,00	21	0,88
13	0,83 **	22	0,88 ***
14	0,83 **	23	0,79 **
15	0,75 ***	24	0,71 ***

* P<0,05 ** P<0,01 *** P <0,001

It is interesting to note that, although item eight in the Dot-Line test is the most difficult item on the test there was no statistical difference between the performance of the normal older person and that of the patients on that item (see Table 10.26). It is also apparent that items three to eight, although easy for normal individuals, are still valuable in determining perceptual problems. As there was no difference between the performance of the patient group and that of the normal group on items one, two and thirteen they should be eliminated from the Dot-Line test. The remaining items should be re-ordered according to their difficulty levels.

TABLE 10.27 DIFFICULTY INDEX - PATIENT GROUP - D-L

ITEM	DIFFICULTY p 20-59 YRS	DIFFICULTY p 60-79 YRS
1	1,00	1,00
2	1,00	0,83
3	0,88 **	0,83 **
4	0,82 **	0,67 *
5	0,94 *	0,75 *
6	0,88 **	0,83 **
7	0,94 *	0,83
8	0,31 **	0,25
9	0,81 **	0,58 ***
10	0,44 **	0,33
11	0,50 ***	0,33
12	0,44 ***	0,50 *
13	,88	,42
14	,31 **	,17 **
15	,38 **	,33
16	,38 ***	,33 *

* P <0,05 ** P <0,01 *** P<0,001

10.6.2. Item Discrimination

This is the extent to which an item discriminates between individuals who do well on the test and those who do not. Two ways of determining the discriminatory level of each item were used.

10.6.2.1. Testing the Difference between Two Proportions

The responses obtained by those individuals whose overall score was in the upper 27% of answers were compared with those in the lowest 27%. Fisher's Exact test was used to make the comparison.

The analysis was carried out on the Arrows test although the high difficulty p of the items indicated that it was doubtful that the test items would discriminate between individuals who did well on the test and those who did not and only items ten (P <0,001), nineteen (P <0,001) and twenty-one (P <0,05) proved to be discriminatory.

The discriminatory levels for the items of the Dot-Line test are shown in Table 10.28.

It would appear that items eight to sixteen discriminate between the subjects who do well on the Dot-Line test and those who do not and should, therefore, be retained in the final test format.

TABLE 10.28 ITEM DISCRIMINATION - D-L

ITEM	20-59 YEARS	60-79 YEARS
1-7	-	-
8	-	0,0031 **
9	0,0005 ***	0,0000 ***
10	0,0000 ***	-
11	-	0,0007 ***
12	-	0,0867
13	-	0,0001 ***
14	0,0000 ***	0,0573
15	0,0002 ***	0,0001 ***
16	0,0013 **	0,0007 ***

* P <0,05 ** P <0,01 *** P <0,001

10.6.2.2. The Correlational Approach

This approach to item discrimination reflects the relationships of the individual responses to the total test score. As the difficulty p was found to be high for most of 'he items of the Arrow test, with only three items being found to be discriminatory in testing the difference between the two proportions, the correlational approach was not ..ed with this test.

The analysis of the items on the Arrows test showed that it was not a very difficult test for normal individuals. Although items nineteen and twenty-one did not discriminate between the performance of normal individuals and patients, they were among the three items that discriminated between the normal individuals who did well on the test and those who did not. There was no satisfactory explanation for this. As it would be useful to shorten the test only the following items should be included: seven, eight, ten, thirteen, fourteen, fifteen, sixteen, eighteen, twenty, twenty-two, twenty-three and twenty-four.

The results of the correlational approach to item discrimination on the Dot-Line test are seen in Table 10.29. Items one to seven did not discriminate at all as the difficulty p was high, nor did items nine, eleven and thirteen for the younger age group, and item nine for the older age group. All other items showed highly significant correlations of the item score with the total test score.

The analysis of the items of the Dot-Line test indicated that the items in the test in which the number of guideline dots present in the test figure were different to those in the model (i.e. items nineteen to sixteen) were the more successful at discriminating performance and therefore should be retained in the final test form. This test should therefore consist of eight items only.

TABLE 10.29 BI-SERIAL CO-EFFICIENT OF CORRELATION BETWEEN ITEM AND TOTAL SCORE - DL

	20-59 YRS			60-79 YRS		
ITEM	UPPER 27%	LOWER 27%	r VALUE	UPPER 27%	LOWER 27%	r VALUE
1-7, 9	-	-	-	-	-	-
8	1,0	0,52	0,70***	0,70	0,00	0,81***
10	1,0	0,33	0,81***	1,0	0,00	0,99***
11	-	-	-	1,0	0,20	0,85***
12	1,0	0,81	0,51***	1,0	0,60	0,69***
13	-	-	-	1,0	0,10	1,00***
14	1,0	0,38	0,78***	0,90	0,40	0,57***
15	1,0	0,48	0,72***	1,0	0,10	1,00***
16	1,0	0,57	0,69***	1,0	0,20	0,85***

*** P <0,001

10.7. TEST RELIABILITY

10.7.1. Split-Half Reliability

The co-efficient of correlation between the two halves of the test for all ages combined for the Arrows test was r = 0,51. In evaluating the split-half reliability for each age group in the Dot-Line test there was a low degree of correlation between the two halves of the test for the age groups twenty to fifty-nine years (r = 0,37) and a high degree of correlation between the two halves for the age group sixty to seventy-nine years (r = 0,83) (see Table 10.30).

10.7.2. Kuder-Richardson Formula 20

In this method the pass fail score of each item was compared with the total test score. The Kuder-Richardson Formula 20 correlation for the items of the Arrows test (all age groups combined) was r=0,40.

The co-efficient of correlation for the younger age group for the <u>Dot-Line test</u> i.e. the twenty to fifty-nine year olds was again lower (r = 0,48) than that for the older age group i.e. the sixty to seventy-nine years (r = 0,80).

A summary of the reliability co-efficients can be found in Table 10.30.

<u>TABLE 10.30 RELIABILITY CO-EFFICIENTS - AR & D-L</u>

TEST	AGE	SPLIT-HALF	KRUDER-RICHARDSON
Arrows	20-79	0,51 ***	0,40 ***
Dot-Lline	20-59	0,37 **	0,48 ***
Dot-Lline	60-79	0,83 ***	0,80 ***

** P <0,01 *** P <0,0001

The co-efficient of correlation, although low for the Arrows test showed statistical significance at P <0,001. The Dot-Line test proved to be most reliable for the older age group although all relationships were significant. Reliability co-efficients of r = 0,80 and above are traditionally accepted as evidence of adequate test reliability, therefore, reliabilities of r = 0,37 and r = 0,48, although statistically significant do not represent adequate internal consistency for the test.

10.7.3. Inter-Rater Reliability

The Friedman two-way analysis of variance showed that there was no statistical difference between either the accuracy (P=0,8825 Arrows test, P=0,7470 Dot-Line test) or time scores (P=0,8465 Arrows test and P=0,1146 Dot-Line test) obtained by three examiners on each of six individuals tested. Both tests could, therefore, be said to have inter-rater reliability.

10.8. TEST VALIDITY

Two types of validity are discussed:-

10.8.1. Content Validity

Content validity was provided by fifteen occupational therapists (group one) who rated the content of the tests, according to the definitions given, on a 0-5 scale (see Table 10.31).

As can be seen from Table 10.31 the Arrows test was assessed to be a very much better than average measure of directionality. The Dot-Line test scored slightly higher on spatial relations than it did on directionality. The candidate is of the opinion that the label of spatial relations that Frostig (1966) attached to a similar test could have influenced the ratings given by the therapists who were all familiar with the Developmental Test of Visual Perception (Frostig 1966).

However, it was felt that both tests were adequate measures of the skill that they were designed to measure. As one requires motor planning and the other does not, the results of performance on both tests can be compared in order to make a more accurate diagnosis.

TABLE 10.31 CONTENT VALIDITY - AR AND D-L

PERCEPTUAL SKILL	ARROWS Mean Rating	DOT-LINE Mean Rating
Form as Mass	0,8	0,7
Overall Contour	0,70	2,1
Sub-Units of Form	0,5	1,9
Perceptual Constancy	1.0	0,7
Directionality	3,8 ***	3,4 **
Space Relations	2,0	3,7 ***
Position in Space	1,70	0,8
Figure-Ground	1,7	2,3
Colour	0,0	0,0
Praxis	not required	essential

** a better than average assessment *** a very much better than average assessment

10.8.2. Construct Validity

Three types of construct validity were considered i.e. age differentiation, correlation with other tests in the battery and the relationship of score to diagnosis in the analysis of patient results.

10.8.2.1. Age Differentiation

Discussion already presented indicated that some percentual skills deteriorate with age. Some tests are sensitive to this age differential and others are not. In the case of the directional tests only the Dot-Line test showed a decrease in the ability to do the test from approximately sixty years of age. The Arrows test is therefore not a valid measure of age related changes whereas the Dot-Line test is a valid means of measuring these changes.

10.8.2.2. Correlation with Other Tests in the Battery

As can bee seen from Table 10.32 these two directional tests have a fairly low correlation with each other. This may be due to the motor component in the Dot-Line test. The highest correlation of the Arrows test is with the Space Relations test (r = 0,48).

The Dot-Line test correlated most highly with the VMI test (r = 0,71) as well as with the Design Construction II test (r = 0,63) the Stick Construction II (r = 0,51) and the Position in Space test (r = 0,49). In fact it correlated with all the tests that measure space concepts, having a greater correlation with those tests that required a motor output than those that did not.

Further analysis was carried out on the tests in the battery to determine whether the tests had related factors. The Arrows test had a high loading for Factor II (0,75) and no loading for Factor I or III. As previously discussed, Factor II appeared to be one of figure-ground and visual discriminative ability, it seemed that, in spite of the therapists rating, the Arrow test might be primarily a test of figure-ground ability containing elements of directional sense.

TABLE 10.32 CORRELATION WITH OTHER TESTS IN THE BATTERY - AR AND D-L

TEST	ARROWS TEST	DOT-LINE TEST
Part whole	0,25	0,53
Figure ground	0,33	0,62
Perceptual Constancy	0,31	0,48
Dot-Line	0,35	-
Arrow	-	0,35
Space relations	0,48	0,44
Position in space	0,25	0,49
Block Construction I	0,23	0,46
Block Construction II	0,16	0,36
Stick Construction I	0,14	0,51
Stick Construction II	0,37	0,53
Design Construction	0,32	0,63
Colour	0,00	0,07
VMI	0,27	0,71

The Dot-Line test has a loading for Factor I (0,60), and Factor II (0,52). Factor I appeared to be a spatial factor, measuring in particular space visualisation. Thus the Dot-Line test appeared to have measures of spatial ability and figure-ground ability.

10.8.2.3. Results of Testing Patients

A statistical difference was found between the scores of patients and normal subjects in both the Arrows ($P < 0,01$) and the Dot-Line tests (twenty to fifty-nine years $P < 0,001$

and sixty to seventy-nine years P <0,05). It would seem therefore that both tests are valid measures of V-P-M dysfunction in the adult C.V.A. patient.

10.9. CONCLUSIONS

In summary the following was found on the Arrows test:-

- The test was one of the more difficult to administer as putting the two test cards simultaneously in the right place required dexterity on the part of the examiner. Nevertheless the test did show inter-rater reliability;

- The Arrows test was not sensitive to age related changes although it did demonstrate dysfunction in the C.V.A. patient;

- It was similar to other tests, as it took longer for the older person to complete;

- Item analysis demonstrated that the items were relatively easy. In spite of this, there was a difference between the normal group and the patients on most items. Excluding the first six extremely easy items from the testing of the normal group and using it in the test of only one patient (patient number two) may have affected the analysis;

- Reliability co-efficients although statistically significant were very low;

- Although the content of the test was validated by the therapists as testing directional concepts its highest correlation was with the Spatial Relations test. It did not correlate well with the other test of directional sense i.e. the Dot-Line test which may have been due to the motor component of the latter;

- The test loaded for Factor II which seemed to be a figure-ground/visual discrimination factor. As directionality is part of the space construct, this loading seemed to indicate that the test was inappropriate for the skill that it was constructed to measure. and that the future version of the test should eliminate the figure-ground component;

- The test did not prove to be particularly useful in the diagnosis of perceptual problems that could not be adequately diagnosed by other tests in the battery other than that of visual discrimination.

In view of the above, it is recommended that the Arrows test be altered to eliminate or decrease the figure-ground component. This could be done by changing the multiple choice sheet so that only two alternatives are offered at a time.

The following was found on the Dot-Line test:-

- The test was easy to administer and to score with good inter-rater reliability;

- The test was sensitive to changes in function in the normal person with increasing age;

- The older person took longer to complete the test;

- Differences between the normal group and the patient group were demonstrated. However, in the age group sixty to seventy-nine years, six patients were found to be untestable, thus proving it to be a difficult test for patients;

- Item analysis demonstrated that several of the items were very easy i.e. the items containing an equal number of dots in the model to that in the test item. These items need to be dropped from the test to make it more useful and less time consuming to administer;

- The content was felt by the therapists to be a more appropriate measure of the skill of spatial relations than directionality. In addition the Dot-Line test correlated best with the other drawing test i.e. the VMI test. The Dot-Line test also loads for both Factor I and II;

- The test seemed to be useful for diagnosing directional dysfunction and apraxia, however, the VMI also demonstrated these problems.

It seemed that the Dot-Line test could be a useful alternative to the VMI test when re-evaluation of a patient is needed within a short space of time. The test, however, must be improved through elimination and re-ordering of items.

CHAPTER 11

THE MOTOR REDUCED SPATIAL TESTS - THE SPACE RELATIONS TEST AND THE POSITION IN SPACE TEST

11.1. INTRODUCTION

The determination of co-ordinates in space is a complex task. There are many aspects to spatial function and many types of problems described in the literature. Due to the complexity of spatial functions, the tests developed in this thesis were divided into two groups:-

- tests measuring grasping space or the subjective group of spatial functions;
- and tests measuring action space or the objective group of spatial functions;

The directional component of the subjective group has been described in Chapter Ten. The Spatial Relations test of the V-P-M Battery was devised to measure the spatial relations component of the subjective group together with the Block Construction I, Stick Construction I and the Design Construction tests. The Position in Space test was designed to measure the space visualisation component of the objective group of space constructs together with the Block Construction II and Stick Construction II tests. As it was believed to be important to distinguish between those problems inherent in motor planning and problems not directly involved in the motor planning process, two types of tests were constructed for each problem, i.e., motor reduced tasks and construction tasks (motor enhanced).

A description of the Spatial Relations test is to be found in Chapter Four (4.6.8) and Appendix 1 Section 2.8 while a description of the Position in Space test will be found in Section 4.6.9. of Chapter Four and Appendix 1 Section 2.9.

11.2. REVIEW OF SIMILAR TESTS

Only those tests requiring minimal motor output will be reviewed here. It is interesting to note that few of the tests that measure spatial function are actually called spatial tests. It was only those tests formulated by Frostig

(1966) and Ayres (1972, 1980a) for children that were specifically called space tests until 1991 when the Visual Object and Space Perception Battery (VOSP) (Warrington and James 1991) used the space label.

The Position in Space component of the Developmental Test of Visual Perception (Frostig 1966) is a motor reduced test in which the subject is required to locate the drawing of an object with a different orientation from the other drawings. Although it has been suggested that this test be used with adults, (Siev and Freishtat, 1976) there are no published studies on adults. Siev, Freishtat and Zoltan (1986) suggested that additional research *"is indicated to justify the use of the Frostig Position in Space Test with the adult brain damaged population"* (pg. 78). The Spatial Relations test of the Developmental test of Visual Perception is reviewed in Chapter Ten, as the candidate felt that it was more of a test of directionality than of spatial relations. In addition it is a motor enhanced method of evaluating this skill requiring the subject to copy lines linking dots.

Ayres (1980a) developed two tests, The Space Visualisation test and the Position in Space test as part of the Southern California Sensory Integration Test Battery. Both of these tests are motor free measures of perceptual function. The Space Visualisation test is a three dimensional test consisting of two plastic form boards, one in an egg shape and the other in a diamond shape. Pegs are fitted into the form boards at different locations. The subject is required to indicate from two alternatives the correct form to fit in the form board (Ayres 1980a). The Position in Space test requires the subject to match a combination of forms to its duplicate which is presented in different orientations. This is a two dimensional test. Taylor (1968) used the Space Visualisation test in analysing dysfunction in left hemiplegia following a stroke. She demonstrated a statistically significant (level not given) difference between a normal group of individuals and left hemiplegics on this test but no age related changes were found. The Position in Space test has been used with adults by Harlowe and Van Deusen (1984) and Van Deusen and Harlowe (1987). However, no normative data on adult function have been published by these authors, only the validity studies on the test.

In order to evaluate spatial relations Harlowe and Van Deusen (1984), and Van Deusen and Harlowe (1987), in the St Mary's C.V.A. Evaluation used spatial relations picture cards purchased from Developmental Learning Materials. Each card depicts an object in relation to another and the subject must orally identify the relationship. Again no normative data for adults were provided.

The Location sub-test of the MacQuarrie Test for Mechanical Ability (MacQuarrie 1953) is a test that makes use of visuo-spatial estimation where the subject is required to locate a letter that is placed in the same position as a spot in a block alongside the letter grid. This is similar to the Position Discrimination, Number Location and Dot Counting tests of the VOSP (Warrington and James 1991). The MacQuarrie test provides percentile norms for each sub-test for men and women aged sixteen and up based on the test performance of 1 000 adults of each sex. This test could well be used to measure dysfunction in the adult stroke patient.

The VOSP provides means, standard deviations and five percent cut off scores obtained on one hundred and fifty subjects. Warrington and James (1991) refer to their test which requires the subject to count dots as a test of single point localisation and spatial scanning. In terms of the definition of spatial relations used in this thesis, the Dot Counting test could be classified as a spatial relations test. The Position Discrimination test of the VOSP is described by Warrington and James (1991) as a test of the ability to perceive the relative position of objects in two dimensional space and is therefore directly comparable to the position in space test of the V-P-M battery. In the candidates' opinion, it also requires space visualisation. The Number Location test of the VOSP is based on the MacQuarrie Test. All the VOSP tests are two dimensional in nature. The VOSP was not available at the time the research was carried out for this thesis.

Tests have been used in which the subject is required to indicate which of several rotated figures matches the stimulus figure. Luria's parallelogram test is a good example of this (Christensen 1974) and has been standardised for adults (Golden 1981). The COTNAB (Tyerman et al 1986) evaluates spatial relations in the Angles Sequencing test and provides normative data as previously mentioned. This test requires the subject to judge ever decreasing angles and place the test cards showing the angles in decreasing order.

The Mental Re-orientation test (Ratcliff 1979) is designed to evaluate mental manipulation in space. It consists of little men with a black disk in one hand and a white disk in the other hand. Each man is presented in different orientations and the subject is required to state whether the black disk is on the subject's right or left side. Although there are no normative data for this test the author found that it distinguished between patients with right posterior hemispheric lesions and other patient groups. Other non-standardised tests such as the figures used by Herrmann and van Dyke (1978) are also used to determine mental rotations.

The Visual Discrimination Test (Wepman et al 1975) is essentially a motor free test for assessing position in space i.e. the subject is required to select the form from four alternatives that which matches the one located in the centre of the page. This test is standardised on children five to eight years of age but could be used with adults if normative data were collected.

Hung, Fisher and Cermak (1987) found the Test of Visual Perceptual Skills (TVPS) 1982 version for children of four to twelve years (Gardner 1982) to be a useful test in discriminating between the perceptual performance of normal and learning disabled young men when a raw score was used. However, the components of the TVPS that have relevance to this chapter i.e. the Visual Discrimination and the Visual-Spatial Relations tests showed no significant differences between these two groups. Gardner (1992) extended the TVPS (non motor) by an additional test for ages twelve to eighteen years. This new test has a component which evaluates visual spatial relations in which the subject is required to pick out from five alternatives the design which has a rotated component.

The Motor Free Visual Perceptual Test (Colarusso and Hammill 1972) can also be used with an adult population. Eiman, Eiman and Cermak (1983) demonstrated a difference in the performance between normal males aged forty to sixty years and schizophrenic male patients in the same age group when using this Motor Free Visual Perceptual Test, however, no normative data are given for the adult population.

The Primary Mental Abilities Test (PMA) (Thurstone 1963:3) is a test *"designed to provide both multifactorial and general measures of intelligence."* As an intelligence test, it contains sub-tests that are inappropriate for Occupational Therapists to use, however, the Space Relations sub-test was thought to be an excellent test of space visualisation and has provided ideas for the Position in Space test designed by the candidate. Although Thurstone labelled his test a Space Relations test, the candidate felt that it measured position in space as defined in this thesis. The PMA itself provides quotients for each sub-test for children up to the age of sixteen years and nine months.

A summary of the aforementioned tests and the spatial skills they measure is to be found in Table 11.1.

TABLE 11.1 SUMMARY OF MOTOR REDUCED AND MOTOR FREE VISUAL-SPATIAL TESTS

AUTHOR	TEST	DESIGNED FOR ADULTS	DESIGNED FOR CHILDREN	SUBJECTIVE SPACE	OBJECTIVE SPACE
Ayres	Space Visualisation,	-	√	-	√
	Position in Space	-	√	-	-
Frostig	Position in Space	-	√	-	√
Van Deusen	St Mary's Space Relations	√	-	not known	not known
MacQuarrie	Location Test	√	-	-	√
Luria	Parallelograms	√	-	-	√
Ratcliff	Mental Reorientation	√	-	-	√
Wepman et al	Visual Discrimination	-	√	√	-
Gardner	Visual Perceptual Skills	-	Child and adolescent	√	-
Colarusso et al	Motor Free Visual Perceptual Test	-	√	√	-
Thurstone	Primary Mental Abilities Test	-	√	-	√

11.3. Reasons For Compiling New Tests

The Space Relations test and the Position in Space test were developed because:

a) Few of the tests described above clearly demarcated the skills of spatial relations, position in space or space visualisation.

b) Tests that did measure the functions of spatial relations and position in space were part of other batteries and therefore could not be used separately.

c) Rigorous administration and scoring criteria for adults were not available as many of the tests were constructed for use with children.

d) Some of the tests were not available in South Africa.

The Space Relations test of the V-P-M battery requires the subject to match designs according to the relationship of the parts to each other while the Position in Space test requires the subject to rotate designs in his mind before the correct alternative can be found.

11.4. Results Of Testing Normal Individuals

11.4.1. Accuracy Scores - Space Relations (SR)

The means and the standard deviations were calculated for the normal group as a whole (Table 11.2) and for each group individually (Table 11.3).

TABLE 11.2 MEAN AND SD OF ACCURACY SCORES - SR

n	MAX SCORE	MEAN	SD	RANGE
118	40	37,82	3,02	22 - 40

In examining the means, the standard deviations and the range of scores for each age group, there was a noticeable difference in the scores obtained by the seventy to seventy-nine year age group when compared to other groups. It was not surprising, therefore, that the

Pearson Chi-square Test showed a statistical difference occurring within the range of scores (see Table 11.4) a cut point of thirty-six being used.

TABLE 11.3 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - SR

AGE	n	MEAN	SD	RANGE
20-29	20	38,25	2,05	34-40
30-39	21	38,10	2,86	30-40
40-49	19	38,21	2,20	34-40
50-59	19	38,47	2,37	32-40
60-69	19	38,47	1,95	32-40
70-79	20	35,50	4,76	22-40

TABLE 11.4 COMPARISON OF ACCURACY SCORES - SR

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 36	25,0	28,6	31,6	21,1	10,5	60,0	29,7
Greater than 36	75,0	71,4	68,4	78,9	89,5	40,0	70,4

Pearson Chi-square Value = 13,087 P<0,05

As a significant difference was found at the P<0,05 level of significance further analysis took place to determine where the differences were to be found. The results of this analysis were as follows:-

a) The Pearson Chi-square Test showed that there was no statistically significant difference between the scores obtained by each of the age groups ranging from twenty years to sixty-nine years. These age groups were therefore grouped together for further analysis.

b) Fischer's Exact test showed that there was a highly significant differer ce between the two new age groups thus formed i.e. the twenty to sixty-nine year olds and the seventy to seventy-nine year olds (P<0,01). The "drop off" in function, therefore, occured at approximately seventy years of age.

TABLE 11.5 COMPARISON OF ACCURACY SCORES FOR AGES 20-69 AND 70-79 YEARS - SR

AGE	20-69	70-79	TOTAL
Less than 36	23,5	60,0	29,7
Greater than 36	76,5	40,0	70,3

Fisher's Exact test P = 0,0026 P<0,01

Percentile ranks were, however, calculated for age groups twenty to thirty-nine, forty to fifty-nine, sixty to sixty-nine and seventy to seventy-nine years (see Appendix 2) in line with the other tests in this battery.

11.4.2. Accuracy Scores - Position in Space Test (PIS)

The means and the standard deviations were calculated for the normal group as a whole (Table 11.6) and for each age group individually (Table 11.7).

TABLE 11.6 MEAN AND SD OF ACCURACY SCORES - PIS

n	MAX SCORE	MEAN	SD	RANGE
118	100	56,84	29,27	-1,00 - 98

If one examines the scores presented in Table 11.7, the large range of scores obtained by each age group with the resultant large standard deviations is noticeable. The fact that the raw score was obtained through subtracting the incorrect responses from the correct responses may have contributed to this and accounted for the negative value seen for age group seventy to seventy-nine years.

The Pearson Chi-square Test was used to determine whether there was any difference in the performance in the various age groups (Table 11.8).

TABLE 11.7 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - PIS

AGE	n	MEAN	SD	RANGE OF SCORES
20-29	20	72,25	28,3	3-98
30-39	21	71,38	24,61	8-95
40-49	19	65,21	18,54	23-90
50-59	18	64,17	22,08	17-98
60-69	18	40,94	23,63	4-86
70-79	20	25,95	25,22	-1-93

TABLE 11.8 COMPARISON OF ACCURACY SCORES - PIS

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 63	25,0	23,8	42,1	33,3	83,3	90,0	49,1
Greater than 63	75,0	76,2	57,9	66,7	16,7	10,0	50,9

Pearson Chi-square value = 34,011 P<0,001

As the Pearson Chi-square Test demonstrated that significant differences occurred in the scores, further analysis was carried out to determine where the differences were to be found as follows:-

a) The Pearson Chi-square Test showed that there was no difference in the performance of the twenty to twenty-nine, thirty to thirty-nine, forty to forty-nine and fifty to fifty-nine year olds (P = 0,5741). These age groups were, therefore, combined for further analysis.

b) Fischer's Exact test showed that there was no statistical difference between the performance of the sixty to sixty-nine and seventy to seventy-nine year olds (P = 0,6525). These two age groups were, therefore, combined for further analysis.

c) Fischer's Exact test showed that there was a highly significant difference between the two new age groups thus formed i.e. twenty to fifty-nine years and sixty to seventy-nine years (Table 11.9).

TABLE 11.9 COMPARISON OF ACCURACY SCORES AGES 20 - 59 AND 60 - 79 YEARS - PIS

AGE	20-59	60-79	TOTAL
Less than 63	30,8	86,8	49,1
Greater than 63	69,2	13,2	50,9

Fisher's Exact test P = 0,0000 $P<0,001$

Thus there was a drop off in performance at the sixty year old level on the Position in Space test.

11.4.3. Time Scores

11.4.3.1. Time Scores - Space Relations Test

The means and the standard deviations were calculated for the time scores for the normal group as a whole (Table 11.10) and for each age group separately (Table 11.11).

TABLE 11.10 MEAN AND SD OF TIME SCORES - SR

n	MEAN/SECS	SD	RANGE
118	239,75	115,2	125-750

Levine's test for equal variance established that there was a significant difference in the variances between the groups (P = 0,0000), therefore, the Welch Test was used to establish that the differences in the mean of the time scores obtained by each group was significant (P = 0,0000). The separate variance t was used to establish where the differences were to be found (Table 11.12).

TABLE 11.11 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - SR

AGE	n	MEAN/SECS	SD	RANGE
20-29	20	176,1	23,7	142-236
30-39	21	179,38	43,11	125-292
40-49	19	190,90	28,22	147-267
50-59	19	208,32	36,49	159-282
60-69	19	269,89	74,83	164-416
70-79	20	414,45	164,90	203-750

On the basis of the differences found, the values for different ages were re-grouped as previously described and a similar analysis carried out. Levine's test demonstrated a difference in the variances between the groups ($p = 0,0000$) and the Welch Test, a difference in the scores obtained by the different groups ($p = 0,0000$), the separate variance t with the Bonferroni correction was used to determine what these differences were (Table 11.13).

As can be seen in Table 11.13, differences existed between all the re-grouped data. Percentiles for the different age groups are to be found in Appendix 2.

TABLE 11.12 DIFFERENCES IN TIME SCORES - SR

AGE		SEPARATE VARIANCE t P VALUE
20 - 29	30 - 39	0,7632
20 - 29	40 - 49	0,0857
20 - 29	50 - 59	0,0028 *
20 - 29	60 - 69	0,0000 ***
20 - 29	70 - 79	0,0000 ***
30 - 39	40 - 49	0,3203
30 - 39	50 - 59	0,0272
30 - 39	60 - 69	0,0001 **
30 - 39	70 - 79	0,0000 ***
40 - 49	50 - 59	0,1090
40 - 49	60 - 69	0,0003 **
40 - 49	70 - 79	0,0000 ***
50 - 59	60 - 69	0.0034
50 - 59	70 - 79	0,0000 ***
60 - 69	70 - 79	0,0014 *

* P<0,05 ** P<0,01 *** P<0,001

TABLE 11.13 DIFFERENCES IN RE-GROUPED TIME SCORES - SR

AGE		SEPARATE VARIANCE t P VALUE
20 - 39	40 - 59	0,0056 *
20 - 39	60 - 69	0,0000 ***
20 - 39	70 - 79	0,0000 ***
40 - 59	60 - 69	0,0008 ***
40 - 59	70 - 79	0,0000 ***
60 - 69	70 - 79	0,0014 **

* P<0,05 ** P<0,01 *** P<0,001

11.4.4. Time Scores - Position In Space Test

The Position in Space test was the only test in the V-P-M battery for which a time limit was set for the completion of the test i.e. four minutes for each page. There were therefore no time scores for this test.

11.4.5. Discussion Of The Results Of Testing Normal Individuals

The scores of the Position in Space test indicated that the normal subjects found this test far more difficult than the Space Relations test. Most subjects commented on this difficulty. As there are very few tests available which evaluate space visualisation one could assume that it is difficult to provide adequate evaluation tools.

Both space tests revealed age related changes, the easier test i.e. the Space Relations test, had a drop off in ability to do the test at sixty years of age whereas the Position in Space test showed a drop off at fifty years of age.

This confirmed the findings of Horn and Catell (1966), Hayslip and Stern (1979) and Botwinick (1981). Observations discussed in Chapter Five and Chapter Nine are also relevant here. The time taken to complete the space relations test showed the expected increase in time.

In the case of both the Space Relations test and the Position in Space test, the null hypothesis is rejected in respect of the accuracy scores. The null hypothesis is rejected in respect of the time scores for the Space Relations test

11.5. Results Of Testing Patients

11 5.1. Accuracy Scores - Space Relations Test

The means and the standard deviations were calculated for the age groups between which no statistical difference was found when testing the normal subjects i.e. twenty to sixty-nine years and seventy to seventy-nine years (Table 11.14).

TABLE 11.14 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - SR

AGE	n	MEAN	SD	RANGE
20 - 69	24	27,96	14,02	1 - 40
70 - 79	3	16,75	22,38	5 - 38

In the twenty to sixty-nine year group two patients were discharged prior to completing this test. In the seventy to seventy-nine year group, the testing of two patients was discontinued prior to the administration of this test (one was uncooperative and the other had severe visual neglect). One patient proved to be untestable on all perceptual measures and one was discharged prior to completing this test. Thus the test results of only twenty-seven patients were available for analysis.

11.5.2. Accuracy Scores - Position In Space Test

The means and the standard deviations were calculated for the age groups between which no statistical differences were found when testing the normal subjects i.e. twenty to fifty-nine years and sixty to seventy-nine years (Table 11.15). In the age group twenty to fifty-nine years two patients were discharged prior to completing this test and one patient was found to be untestable. In the sixty to seventy-nine year old group, six patients were found to be untestable and three were discharged prior to completing this test.

TABLE 11.15 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - PIS

AGE	n	MEAN	SD	RANGE
20 - 59	14	33,64	23,22	2 - 80
60 - 79	7	22,71	28,87	4 - 81

11.5.3. Time Scores - Space Relations Test

The means and the standard deviations were calculated for the age groups in which no statistical difference was found when testing the normal subjects i.e. twenty to thirty-nine years, forty to fifty-nine, sixty to sixty-nine and seventy to seventy-nine years (Table 11.16).

TABLE 11.16 MEAN AND SD OF TIME SCORES FOR PATIENTS - SR

AGE	n	MEAN/SECS	SD	RANGE
20 - 39	6	378,17	191,45	172 - 699
40 - 59	9	381,33	139,99	227 - 679
60 - 69	9	571,56	290,17	233 - 985
70 - 79	3	1013,67	60,99	960 - 1000

11.5.4. Comparison Between Patients and Normal Subjects

11.5.4.1. Accuracy Scores - The Space Relations Test

Comparisons were made between the patients and the normal group in the ages twenty to sixty-nine and seventy to seventy-nine. As Levine's test for variances between the groups was found to be significant in both instances (P = 0,0000), the Welch Test was used to test for significant differences between the scores obtained by the normal and patient group (Table 11.17).

As can be seen from Table 11.17, the patients in the younger age category obtained scores significantly lower than the normal group. No differences were found between the patients and the normal group in the age range seventy to seventy-nine. This may be due to the small number of patients tested (n = 3).

TABLE 11.17 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - SR

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20 - 69	0,0000	0,0007 ***
70 - 79	0,0000	0,1939

*** P<0,001

11.5.4.2. Accuracy Scores - The Position in Space Test

Testing for significant differences between the accuracy scores obtained by the normal group and the patient group, was carried out as for the Space Relations test (Table 11.18).

TABLE 11.18 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - PIS

AGE	LEVINE'S TEST P VALUE	F TEST P VALUE
20 - 59	0,8382	0,0000 ***
60 - 79	0,9934	0,3286

*** P<0,001

As in the space relations test, there were no statistical differences in the performance of the elderly group of normal subjects and the patients of the same age. This was surprising for a test of this level of difficulty but may again be due to the small number of patients tested (n=7). Six patients were found to be untestable in this age range which is probably a more accurate reflection of its difficulty level.

11.5.4.3. Time Scores - The Space Relations Test

Comparisons were made between the normal and the patient group within the age categories already established i.e. twenty to thirty-nine, forty to fifty-nine, sixty to sixty-nine and seventy to seventy-nine years. The results are to be seen in Table 11.19. There was no difference between the normal and the patient group at the younger age in the time taken to complete the test.

TABLE 11.19 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - SR

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20 - 39	0,0000	0,0518
40 - 59	0,0000	0,0046 *
60 - 69	0,0000	0,0129 *
70 - 79	0,0670	0,0000 ***

* P<0,05 *** P <0,001

11.5.5. Analysis of Tests Executed by Patients

11.5.5.1. Space Relations Test

a) In the age group twenty to thirty-nine, two of the seven patients performed below the fifteenth percentile and therefore could be said to have a problem with space relations. One patient was discharged prior to completing the test.

b) In the age group forty to fifty-nine, six of the ten patients demonstrated problems as they obtained scores below the fifteenth percentile and one was discharged prior to completeing the test.

c) In the age group sixty to sixty-nine, three patients received scores above the fifteenth percentile, five below and one was discharged prior to completing the test.

d) Of the six patients in the seventy to seventy-nine year group, two obtained scores above the fifteenth percentile, one below and three were untestable.

11.5.5.2. *Position in Space Test*

a) In the age group twenty to thirty-nine years, two patients were found to have problems with position in space as they received scores below the fifteenth percentile and one was discharged without completeing the test.

b) In the age group forty to fifty-nine years, five patients received scores below the fifteenth percentile, one was untestable and one discharged prior to completing the test.

c) Of the ten patients in the sixty to sixty-nine age group, only one received a score within normal limits, one was discharged prior to completing the test, one was untestable and the rest received scores below the fifteenth percentile thereby demonstrating a problem with position in space.

d) In the seventy to seventy-nine year age group, the only patient who completed the test, received a score below the fifteenth percentile and one was discharged.

Although the Space Relations test appeared to be the easier of the two tests, there was no real difference in the number of patients passing this test particularly in the younger age groups, compared to the Position in Space test.

It was also interesting to note that no differences existed in the time scores obtained in the twenty to thirty-nine year old group on the Space Relations test, the patient and the normal group taking equally as long to complete the test. Differences between the forty to fifty-nine year olds and the sixty to sixty-nine year olds were probably significant ($P<0,05$). However, the differences between the patients and the normal group once they had reached seventy was almost certainly significant ($P<0,001$). Caution should be used in interpreting these results due to the small number of patients in the seventy year old group.

11.5.6. Problems That May Be Diagnosed From Performance On The Space Relations Test And The Position In Space Test

The analysis of the performance of patients on a multiple choice test, takes a different form to that of drawing and construction tests, as making a diagnosis is dependent on determining the skill that the test measures and comparing the patients' scores to those of normal individuals of the same age.

However, comparisons can be made between the performance of the individual patient on tests that require motor output and those that do not (See Chapter 16 on Apraxia and Chapter 19 on Spatial Discrimination).

The following types of problems were diagnosed in the execution of the Spatial Relations and the Position in Space tests:-

a) Neglect of the left visual field was seen in the performance of patient twenty-seven in both tests.

b) Poor spatial ability in both aspects of space i.e. position in space and spatial relations (patients two, eight, nine, ten and twenty-two).

c) Poor space visualisation was seen in the performance of patient twenty. This patient performed within normal limits in the space relations test, but poorly in the position in space test.

d) Visual inattention - Patient thirteen appeared to be unable to see all the different designs at one time, therefore, was unable to select the correct design. Other problems such as poor visual memory and poor concentration (patient eleven and twelve) cannot always be dissociated from visual attention and require an experienced tester to make a diagnosis.

e) Perseveration - this was seen in the results of patient nine as a design was repeatedly marked.

11.6. ITEM ANALYSIS

Item analysis was only carried out on the Space Relations test. The Position in Space test is a speed test i.e. the subject had a limited time to complete his response therefore not every item wast completed by every subject. Thus the results had to do with the speed at which the subject worked rather than his overall ability to do the test (Downie and Heath 1974).

11.6.1. The Difficulty Index

The difficulty index was calculated using the responses from the normal group of individuals, in the age groups twenty to sixty-nine and seventy to seventy-nine years (see Table 11.20).

As can be seen from Table 11.20, very little variation occurred between the items, each item demonstrating a high difficulty p, therefore being very easy to complete. The older age group found the first twenty items slightly more difficult than the last twenty. This was to be expected as the last twenty items used more familiar letter and number relationships. Accordingly the items should be re-ordered so that the test begins with items twenty-one to forty.

The difficulty index for the responses obtained from the patient group was also calculated. Fischer's Exact test was used to calculate whether there were statistical differences between the normal and patient group in the responses obtained on each item (See Table 11.21).

TABLE 11.20 DIFFICULTY INDEX - SR

ITEM	DIFFICULTY p 20-69 YRS	DIFFICULTY p 70-79 YRS	ITEM	DIFFICULTY p 20-69 YRS	DIFFICULTY p 70-79 YRS
1	0,97	0,90	21	1,00	1,00
2	0,96	0,80	22	0,95	0,90
3	0,97	0,75	23	0,91	0,90
4	0,99	1,00	24	1,00	1,00
5	1,00	1,00	25	0,95	0,90
6	0,99	0,95	26	0,99	0,90
7	0,99	0,95	27	0,86	0,75
8	0,97	0,95	28	0,99	0,90
9	0,98	1,00	29	0,97	1,00
10	0,97	0,90	30	0,93	0,90
11	0,99	0,90	31	1,00	1,00
12	0,99	0,95	32	0,94	0,90
13	0,99	0,85	33	0,98	1,00
14	0,99	1,00	34	1,00	1,00
15	1,00	1,00	35	1,00	1,00
16	1,00	0,90	36	1,00	1,00
17	1,00	1,00	37	0,99	1,00
18	0,99	0,90	38	1,00	1,00
19	1,00	1,00	39	1,00	1,00
20	0,98	1,00	40	1,00	1,00

The difficulty p for the seventy to seventy-nine year old group was not calculated as scores from only three patients in this group were available. As previously stated three patients in this group were untestable on this component of the test battery. Patients in the older age group appeared to have severe difficulty with spatial type tasks not experienced in the other tests (see discussion in Chapter Eighteen) and the zero scores of the patients who were untestable should probably have been included in the results.

From the difficulty index calculated on the results of testing the normal subjects it appeared that only items one to twenty should be retained in the test and re-ordered in terms of difficulty levels. The remaining items were easy, in addition there were more items which demonstrated a statistical difference between the patient group and the normal group in the first twenty items.

11.6.2. Item Discrimination

This is the extent to which the item discriminates between the subjects who can do well on the test and those who can not. The procedures for establishing the discriminatory levels of each item i.e. testing the significant differences between two proportions and the correlation approach were impossible to determine. This was due to the fact that the difficulty levels were high and cut points for the upper 27% and lower 27% of the overall score could not be easily established.

ABLE 11.21 DIFFICULTY INDEX - PATIENT GROUP - SR

ITEM	DIFFICULTY p 20-69 YRS	ITEM	DIFFICULTY p 20-69 YRS
1	0,92	21	1,00
2	0,76 **	22	0,96
3	0,92	23	0,88
4	0,84 **	24	0,96
5	0,96	25	0,80 *
6	0,84 **	26	0,96
7	0,76 **	27	0,86
8	0,84 *	28	1,00
9	0,88	29	0,96
10	0,80 **	30	0,92
11	0,96	31	0,88 **
12	0,96	32	0,84
13	0,80 **	33	0,84 *
14	0,80 **	34	1,00
15	0,96	35	0,80 ***
16	0,88 **	36	1,00
17	0,84 **	37	0,96
18	0,96	38	0,92 *
19	0,32 *	39	0,88 **
20	0,96	40	0,96

* P < 0,05 ** P < 0,01 *** P < 0,001

11.7. TEST RELIABILITY

11.7.1. Split-Half Reliability

Split-half reliability provided a co-efficient of internal consistency for the test with regard to content ampling. The resultant co-efficient of correlation for the Space Rela ons test was r = 0,43 in the age group twenty to sixty-nine and r = 0.73 for the age group seventy to seventy-nine.

Split-half reliability co-efficients can not be calculated for the Position In Space test as the time limit meant that not every item was answered by all subjects.

11.7.2. Kuder-Richardson Formula 20

In this method of calculating test reliability, the pass-fail score for each item was compared with the total score using the Kuder-Richardson Formula 20. This procedure determined the inter-item consistency of the test.

For the age group twenty to sixty-nine the co-efficient of correlation was r = 0,87 and for the age group seventy to seventy-nine, r = 0.94. Thus reliability co-efficients for the Space Relations test using this method were high. Table 11.22 shows the range of correlation co-efficients which are all highly significant. Only the split-half correlation for the younger group is well below the r = 0,80 considered desirable.

TABLE 11.22 RELIABILITY CO-EFFICIENTS - SR

AGE	SPLIT-HALF	KUDER-RICHARDSON
20-69	0,43 ***	0,87 ***
70-79	0,73 ***	0,94 ***

*** P < 0,001

Again the Kuder-Richardson Formula was not calculated for the Position in Space test for the same reasons as those given in 11.7.1.

11.7.3. Inter-Rater Reliability

The Friedman two-way analysis of variance showed that there was no statistical difference between either the accuracy (p=0,8825 in the Space Relations test and p=0,8465 in the Position in Space test) or time scores (p=0,3529 in the Space Relations test) obtained by three examiners on each of six individuals tested. Both tests could therefore be said to have inter-rater reliability.

11.8. TEST VALIDITY

Two types of validity were determined i.e. content validity and construct validity.

11.8.1. Content Validity

The data on content validity were provided by fifteen occupational therapists (group one) who rated the Space Relations and Position in Space tests on a 0 to 5 scale, according to how well they felt the test measured the defined skills (see Table 11.23).

TABLE 11.23 CONTENT VALIDITY - SR AND PIS

PERCEPTUAL SKILL	SPACE RELATIONS Mean Rating	POSITION IN SPACE Mean Rating
Form as mass	1,4	2,0
Overall Contour	1,5	1,3
Sub-Units of Form	1,6	1,3
Perceptual Constancy	2,3	3,4 **
Directionality	2,1	2,5
Space Relations	3,3 **	1,6
Position in Space	2,9	3,7 ***
Figure-Ground	2,1	1,4
Colour	0,1	0
Praxis	not required	not required

** a better than average assessment of skill *** a very much better than average assessment

As can be seen, the Space Relations test was estimated to be only a better than average way of evaluating the skill of space relations. The Position in Space test was thought to be a very much better than average way of measuring the skill of position in space but was also thought to measure perceptual constancy.

11.8.2. Construct Validity

11.8.2.1. Age Differentiation

Discussion already presented indicated that some perceptual skills deteriorate with age. In the case of the two motor reduced space tests, both were found to be sensitive to age related changes. The Spatial Relations test demonstrated a difference in the performance between the twenty to sixty-nine year old group and the seventy to seventy-nine year old group ($P < 0{,}01$). The Position in Space test showed a drop off in performance at a younger age group i.e. at fifty-nine years ($P < 0{,}0000$). Thus both were valid measures of age related changes in the ability to perceive space.

11.8.2.2. Correlation with Other Tests in the Battery

The highest correlation of the Space Relations test was with the Design Construction test ($r = 0{,}56$). This is realistic as both tests were constructed to measure spatial relations. Other correlations although lower i.e. with the Arrows test and the test of Visual Motor Integration, can be explained on the basis that some spatial ability is needed to complete the other tests mentioned. The low correlation between the Space Relations test and the Block Construction I and Stick Construction I tests can only be explained by taking the motor component of the latter two tests into account.

TABLE 11.24 CORRELATION WITH OTHER TESTS IN THE BATTERY - SR AND PIS

	SPACE RELATIONS TEST	POSITION IN SPACE TEST
Arrows	0,48 ***	0,25
VMI	0,43 ***	0,56 ***
Block Construction I	0,31	0,28
Block Construction II	0,24	0,48 ***
Stick Construction I	0,21	0,37
Stick Construction II	0,39 ***	0,23
Design Construction	0,56 ***	0,52 ***
Dot-Line	0,44 ***	0,49 ***
Figure-Ground	0,48 ***	0,42 ***
Perceptual Constancy	0,43 ***	0,37
Position in Space	0,27	1,00
Space Relations	1,00	0,27
Part-Whole	0,27	0,60 ***
Colour	0,12	0,21

*** P <0,001

The correlation of the Position in Space test with the VMI (r = 0,56) is interesting. It is possible that copying drawings of designs freehand requires a degree of space visualisation. Correlation with the Block Construction II test was expected as the tests were designed to measure the same skill as does the Design Construction test. Aspects of the Dot-Line test also measure space visualisation hence the correlation with that test.

However, the highest correlation (r = 0,60) was with the Part-Whole test indicating the extent to which space visualisation was needed to complete this task.

It is interesting to note that the correlation of the two motor reduced space tests is only (r = 0,27) which seemed to show that different components of space perception were being measured as intended.

A factor analysis was carried out on the tests in the battery to determine whether the tests had factors related to each other. It can be seen that the Position in Space test did not load at all for factor II which is probably a figure-ground/visual discrimination factor.

TABLE 11.25 SORTED ROTATED FACTOR LOADINGS - SR AND PIS

	FACTOR I	FACTOR II
Position in Space	0,80	0,0
Space Relations	0,56	0,78
Block Construction I	0,74	0,0
Part-Whole	0,73	0,0
VMI	0,63	0,41
Dot-Line	0,61	0,52
Figure-Ground	0,58	0,54
Design Construction	0,51	0,57

It appears that tests that are loaded for Factor I require spatial ability, in particular the type of ability that requires space visualisation. The Position in Space test had a high loading for this factor whereas the Spatial Relations test had a higher loading for Factor II which appeared to be more a measure of visual discrimination and figure-ground ability.

As no marker variables were included in the factor analysis, results cannot be readily used for the purpose of construct validity.

11.8.2.3. Results of Testing Patients

Both tests were found to be valuable in measuring the difficulties experienced by patients in the younger age groups, a statistically significant difference being found in the scores obtained by the patients and normal individuals (Space Relations test twenty to sixty-nine years $P < 0,001$ and the Position in Space test twenty to fifty-nine years $P < 0,001$). The lack of differences found between the scores obtained by older groups could be due to the small number of patients tested. However, as several patients in the older age groups could not do the tests at all, it would seem that performance, or lack thereof, could indicate problems in the area of space perception.

11.9. CONCLUSIONS

In summary the following has been found on the <u>spatial relations test</u> :

- it was an easy test to administer and judging by the difficulty p, it was a fairly easy test to complete although it was sensitive to age related changes;
- as in al' other tests it took longer for the older group of subjects as well as the patients to complete the test;
- there was also a difference in the performance of younger people between the normal group and the patients, thus the test was sensitive to changes which occur as the result of a stroke in this group;
- it was not clear why there was no difference in the performance between the normal elderly and the elderly patients. Small numbers in the patient group could not be the only reason for this lack of difference as many of the other tests with small numbers of elderly did show a difference. Taking out all the items that did not discriminate between subjects who could and those who could not do the test may make the test more useful for diagnosing problems in the elderly;

- reliability co-efficients were acceptable, except for the split-half reliability in the twenty to sixty-nine year old group;

- the rating for content as evaluated by experts was not as high as was expected. Although it was shown to have valid content for the skills that it was constructed to measure, the lower construct rating might have been due to the fact that the test was also found to evaluate perceptual constancy, directional, position in space and figure-ground skills;

- the test loaded for Factor II which seemed to be a figure-ground and visual discrimination factor. The results of this analysis seen together with the content validity study seemed to indicate that this test should be used to determine visual discrimination ability.

- the Block Construction I and Stick Construction I tests were compiled to provide information on the motor enhanced aspects of space relations. The Space Relation test, however, did not correlate well with either of these tests.

It would seem, therefore, that the test was not a good measure of spatial relations, but after undergoing revision and re-ordering of the items could be used to evaluate visual discrimination.

In contrast the following has been found on the <u>position in space test</u>:-

- it was an easy test to administer. However, it was not easy for the subjects to complete or easy for the subjects to understand what was required;

- because it was a timed test no item analysis was done;

- there was a "drop off" in the ability of normal subjects to complete the test at a fairly early age;

- the test was shown to be sensitive to changes that occur in the younger stroke patient, but not in the older group. A large number of patients were unable to do this test. This might be the reason for the lack of difference in the scores between the normal and the patient group. If a zero score had been recorded for these patients a different picture would probably have been found;

- the test demonstrated inter-rater reliability;
- construct validity for position in space was demonstrated although it also has some measure of perceptual constancy;
- it correlated with the Block Construction II test, a test which was constructed to measure position in space and space visualisation using motor output;
- it has a very high loading for Factor I which was thought to be a space visualisation factor;

It would appear that the Position in Space test was a good measure of this skill and should be retained in the battery. However, in constructing a new test, the Position in Space test should be reduced by half. Before the test can be reconstructed it should be administered in its entirety without the time limits, so that an item analysis can be conducted. The items can then be graded from easy to more difficult.

THE FORMULATION AND STANDARDISATION OF A BATTERY OF TESTS OF VISUAL-PERCEPTUAL-MOTOR FUNCTION FOR THE ADULT CEREBRO - VASCULAR ACCIDENT PATIENT

Marjorie Eileen Concha

VOLUME II

A Thesis Submitted to the Faculty of Medicine, University of the Witwatersrand, Johannesburg for the Degree of Doctor of Philosophy.

Johannesburg 1998

CONTENT

TABLES

FIGURES

CHAPTER 12

THE BLOCK CONSTRUCTION TESTS

12.1. INTRODUCTION

Construction with blocks has long been used as a method of determining three dimensional constructional ability, spatial concepts and the ability to motor plan. As far as it can be ascertained, the first person to use Block Construction in assessment was Poppelreuter in 1917 (Benton 1979).

Block Construction originated in the play of the child. Gesell (1966) stated that block play was one of the common interests of the pre-school child and that the child used this method to develop a space structure within the sphere of grasp and release i.e. within the realm of grasping space. According to Lezak (1983), building and assembly tasks involve the spatial component of perception at the conceptual level as well as in motor execution.

Although there is a tendency for drawing and constructional disturbances to occur together, this occurrence is variable (see Chapter Sixteen on apraxia), so that the two types of activity need to be evaluated separately (Lezak 1983). Two types of constructional tasks were used in the V-P-M battery i.e. construction with blocks involving mainly three dimensional constructions and two dimensional construction as in the stick and design construction tests. Only the Block Construction tests are discussed in this chapter.

The Block Construction tests were drawn up by the candidate for the purpose of evaluating spatial relationships and space visualisation and the link between spatial skills and motor output. Two tests were used i.e. construction from a three dimensional model and construction from a two dimensional drawing on a card.

A description of the Block Construction I test is to be found in Chapter Four (4.6.10.) and Appendix 1, Section 2.10.1. while a description of the Block Construction II test will be found in section 4.6.11. of Chapter Four and Appendix 1 Section 2.10. 2.

12.2. REVIEW OF SIMILAR TESTS

Block Construction tasks may involve horizontal or vertical arrangements on a flat surface or they may use a combination of planes. The model may be three dimensional or presented in abstract two dimensional form. Benton (1979:221) stated that these tasks are not *"... equivalent in their demands on sustained attention ... perceptual acuity, the apprehension of spatial relationships, judgement of perspective and motor skill"*. It is for this reason that different types of Block Construction tests have been developed.

12.2.1. Two Dimensional Horizontal And Vertical Arrangements

Several tests that involved construction in the horizontal and vertical on a flat surface were available. However, they all included matching colour to make a design.

Kohs' Blocks were the forerunners of this type of block design and many tests have since been based on this test (Lezak 1983). The most widely used replacement of the Kohs' Block test is the Block Design sub-test of the Wechsler Intelligence Scales for Adults (WAIS) (Wechsler 1981). In the WAIS sub-test the subject used blocks coloured red and white to copy a design presented on a smaller scale in a booklet. The WAIS sub-test is purported to measure visuo-spatial ability, conceptualisation and the ability to form a "gestalt".

The Goldstein and Sheerer Cube Test is another block construction test based on the concept of Kohs' blocks (Grassi 1970). There are, however, no normative data for the Goldstein and Sheerer Cube Test.

The Cube Copying sub-test of the RPAB (Whiting et al 1985), is similar to the WAIS test, but in this case the blocks are coloured blue and yellow (see Fig 12.1). The LOTCA Test (Itzkovich et al 1990) has a similar sub-test using blocks with colours.

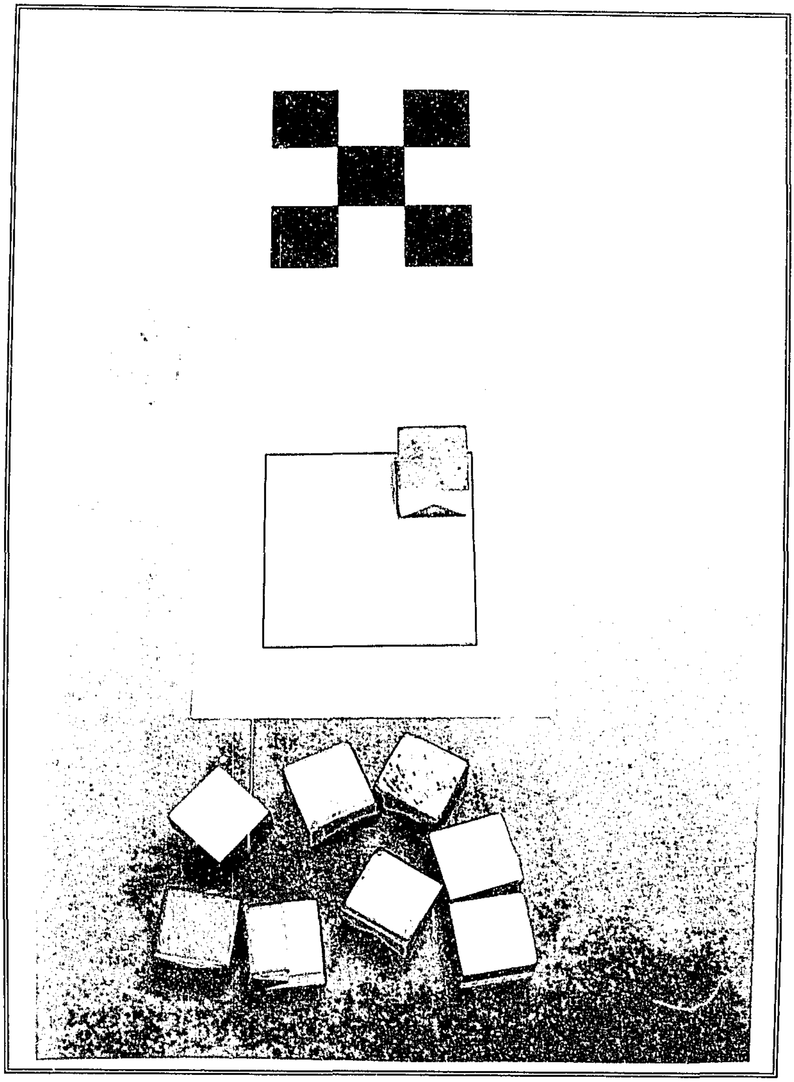


FIGURE 12.1 CUBE COPYING TEST - RIVERMEAD PERCEPTUAL ASSESSMENT BATTERY

On analysing the perceptual skills evaluated by the above mentioned tests it seemed that, apart from spatial ability and the ability to form a "gestalt" (Part-whole perception), colour discrimination ability was also required. As these tests are also used to assess conceptualisation it seemed desirable to compile a new Block Construction test, the execution of which required mainly spatial and constructional ability. Construction in the horizontal and vertical on a flat surface was not included in the V-P-M battery due to the limitations of this dimension should colour not be incorporated into the design. However, there are two designs, one in the Block Construction I test and the other in the Block Construction II test which require the construction to be made on the flat surface.

12.2.2. Three Dimensional Construction

Gesell (1966) described various block constructions used with children. These increase in complexity as the child gets older. Gesell's constructions have been used in a test for the evaluation of the pre-school child (Ames et al 1980). A tower of ten blocks and a train can be produced by a child of two and a half years old while a bridge can be constructed by a three and a half year to four year old child, a gate can be constructed at five years and steps at six years. Two of these Block Constructions tasks, i.e. the tower and the bridge, are included in the Standford-Binet Battery (Terman and Merril 1973), an intelligence scale with norms for ages ranging from two years through to the young adult.

Grassi (1970) tried to improve the Goldstein and Sheerer Cube Test in his Grassi Block Substitution Test which required the subject to reproduce a three dimensional model instead of the two dimensional construction demanded by the Goldstein and Sheerer test. The Grassi test called for the subject to reproduce the design on all sides i.e. top, side, and bottom of the construction thus requiring the complex skill of space visualisation.

Non-standardised three dimensional Block Construction tasks have been described by Hecaen et al (1951) and Goodglass and Kaplan (1972), however, the most commonly used three dimensional construction task is that formulated by Benton et al (1983). This test, originally called the Three Dimensional Construction Praxis Test (Benton and Fogel 1962) consists of three different constructions of increasing complexity. Benton et al (1983) obtained data for this test on one hundred control patients between the ages of sixteen and sixty-three years, who had no history of brain disease or injury. They found that the mean score for the older patienst i.e. those above fifty years was higher than for those below fifty years, however, no statistical analysis was reported on the data. A patient was said to have intact function if he scored twenty to twenty-two points or above as ninety percent of normal controls obtained this score.

Benton et al (1983) also evaluated patient performance by comparing the copying of a three dimensional model with that of a photographic model. He found that the performance improved when the three dimensional model was used.

Farver and Farver (1982) used the Block Construction test described in the Boston Diagnostic Aphasia Examination (Goodglass and Kaplan 1972) to determine age related changes in constructional tasks. This test consists of ten designs presented in the form of photographs. The subject is given one minute to complete each construction which is then scored correct or incorrect. Farver and Faver (1982) used this test to evaluate sixty-eight normal individuals ranging in age from forty years to eighty years. They found a statistically significant age effect with increasing age. Differences were found in the performance of the three age groups of sixty to sixty-nine, seventy to seventy-nine and eighty to eighty-nine years.

Three tests have more recently been formulated for three dimensional Block Construction i.e. the Three Dimensional Copying test of the RPAB (Whiting et al 1985), the Three Dimensional Block Construction test of the COTNAB (Tyerman et al 1986) and the Plain Block Design of the LOTCA (Itzkovich et al 1992).

The RPAB consists of only two tasks, the first a fairly simple one which is used as a demonstration item (Fig 12.2). The second task is a more complicated construction made up of several different sized and shaped blocks. The total test score is obtained from one construction which the candidate feels is undesirable as one mistake can be easily compounded. The COTNAB Block Construction test is

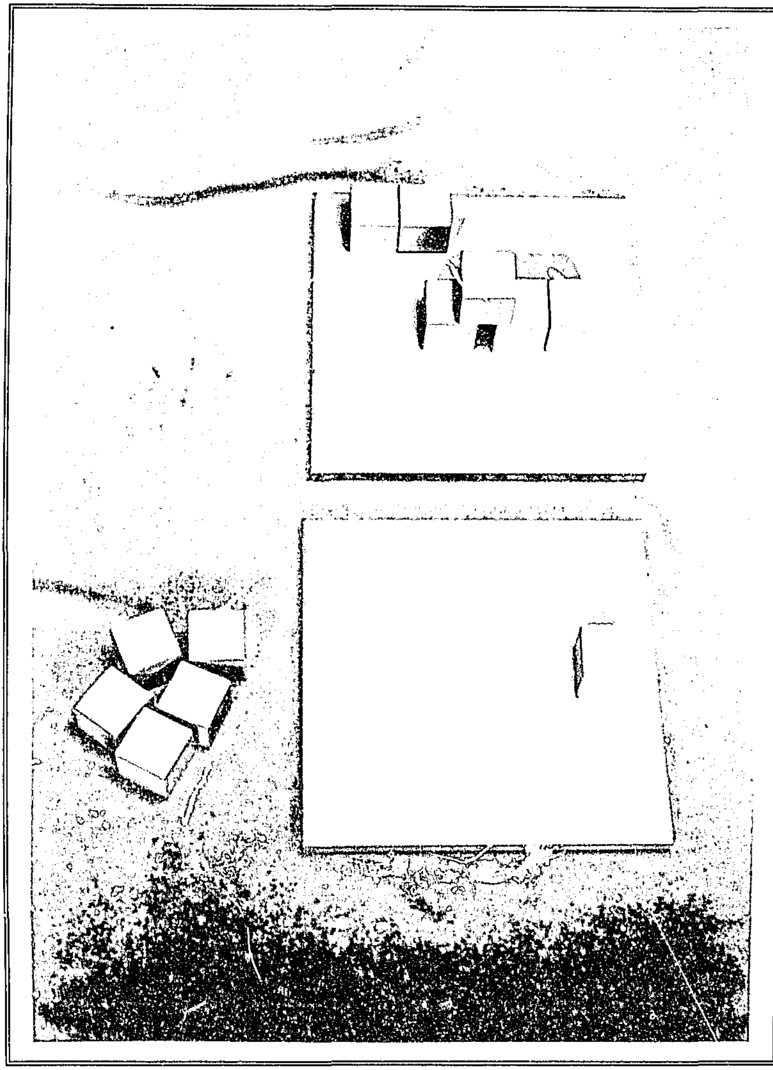


FIGURE 12.2 TRIAL ITEM - RIVERMEAD PERCEPTUAL BATTERY - THREE DIMENSIONAL BLOCK CONSTRUCTION

very similar, the only difference being that the blocks supplied to complete the construction are all of the same size and shape as those of the model. The LOTCA has the same characteristics but the model is depicted on a card. No age related changes were found in normal performance on either the RPAB test or the LOTCA. The COTNAB has different tables for each of three age groups i.e. sixteen to thirty years, thirty-one to forty-nine years and fifty to sixty-five years. However no statistical comparisons have been made between the scores for each of these groups.

12.3. Reasons For Compiling New Tests

Tests developed by Benton and Fogel (1962) and Goodglass and Kaplan (1972) could have been used by the candidate, however, as these tests required the subject to identify the size and shape as well as the spatial location of each piece, it was decided to construct a test using blocks of a uniform size, shape and colour. In addition the available administration and scoring instructions for the above mentioned tests were felt to be inadequate to assure test-retest reliability.

A test was therefore constructed that:

a) evaluated the ability to construct from both a two and a three dimensional model in order to evaluate grasping and action space.

b) provided rigorous scoring and administration criteria.

c) provided an evaluation of normal function against which patient performance could be evaluated.

Two Block Construction tests were designed. The Block Construction I test requires the subject to copy a three dimensional model and the Block Construction II test, the subject to copy a two dimensional drawing.

12.4. Results Of Testing Normal Individuals

12.4.1. Accuracy Scores - Block Construction I Test

The means and the standard deviations were calculated for the normal group (i.e. group three) as a whole (Table 12.1) and for each age group individually (Table 12.2).

TABLE 12.1 MEAN AND SD OF ACCURACY SCORES - BCI

n	MAX SCORE	MEAN	SD	RANGE
118	12	11,5	0,96	8-12

TABLE 12.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - BCI

AGE	n	MEAN	SD	RANGE
20-29	20	11,9	0,31	11-12
30-39	21	11,38	1,07	9-12
40-49	19	11,90	0,32	11-12
50-59	19	11,53	0,70	10-12
60-69	19	11,37	1,30	8-12
70-79	20	10,95	1,23	9-12

The Pearson Chi-square Test was used to determine whether any difference existed in the performance within the group as a whole. A cut point of eleven was used with the percentage obtaining a score greater than and lower than the cut point being used to calculate the Chi-square value (Table 12.3).

TABLE 12.3 COMPARISON OF ACCURACY SCORES - BCI

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
less than 11	10,0	28,6	10,5	36,8	21,1	50,0	26,3
greater than 11	90,0	71,4	89,5	63,2	78,9	50,0	73,7

Pearson Chi-square Value = 12,40 $P < 0,05$

As a significant difference was found at the $P < 0,05$ level of significance further analysis took place to determine where the differences were to be found. The results of this analysis were as follows:-

a) The Pearson Chi-square Test showed that no difference existed between the age groups twenty to twenty-nine, thirty to thirty-nine, forty to forty-nine, fifty to fifty-nine and sixty to sixty-nine ($P = 0,1839$). These age groups were, therefore, combined for the purpose of further analysis.

b) Yates' Corrected Chi-square test showed that there was a difference between the performance of the twenty to sixty-nine year old group and that of the seventy to seventy-nine year old group (P = 0,0179) at the P< 0,05 level of significance (Table 12.4).

TABLE 12.4 COMPARISON OF ACCURACY SCORES FOR AGES 20 - 69 AND 70 - 79 YEARS - BCI

AGE	20-69	70-79	TOTAL
Less than 11	21,4	50,0	26,3
Greater than 11	78,6	50,0	73,7

Yates' Corrected Chi-square Value = 0,0124 P <0,05

12.4.2. Accuracy Scores - Block Construction II Test

The means and the standard deviations were calculated for the normal group i.e. Group three, as a whole (Table 12.5) and for each age group individually (Table 12.6).

TABLE 12.5 MEAN AND SD OF ACCURACY SCORES - BCII

n	MAX SCORE	MEAN	SD	RANGE
118	12	9,84	2,42	2-12

TABLE 12.6 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - BCII

AGE	n	MEAN	SD	RANGE
20-29	20	10,9	1,52	7-12
30-39	21	10,38	1,86	6-12
40-49	19	9,84	2,12	6-12
50-59	19	9,90	2,28	4-12
60-69	19	9,86	2,87	4-12
70-79	20	8,30	3,03	2-12

The Pearson Chi-square Test was used to determine whether any differences in the accuracy scores existed between the various age groups (Table 12.7).

TABLE 12.7 COMPARISON OF ACCURACY SCORES - BCII

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
less than 10	35,0	47,6	52,6	57,9	52,6	75,0	53,4
greater than 10	65,0	52,4	47,4	42,1	47,4	25,0	46,6

Pearson Chi-square Value = 6,916 = Not significant

As there were no differences amongst the age groups no further analysis was carried out. The data on both the Block Construction I and II tests were, for the sake of convenience, grouped as per the other data in the battery for the calculation of percentiles i.e. twenty to thirty-nine, forty to fifty-nine, sixty to sixty-nine and seventy-seventy-nine.

Thus it can be seen that the null hypothesis which stated that decreased ability will not be found with advancing age in visual-perceptual motor function was rejected in the case of Block Construction from a three dimensional model (i.e. Block Construction I) at the 5% level of significance, and accepted in the Block Construction II test.

Although the Block Construction I test demonstrated a drop off in function in the elderly subjects, the mean for the groups as a whole ($\bar{x}$ = 11,5) was considerably higher in this test than in the Block Construction II test ($\bar{x}$ = 9,84). This seemed to reinforce the findings of Benton et al (1983) and Fall (1987) that Block Construction from a model is easier than that from a drawing or photograph and possibly makes different perceptual demands.

12.4.3. Time Scores - Block Construction I Test

The means and the standard deviations were calculated for the time scores for the group as a whole (Table 12.8) and for each group separately (Table 12.9).

TABLE 12.8 MEAN AND SD OF TIME SCORES - BCI

n	MEAN/SECS	SD	RANGE
118	83,08	34,35	41-193

TABLE 12.9 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - BCI

AGE	n	MEAN/SECS	SD	RANGE
20-29	20	60,35	13,52	41-89
30-39	21	67,91	24,81	41-148
40-49	19	74,74	22,17	45-138
50-59	19	76,53	16,95	52-110
60-69	19	95,74	34,88	52-189
70-79	20	123,85	41,55	62-193

Levine's test for equal variance established that there was a significant difference in the variances among the group (P=0,0001) and the Welch Test established that there was a significant difference between the means of the time scores obtained by the different groups (P=0,0000). The separate variance t was, therefore, used with the Bonferroni correction for multiple comparisons, to establish where the differences were (Table 12.10).

On the basis of the differences found, further analysis was carried out (Table 12.11).

Although no differences existed among some of the regrouped ages, percentiles were calculated in the twenty to thirty-nine year, forty to fifty-nine year, sixty to sixty-nine year and seventy to seventy-nine year age groups.

TABLE 12.10 DIFFERENCES IN TIME SCORES - BCI

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,2322
20-29	40-49	0,0213
20-29	50-59	0,0024 *
20-29	60-69	0,0004 **
20-29	70-79	0,0000 ***
30-39	40-49	0,3635
30-39	50-59	0,2042
30-39	60-69	0,0070
30-39	70-79	0,0000 ***
40-49	50-59	0,7815
40-49	60-69	0,0344
40-49	70-79	0,0001 **
50-59	60-69	0,0402
50-59	70-79	0,0001 **
60-69	70-79	0,0277

* P<0,05 ** P<0,01 *** P<0,001

TABLE 12.11 DIFFERENCES IN RE-GROUPED TIME CORES - BCI

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-49	0,0126
20-39	60-69	0,0012 **
20-39	70-79	0,0000 ***
40-59	60-69	0,0282
40-59	70-79	0,0001 ***
60-69	70-79	0,0277

** P<0,01 *** P<0,001

12.4.4. Time Scores - Block Construction II Test

The means and the standard deviations were calculated for the time scores for the group as a whole (Table 12.12) and for each group separately (Table 12.13)

TABLE 12.12 MEAN AND SD FOR TIME SCORES - BCII

n	MEAN/SECS	SD	RANGE
118	94,07	80,05	30-576

TABLE 12.13 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - BCII

AGE	n	MEAN/SECS	SD	RANGE
20-29	20	51,80	12,12	30-74
30-39	21	67,14	30,54	33-162
40-49	19	83,95	60,45	39-316
50-59	19	82,84	61,07	41-311
60-69	19	126,42	94,72	52-360
70-79	20	154,15	122,80	48-576

Levine's test for equal variance established that there was a significant difference between the variances of the different age groups ($P = 0{,}0002$), and the Welch Test demonstrated a significant difference between the mean of the time scores ($P = 0{,}0001$). The separate variance t was, therefore, used with the Bonferroni correction, to establish where the differences were (Table 12.14).

The only difference found was between the twenty to twenty-nine year olds and the sixty to sixty-nine year olds and the seventy to seventy-nine year olds at $P<0{,}05$. The statistician felt that the scores should be regrouped into age categories as previously described and similar statistical procedures used to investigate any possible differences between the regrouped data. Levine's test ($P = 0{,}0000$) and the Welch Test ($P = 0{,}0003$) showed significant differences between the groups, therefore, the separate variance t was used to demonstrate where those differences were (Table 12.15).

TABLE 12.14 DIFFERENCES IN TIME SCORES - BCII

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,0424
20-29	40-49	0,0345
20-29	50-59	0,0422
20-29	60-69	0,0030 *
20-29	70-79	0,0015 *
30-39	40-49	0,2848
30-39	50-59	0,3209
30-39	60-69	0,0163
30-39	70-79	0,0056
40-49	50-59	0,9556
40-49	60-69	0,1097
40-49	70-79	0,0303
50-59	60-69	0,1020
50-59	70-79	0,0232
60-69	70-79	0,4337

* P<0,05

On the basis of the results it was possible to calculate percentiles for each of the age categories previously described.

12.4.5. Discussion Of The Results Of Testing Normal Individuals

It was interesting to note that block construction from a two dimensional model depicted on a card (Block Construction II test), although seemingly the more difficult of the two tests (the means of the scores for each age group were lower than for the Block Construction I test), did not show age related changes. Comparable

TABLE 12.15 DIFFERENCES IN RE-GROUPED TIME SCORES - BCII

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-49	0,0275
20-39	60-69	0,0069 *
20-39	70-79	0,0028 *
40-59	60-69	0,0826
40-59	70-79	0,0230
60-69	70-79	0,4337

* P<0.05

tests in which the two dimensional model was depicted in a photograph such as the Three Dimensional Construction Test (Benton et al 1983) and the Block Construction test from the Boston Diagnostic Aphasia Battery (Goldglass and Kaplan 1972) demonstrated that changes occurred with age. Farver and Farver (1982) as already mentioned, found a statistically significant age effect on the Boston Diagnostic Aphasia Battery. The age effects found were probably due to the fact that both these tests used blocks of different sizes and shapes and thus involved a variety of skills, unlike the Block Construction II test of the V-P-M battery.

The type of mistakes made by normal individuals consisted mainly of spatial alignment. The design was not squared to the edge of the construction pad and the blocks were unequally aligned.

The time scores refl ted an expected increase in the time needed to complete each test with advancing age. The statistical difference between the different ages was greater for the Block Construction I test.

12.5. Results Of Testing Patients

12.5.1. Accuracy Scores - Block Construction I Test

The means and standard deviations were calculated for the age groups between which no statistical difference was found when testing the normal subjects i.e. twenty to sixty-nine years and seventy to seventy-nine years (see Table 12.16). No standard deviation was calculated for the older group as the scores of only two patients were available.

TABLE 12.16 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - BCI

AGE	n	MEAN	SD	RANGE
20-69	22	9,68	2,10	2-12
70-79	2	6,5	-	4-9

12.5.2. Accuracy Scores - Block Construction II Test

The mean and standard deviation was calculated for the age group for which no statistical difference was found in the normal subjects. For this test, therefore, the scores of patients of all age groups were combined (see Table 12.17).

TABLE 12.17 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - BCII

AGE	n	MEAN	SD	RANGE
20-79	24	7,5	4,16	0-12

12.5.3. Time Scores - Block Construction I

The means and the standard deviations were calculated for the age groups between which no statistical difference was found in the normal groups i.e. twenty to fifty-nine years and sixty to seventy-nine years (see Table 12.18).

TABLE 12.18 MEAN AND SD FOR TIME SCORES FOR PATIENTS - BCI

AGE	n	MEAN/SECS	SD	RANGE
20-59	14	138,79	50,57	84-276
60-79	10	219,40	98,42	117-379

12.5.4. Time Scores - Block Construction II

The means and the standard deviations were calculated for the age groups between which no statistical difference was found in the normal group i.e. twenty to fifty-nine years and sixty to seventy-nine years (see Table 12.19).

TABLE 12.19 MEAN AND SD FOR TIME SCORES FOR PATIENTS - BCII

AGE	n	MEAN/SECS	SD	RANGE
20-59	14	127,86	42,97	69-202
60-79	10	259,80	212,29	124-443

12.5.5. Comparison Between Patients And Normal Subjects

12.5.5.1. Accuracy Scores - Block Construction I Test

Comparisons were made between the patients and the normal subjects in the age groups previously defined. As Levine's test for variances between the groups was found to be significant, the Welch Test was used to test for significant differences between the means of the scores obtained by the normal group and the patient group twenty to sixty-nine years (Table 12.20). The difference between the performance of the two groups was significant at the $P < 0,01$ level of significance. No analysis was carried out for differences experienced by the seventy to seventy-nine year group as only two patients completed the test.

TABLE 12.20 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - BCI

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20-69	0,0000	0,0068 **
70-79	-	-

** P<0,01

12.5.5.2. Accuracy Scores - Block Construction II Test

Testing for significant differences between the accuracy scores obtained by the normal group and the patient group was carried out on the same basis as for the Block Construction I test. In this instance the difference between the normal and the patient group was only probably significant at the P < 0,05 level of significance (Table 12.21).

TABLE 12.21 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - BCII

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20-79	0,0000	0,0116 *

* P<0,05

12.5.5.3. Time Scores - Block Construction I Test

Comparisons were made in the age groups between which no statistical difference was found in the time taken by the normal group to complete the test (Table 12.22). Differences in both groups were found to be significant at the p < 0,01 level of significance.

TABLE 12.22 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - BCI

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20-59	0,0001	0,0002 **
60-79	0,0000	0,0065 **

** P <0,01

12.5.5.4. Time Scores - Block Construction II

Comparisons between the time scores obtained by the patients and those obtained by the normal individuals were made on the Block Construction II test as described above (see Table 12.23).

TABLE 12.23 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - BC II

AGE	LEVINE'S TEST P VALUE	F TEST P VALUE
20-59	0,2973	0,0000 ***
60-79	0,4122	0,0059 **

** P<0,01 *** P<0,001

12.5.6. Analysis of Tests Executed by Patients

12.5.6.1. Block Construction I Test

a) In the age group twenty to thirty-nine years only one patient obtained a score below the fifteenth percentile and could therefore be said to have a problem with spatial relationships. Two patients were discharged prior to completing the test.

b) In the age group forty to fifty-nine years, five of the ten patients received scores below the fifteenth percentile, two were discharged prior to completing the test and three received a full score.

c) Of the ten patients in the sixty to sixty-nine year old group, only three demonstrated spatial relationship problems as they obtained a score below the fifteenth percentile and one was untestable.

d) Of the five patients who completed the test in the seventy to seventy-nine year age group one obtained a score below the fifteenth percentile, three were untestable and one was discharged prior to completing the test.

12.5.6.2. Block Construction II Test

a) In the twenty to thirty-nine year age group, three patients obtained scores above the fifteenth percentile, two were discharged and two had spatial relationship problems as they obtained scores below the fifteenth percentile.

b) In the forty to fifty-nine age group, five patients received scores above the fifteenth percentile, three below and two were discharged.

c) Of the ten patients in the sixty to sixty-nine year age group, five patients received scores above the fifteenth percentile, four demonstrated problems with a score below the fifteenth percentile. One was untestable.

d) In the seventy to seventy-nine year age group of the two patients who completed the test, one received a score above the fifteenth percentile and one below.

Three patients in the group were untestable and one was discharged.

Sixteen of the twenty-four patients who completed the Block Construction tests obtained a higher score on the Block Construction I test, five an identical score and only three a higher score on the Block Construction II test, an indication that the Block Construction II test is more difficult for patients.

The patients also took a longer period of time to complete the Block Construction tests than did the normal subjects.

12.5.7. Problems That May Be Diagnosed From Performance On The Block Construction Tests

The analysis of the Block Construction tests was more difficult than that of paper and pencil tests as no permanent record was kept of each patient's construction. An attempt was made to record the type of mistakes made either through drawing or describing in written form the appearance of each patient's construction and the process used by the patient to construct the design. Immediately the patient completed his construction it was dismantled by the examiner to make way for the next item. Due to time constraints it was not always possible to record all the observations.

The following types of performance and mistakes were recorded in the Block Constructions tests:-

a) Additions to the construction i.e. the patient used too many blocks in the construction. This type of mistake occurred in the construction of patient fifteen, and could be attributed to a problem with motor planning or space relationships.

b) Blocks were omitted from the design. Patient fifteen also made this type of mistake.

c) General untidiness of a design which was basically correct but which made it difficult to score the construction accurately. This type of performance was seen in the constructions of patient four and could be attributed to a motor planning or inco-ordination problem.

d) Designs were rotated (Fig 12.3) as seen in the constructions of patients eight and fifteen (position in space problem).

e) Use of trial and error to complete the design i.e. the patient re-orders, removes and replaces blocks until the construction is correct (patient sixteen) (planning problem).

f) Constant restructuring of a design which does not result in a correct construction in the end (patient nine) (planning problem).

g) Poor and incorrect spatial placement of a block (Fig 12.4 and 12.5 and 12.6) as seen in the construction of patient one, two and nine (spatial relations problem).

Space Relations i.e. Grasping Space. The Block Construction tests were said to measure this skill as rated by the fifteen occupational therapists. The Block Construction I test was specifically constructed for this purpose. As the difference between the normal group and the patient group was almost certainly significant at the 1% level it was felt that this was a good test for space relations. Problems can be seen in figures 12.6 and 12.7.

Position in Space i.e. Action Space. The Block Construction II test was constructed for the purpose of evaluating Action Space. However, the therapists did not feel that it was an adequate measure of this skill. There was only a probably significant difference between the normal and the patient group at the 5% level of significance on this test. It was felt that this difference was sufficient to make this test an adequate method of testing patients. However, modifications should take place meet the requirements of construct validity. Mistakes made , ' the patient are seen in Figures 12.8 and 12.9.

Apraxia. The Block Construction tests were designed to evaluate constructional apraxia. The type of problems described by Critchley (1953) as being symptomatic of apraxia was clearly seen in the performance of patient nine, i.e. constant re-ordering of blocks, perseveration, the building of her construction on top of the model and difficulties in the correct spatial placement of blocks (Fig 12.9).

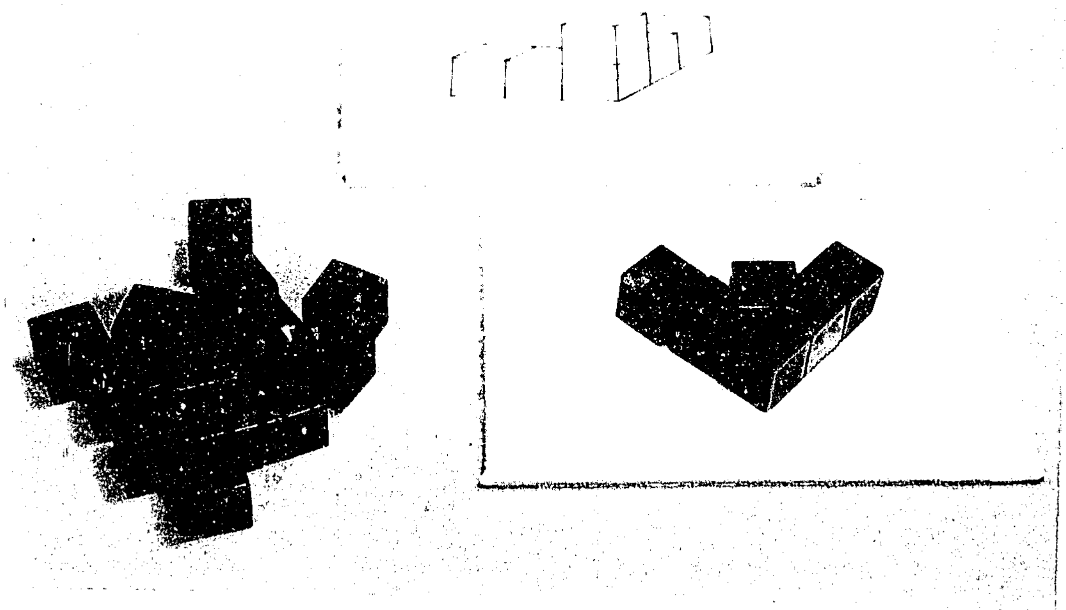


FIGURE 12.3 CONSTRUCTION ROTATED

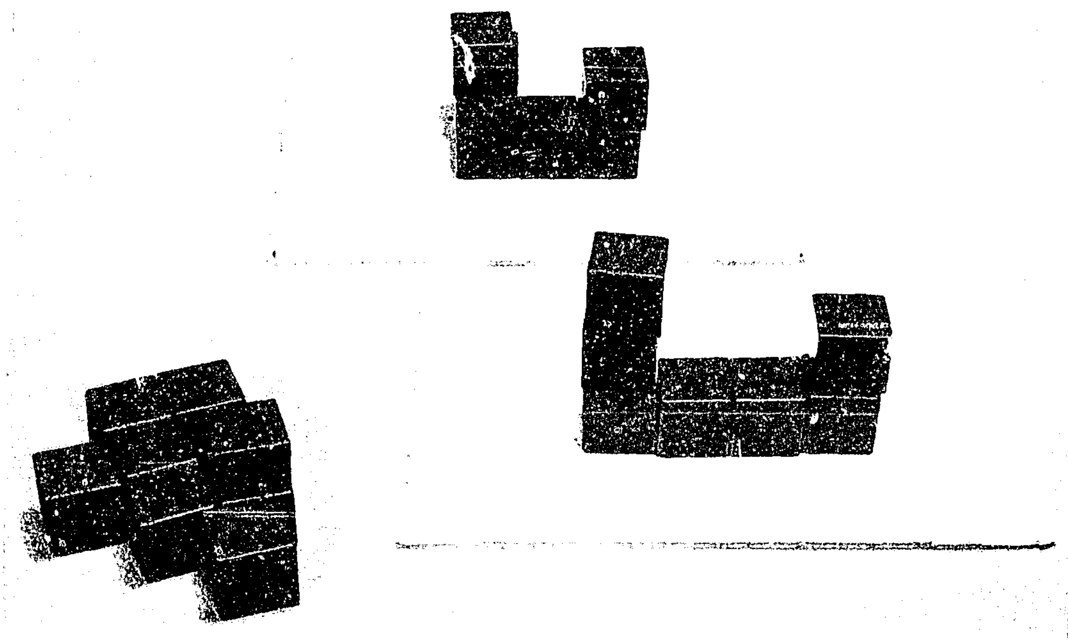


FIGURE 12.4 POOR SPATIAL PLACEMENT

FIGURE 12.5 INCORRECT SPATIAL PLACEMENT

FIGURE 12.6 INCORRECT SPATIAL PLACEMENT

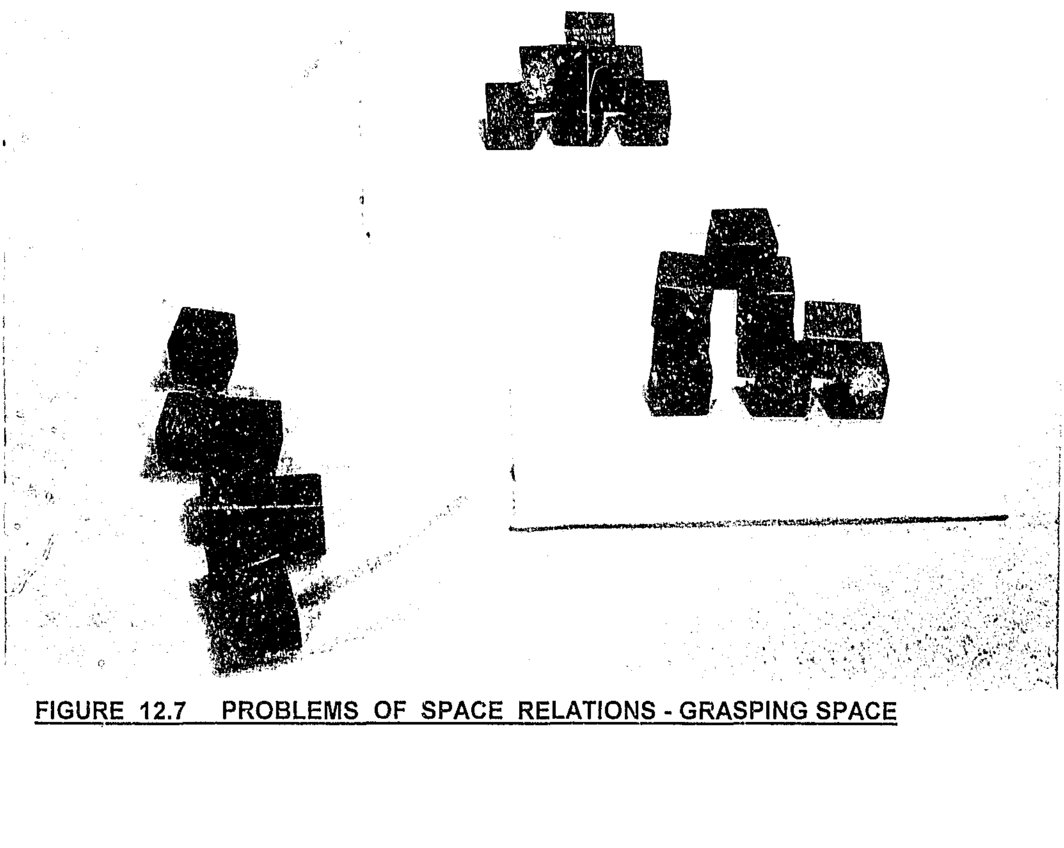


FIGURE 12.7 PROBLEMS OF SPACE RELATIONS - GRASPING SPACE

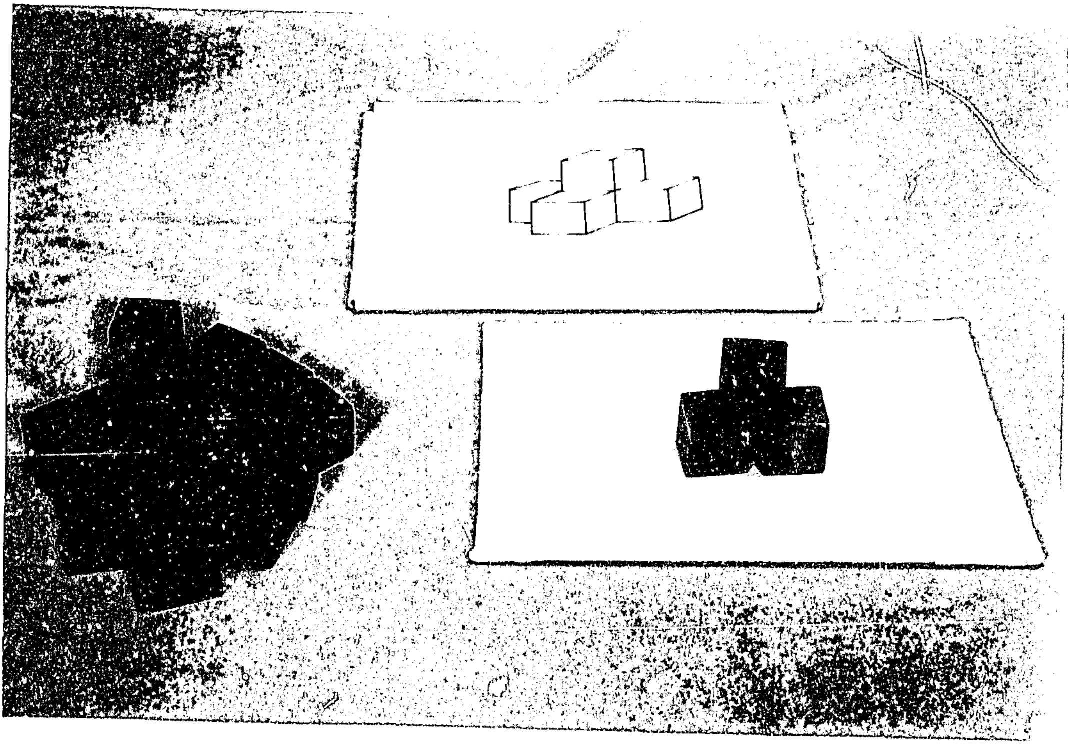


FIGURE 12.8 PROBLEMS OF POSITION IN SPACE - ACTION SPACE

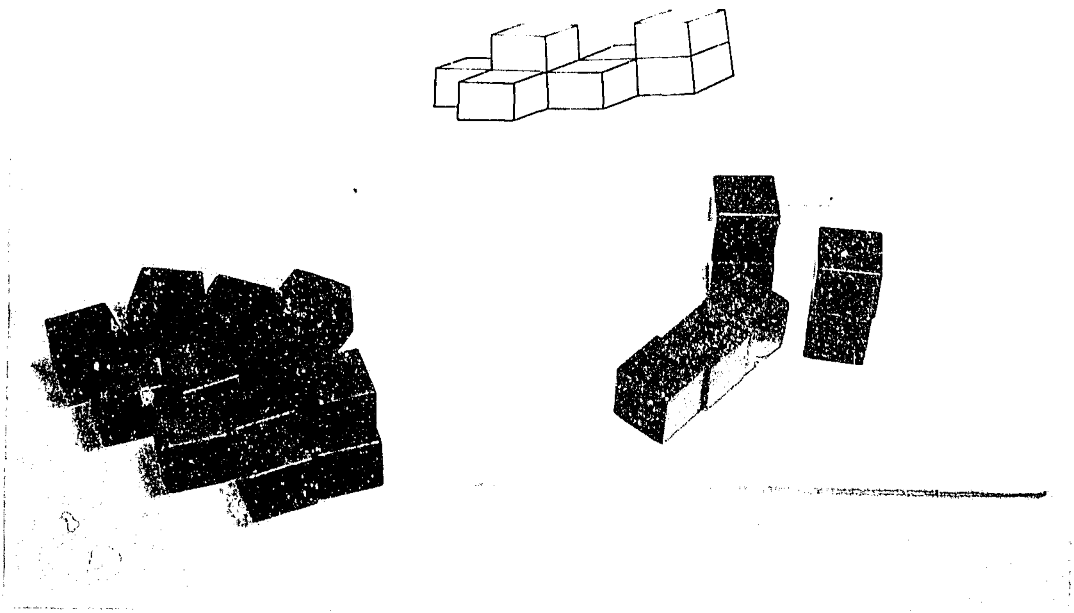


FIGURE 12.9 PROBLEMS OF POSITION IN SPACE - ACTION SPACE

12.6. ITEM ANALYSIS

12.6.1. Difficulty Index

The difficulty index was calculated using the responses from the normal group of individuals in the age groups between which no statistical differences were found in the accuracy scores i.e. Block Construction I test ages twenty to sixty-nine and seventy to seventy-nine years and Block Construction II test - all age groups combined (see Table 12.24 and 12.25).

Some small variations occurred between the two age groups in the Block Construction I test but generally the items demonstrated a high difficulty p and were, therefore, fairly easy. The items of the Block Construction II test showed a greater variation in the difficulty index and could, therefore, be presented in a different order i.e. item two, one, three, five, four and six.

TABLE 12.24 DIFFICULTY INDEX - BCI

ITEM	DIFFICULTY p 20-69 YRS	DIFFICULTY p 70-79 YRS
1	0,94	0,93
2	0,96	0,93
3	0,98	1,00
4	1,00	0,93
5	0,91	0,79
6	0,97	1,00

The difficulty index for the responses obtained from the patient group was also calculated (Table 12.26 and 12.27). The Pearson Chi-square Test was used to calculate whether there were statistical differences between the normal and the patient group in the response obtained on each item. This was done to help to determine the final test form. The probability values refer to the probable differences between the patients and the normal subjects on each item of the tests.

TABLE 12.25 DIFFICULTY INDEX - BCII

ITEM	DIFFICULTY p 20-79 YEARS
1	0,90
2	0,92
3	0,79
4	0,69
5	0,70
6	0,54

No calculations were carried out for the seventy year old group on the Block Construction I test due to the small number of patients in the group. The younger groups of patients found items three, four, five and six more difficult than the normal subjects on this test.

TABLE 12.26 DIFFICULTY INDEX - PATIENT GROUP - BCI

ITEM	DIFFICULTY p 20-69 YEARS
1	0,77
2	0,86
3	0,68 ***
4	0,68 ***
5	0,77 **
6	0,59 ***

** P< 0,01 *** P< 0,001

TABLE 12.27 DIFFICULTY INDEX - PATIENT GROUP - BCII

ITEM	DIFFICULTY p 20-79 YEARS
1	0,79 **
2	0,71 ***
3	0,63
4	0,29 ***
5	0,58 **
6	0,33

** P<0,01 *** P<0,001

In the Block Construction II test the Pearson Chi-square Test demonstrated that statistical differences were found between the patient group and the normal group items three to six (BC I test) and items one, two, four and five (BC II test). These items should therefore be retained in the final test form.

12.6.2. *Item Discrimination*

This is the extent to which the item discriminates between those who do well on the test and those who do not. It was, however, impossible to carry out tests for determining the discriminatory level of each item for the following reasons:-

12.6.2.1. Testing the Significance of the Difference between two proportions

Due to the fact that the difficulty p of each test item was high it was difficult to establish cut-points, therefore, this procedure was not carried out on either of the Block Construction tests.

12.6.2.2. Correlational Approach

For the reasons given above, it was also inappropriate to attempt this analysis on both Block Construction tests.

12.7. Test Reliability

12.7.1. Split-Half Reliability

This determines the internal consistency of the test with regard to content sampling. The resultant co-efficient of correlation for the Block Construction I test was r = 0,47 for age groups twenty to sixty-nine years and r = 0,32 for the age group seventy to seventy-nine years. This demonstrated a low degree of correlation between the two halves of the test (Table 12.28). This test was unusual in that the younger age group demonstrated a slightly higher reliability co-efficient than that of the older group, contrary to the findings in other tests.

The split-half reliability for the Block Construction II test for all age groups combined, was r = 0,73.

12.7.2. Kuder-Richardson Formula 20

In this method, the pass fail score for each item was compared with the total test score using the Kuder-Richardson Formula 20. This procedure determined the homogeneity of the test items or the inter-item consistency of the test. For both the Block Construction I test and the Block Construction II test, the Kuder-Richardson Formula was calculated for all age groups combined. This gave a co-efficient of r = 0,73 for the Block Construction I test and r = 0,86 for the Block Construction II test which indicated a high degree of correlation between each item and the total test score.

Thus the co-efficient of correlation ranged between r = 0,32 and r = 0,86. All reliabilities except for the split-half reliability of the age group seventy to seventy-nine years in the Block Construction I test showed statistical significance (Table 11.28). Reliability co-efficients of 0,80 or above are traditionally accepted as evidence of adequate test reliability although many tests are published with reliabilities lower than this. It would seem, therefore, that the Block Construction II Test was the more reliable of the two.

TABLE 12.28 RELIABILITY CO-EFFICIENTS - BCI AND BCII

TEST	AGE	SPLIT-HALF	KUDER-RICHARDSON
BC I test	20-69	0,47 ***	
	70-79	0,32	
	all ages	0,43	0,73 ***
BC II test	all ages	0,73 ***	0,86 ***

** P <0,01 *** P <0,001

12.7.3. Inter-Rater Reliability

The Friedman two-way analysis of variance demonstrated that there was no statistical difference between either the accuracy (p=0,1353 Block Construction I and p=0,3114 Block Construction II test) or time (p=0,0695 Block Construction I and p=0,1298 Block Construction II test) scores obtained by three examiners on each of six individuals tested. Both tests could, therefore, be said to have inter-rater reliability.

12.8. TEST VALIDITY

Two types of validity are discussed:-

12.8.1. Content Validity

The data on content validity were provided by fifteen occupational therapists who rated the Block Construction tests on 0-5 scale (Table 12.29).

As can be seen the rating came close to the upper quartile in the skill of space relations for both tests. Both block tests could only be considered as a better than average assessment of this perceptual skill. The Block Construction II test was compiled for the purpose of measuring position in space (the space visualisation component) but it seemed that the therapists did not feel that space visualisation was measured to any degree in this test. This lack of congruity between the definition provided and the rating could be due to the lack of emphasis placed on the spatial constructs of action space in the perceptual literature and in treatment programmes. It could also indicate that the Block Construction II test was a poor test of this skill. Both tests were found to require motor planning ability and were, therefore, felt to be valid measures of constructional praxis and space relations.

TABLE 12.29 CONTENT VALIDITY - BC I AND BC II

PERCEPTUAL SKILL	BLOCK CONSTRUCTION 1 MEAN RATING	BLOCK CONSTRUCTION II MEAN RATING
Form as Mass	0,7	0,8
Overall Contour	2,0	2,0
Sub-Units of Form	1,9	1,9
Perceptual Constancy	0,8	1,1
Directionality	1,9	1,9
Space Relations	3,66 **	3,73 **
Position in Space	1,1	1,1
Figure-Ground	0,7	0,8
Colour	0,0	0,0
Praxis	Essential	Essential

**a better than average assessment.

12.8.2. Construct Validity

12.8.2.1. Age Differentiation

As already stated in this thesis some perceptual skills deteriorate with age. Some tests were sensitive to this age differential and others were not. In the case of the Block Construction tests of the V-P-M battery only the Block Construction I test i.e. Block Construction from a three dimensional model was sensitive to this differential, with only a probably significant difference (P = <0,05) being found between age groups twenty to sixty-nine and seventy to seventy-nine years. The Block Construction II test did not appear to be a valid measure of changes occurring with age.

12.8.2.2. Correlation with Other Tests in the Battery

The highest correlation of the Block Construction I test was with the Design Construction test (r = 0,64). This was expected as both tests were compiled to measure spatial relationships. However, the correlation with the Space Relations test was only r = 0,31. This was probably due to the motor reduced nature of the Space Relations test.

The correlation of the Block Construction II test with the Position in Space test was r = 0,48, the test with which it was expected to correlate. This correlation is low, again probably due to the motor reduced nature of the Position in Space test.

It was interesting to note that the correlation between the two Block Construction tests was only r = 0,22 (see correlation matrix in Chapter Twenty-One) indicating the possibility of different skills inherent in each.

A factor analysis was carried out on the tests in the battery to determine whether tests had factors related to each other. Block Construction II is strongly loaded toward Factor I which seems to be a space factor (see table 12.30).

TABLE 12.30 SORTED ROTATED FACTOR LOADINGS - BCI AND BCII

TEST	FACTOR I	FACTOR II	FACTOR III
Block Construction I	0,28	0,48	0,42
Block Construction II	0,74	0,00	0,00

The Block Construction I test has loadings below 0,50 and would thus appear to belong to none of the existing factors. As the factor analysis was an exploratory analysis and a no marker variables were included, the results can not be readily used for construct validity.

12.8.2.3. Results of Testing Patients

A statistical difference was found between the scores of patients and normal subjects in the Block Construction I test (twenty to sixty-nine years) ($P<0,01$) and in the Block Construction II test (twenty to seventy-nine years) ($P<0,05$). It would seem, therefore, that both tests are a valid way of determining dysfunction in V-P-M skills in the C.V.A. patient.

12.9. CONCLUSIONS

The following was found on the Block Construction I test:-

- It seemed that this test was a fairly easy test to complete judging by the levels of the item difficulty p values. However, it did show a drop off in the ability of normal elderly subjects at age seventy, therefore, these subjects found the test more difficult to complete than the younger group of subjects;

- As in all other tests the normal elderly and the patients took longer to complete the test;

- The test was sensitive to changes that took place as a result of a C.V.A., there being a statistically significant difference in the performance of the patient and the normal group twenty to sixty-nine years. In the seventy year old group several patients were unable to complete the test;

- The split-half reliability, although significant was low and the Kuder-Richardson formula only approached an acceptable level. It seemed, therefore, that apart from inter-rater reliability, the Block Construction I test was not a reliable test;

- The therapists evaluated the content only as a better than average measurement of space relations;

- The Block Construction I test did not appear to fall into any of the three factors identified in the factor analysis and its highest correlation was with the Design Construction test. It has very low correlations with the Space Relations test. It would seem, therefore, that this test was not a good measure of the skill that it was constructed to measure.

The Block Construction II test demonstrated the following characteristics:-

- It was also a fairly easy test to complete, although the difficulty levels were lower than those of the Block Construction I test. In contrast to the Block Construction I, this test showed no difference between the performance of the different age groups. However, the test was sensitive to changes that took place as a result of a C.V.A. but the difference between normal subjects and patients was only at the 5% level of significance;

- The Block Construction II test was shown to be far more reliable than the Block Construction I test with acceptable levels of reliability;

- Although constructed to evaluate position in space (space visualisation), the therapists rated it extremely low on this skill. According to them, the Block Construction II test was a test of space relations. The manner in which the model was depicted for the patient to copy, although different to that of the Block Construction I test, did not change the skill being measured in the opinion of the therapists;

- This test did not correlate highly with the position in space test, probably because of the differences in the motor component. The Block Construction II test, unlike the Block Construction I test was strongly loaded for factor I which was thought to be a space factor. The two Block Construction tests also did not correlate well with each other;

Taking all the foregoing into account it would seem that neither Block Construction tests have strong and consistent indicators for their retention in their present form in the battery apart from the fact that both tests highlighted some of the problems known to be experienced by patients with spatial problems.

Block Construction tests form part of most perceptual assessment batteries and it seems, therefore, that further research must be carried out to find an acceptable Block Construction test for the V-P-M battery. Either the Block Construction II test should be retained and improved or consideration should be given to including the Block Construction used by Benton et al (1983) i.e. the Three Dimensional Construction Praxis test, in any future version of the V-P-M battery.

CHAPTER 13

THE STICK CONSTRUCTION TESTS

13.1. INTRODUCTION

Unlike the Block Construction tasks, construction using sticks is used relatively rarely as a testing medium. However, construction of a design or pattern with sticks does enable the spatial components of perception to be tested on a two dimensional level instead of on a three dimensional level as in the Block Construction tasks.

Stick Construction tasks were devised by the candidate for the purpose of evaluating spatial relationships and the link between spatial skills and motor output at the two dimensional level of space concepts. Critchley (1953) described Stick Construction as a way of evaluating constructional apraxia which was preferable to a drawing task as the "*... the patients variations and changes in mind are more in evidence to the close observer*" (Page 180).

A description of the Stick Construction I test is to be found in Chapter Four Section 4.6.12. and Appendix 1 Section 2.10.3. and that of the Stick Construction II test in Chapter Four Section 4.6.13. and Appendix 1 Section 2.10. 4.

13.2. REVIEW OF SIMILAR TESTS

The usual format for Stick Construction is a copying task. Critchley (1953) described an unstandardised Stick Construction test containing sticks of different lengths. According to Lezak (1983), Fogel also described a test of copying designs with sticks. The most used Stick Construction tests were those constructed by Benson and Barton (1970) and by Goodglass and Kaplan (1972). Goodglass and Kaplan (1972) used the Goldstein and Sheerer (1953) test as the basis for their own (see Figure 13.1). Although the designs of the Goodglass and Kaplan test are suitable for the evaluation of perceptual function as described in this thesis, the memory component in the execution of the test

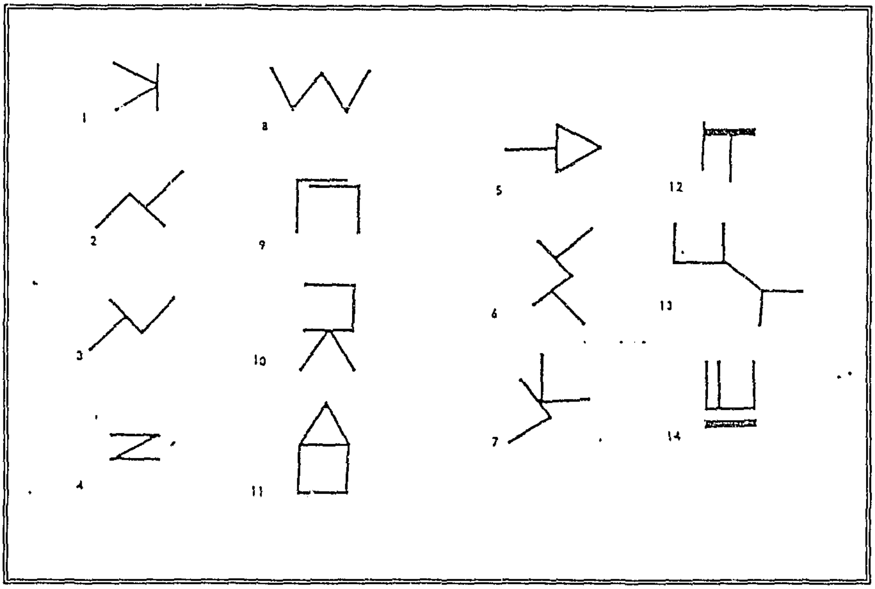


FIGURE 13.1 THE GOODGLASS AND KAPLAN STICK CONSTRUCTION TEST

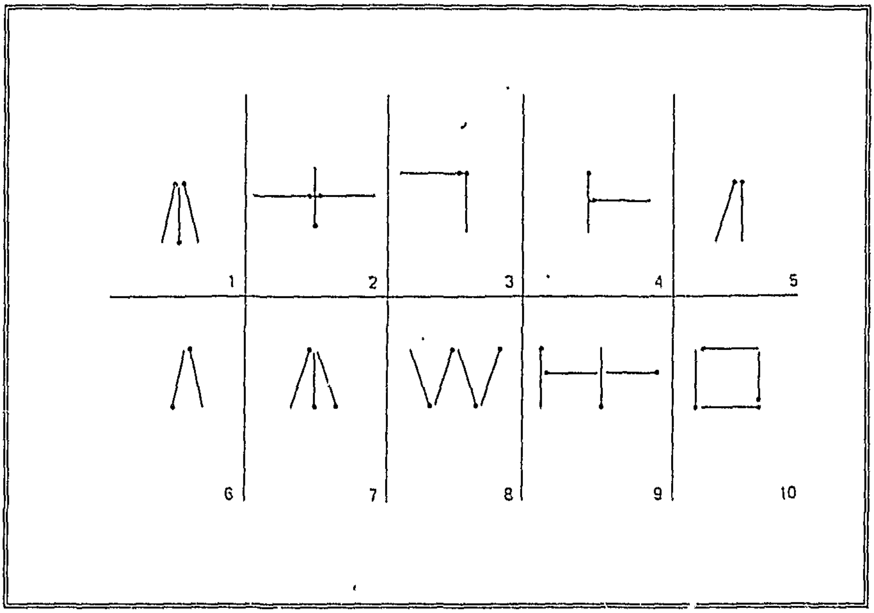


FIGURE 13.2 THE BUTTERS AND BARTON STICK CONSTRUCTION TEST

made it unsuitable for the V-P-M battery. There were also no normative data available for this test.

The Stick Construction tests of Benson and Barton (1970) and Butters and Barton (1970) contained designs that could have been used in the V-P-M battery (see Figure 13.2). Sticks constructed to resemble matches were used for this purpose. Again there were no normative data but some information was available on the different performance of patients with left and right hemisphere lesions as well as on a technique requiring the subject to rotate the design (Benson and Barton 1970, Butters and Barton 1970).

Hecaen and Assal (1970 : 293) used *"reproduction of a series of non-figurative and figurative models"* with zigzags of three, four and five elements as well as a complex form. Their test also required the reproductions of a cube model, a house and a star of eight points. There were no normative data for this test although control subjects, whose performance was compared with that of left and right brain damaged patients, produced copies free of errors.

13.3. Reasons For Compiling New Tests

The Stick Construction designs used by Benson and Barton (1970), Hecaen and Assal (1970) and Goodglass and Kaplan (1972), could have been used for the Stick Construction test in this thesis. However, as none of these tests offered any normative data and little information was available on them, it was thought to be appropriate to develop a new test. Designs were therefore constructed using two types of models for the subject to copy i.e. a model of the sticks themselves (Stick Construction I test) and that of a design depicted on a card (Stick Construction II test).

13.4. Results Of Testing Normal Individuals

13.4.1. Accuracy Scores - Stick Construction I Test

The means and the standard deviations of the accuracy scores were calculated for the normal group as a whole (Table 13.1) and for each age group individually (Table 13.2).

TABLE 13.1 MEAN AND SD OF ACCURACY SCORES - SCI

n	MAX SCORE	MEAN	SD	RANGE
118	12	11,36	1,03	7-12

TABLE 13.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - SCI

AGE	n	MEAN	SD	RANGE
20-29	20	11,55	0,51	11-12
30-39	21	11,71	0,56	10-12
40-49	19	11,26	1,33	7-12
50-59	19	11,42	0,96	9-12
60-69	19	11,26	1,10	9-12
70-79	20	10,90	1,37	7-12

The Pearson Chi-square test was used to determine whether differences existed in the performance within the group as a whole. A cut point of eleven was used, with the percentage obtaining a score greater than and less than the cut point being used to calculate the Chi-square value (Table 13.3).

TABLE 13.3 COMPARISON OF ACCURACY SCORES - SCI

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 11	45,0	23,8	36,8	36,8	36,8	55,0	39,1
Greater than 11	55,0	76,2	63,2	63,2	63,2	45,0	61,0

As there were no significant differences among the scores obtained by the different age groups, no further analysis was carried out.

13.4.2. Accuracy Scores - Stick Construction II Test

The means and the standard deviations were calculated for the normal group as a whole (Table 13.4) and for each group separately (Table 13.5).

TABLE 13.4 MEAN AND SD FOR ACCURACY SCORES - SCII

n	MAX SCORE	MEAN	SD	RANGE
118	12	10,26	1,82	3-12

TABLE 13.5 MEAN AND SD FOR ACCURACY SCORES FOR EACH AGE GROUP - SCII

AGE	n	MEAN	SD	RANGE
20-29	20	10,75	1,29	7-12
30-39	21	10,86	1,15	8-12
40-49	19	10,68	1,38	7-12
50-59	19	10,26	1,45	6-12
60-69	19	9,58	2,43	4-12
70-79	20	9,40	2,44	3-12

The Pearson Chi-square test was used to determine whether differences existed in the accuracy scores obtained by the different age groups. A cut point of ten was used, with the percentage of subjects obtaining a score greater or less than the cut point being used to calculate the Chi-square value (Table 13.6).

Again no differences were found in the performance of the different age groups therefore the null hypothesis which stated that decreased ability will not be found with advancing age in visual-perceptual-motor function, was retained in the case of both Stick Construction tests. Percentiles were calculated for the accuracy scores in the age groups determined for the other tests in the battery(see Appendix 2).

TABLE 13.6 COMPARISON OF ACCURACY SCORES - SCII

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 10	40,0	38,1	31,6	57,9	47,4	60,0	45,8
Greater than 10	60,0	61,9	68,4	42,1	52,6	40,0	54,2

Pearson Chi-square Value = 5,085 not significant P = 0,4056

13.4.3. Time Scores - Stick Construction I Test

The means and the standard deviations were calculated for the time scores for the normal group as a whole (Table 13.7) and for each age groups separately (Table 13.8).

TABLE 13.7 MEAN AND SD OF TIME SCORES - SCI

n	MEAN/SECS	SD	RANGE
118	74,19	26,15	19-184

TABLE 13.8 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - SCI

AGES	n	MEAN/SECS	SD	RANGE
20-29	20	65,2	21,99	36-114
30-39	21	64,43	21,24	37-131
40-49	19	64,90	15,60	19-89
50-59	19	68,58	17,92	39-100
60-69	19	84,21	25,32	48-144
70-79	20	98,05	33,29	50-184

It was interesting to note that the age group forty to forty-nine obtained the lowest time scores in this test unlike other tests where they appeared to take longer than younger groups.

Levine's test for equal variance was not significant (P=0,1054), therefore, the variances between the groups were equal and the F test was used to determine whether the means differed from one another. The means were found to be significantly different (P=0,0001), therefore, the separate variance t with the Bonferroni correction for multiple comparisons was used to establish where the differences were (Table 13.9).

TABLE 13.9 DIFFERENCES IN TIME SCORES - SCI

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,9158
20-29	40-49	0,9675
20-29	50-59	0,6518
20-29	60-69	0,0133
20-29	70-79	0,0000 ***
30-39	40-49	0,9498
30-39	50-59	0,5751
30-39	60-69	0,0085
30-39	70-79	0,0000 ***
40-49	50-59	0,6272
40-49	60-69	0,0130
40-49	70-79	0,0000 ***
50-59	60-69	0,0411
50-59	70-79	0,0001 **
60-69	70-79	0,0556

* P<0,05 ** P<0,01 *** P<0,001

As can be seen from Table 13.9 differences occurred in the time scores obtained by the age groups twenty to twenty-nine, thirty to thirty-nine, forty to forty-nine, fifty to fifty-nine and seventy to seventy-nine. The data were, therefore, regrouped and the procedures repeated.

As Levine's test demonstrated significant differences in the variances between the groups (P = 0,0385), the Welch Test was used to determine if there were any differences between the newly established groups. This test showed that there were significant differences (P = 0,0003) and the separate variance t with the Bonferroni correction was used to determine where these differences were to be found (Table 13.10).

TABLE 13.10 DIFFERENCES IN RE-GROUPED TIME SCORES - SCI

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-59	0,6540
20-39	60-69	0,0069 *
20-39	70-79	0,0004 **
40-59	60-69	0,0113
40-59	70-79	0,0006 **
60-69	70-79	0,1510

* P<0,05 ** P < 0,01

Thus the ages twenty to fifty- nine years and sixty to seventy-nine year could be grouped together. Percentiles were calculated for the regrouped ages as already described (see Appendix 2).

13.4.4. Time Scores - Stick Construction II Test

The means and standard deviations were calculated for the time taken to complete the test for the group as a whole (Table 13.11) and for each group separately (Table 13.12).

TABLE 13.11 MEAN AND SD OF TIME SCORES - SCII

n	MEAN/SECS	SD	RANGE
118	84,80	37,55	40-388

TABLE 13.12 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - SCII

AGE	n	MEAN/SECS	SD	RANGE
20-29	20	69,75	17,88	41-113
30-39	21	72,52	19,80	60-108
40-49	19	76,16	13,89	60-98
50-59	19	80,63	21,02	54-150
60-69	19	89,95	29,16	42-116
70-79	20	130,0	68,86	57-388

The same procedure for analysing the scores was used as that used for the Stick Construction I time scores. Levine's test for equal variance was significant ($P = 0{,}0133$) therefore the Welch Test was used to establish if there were any differences in the scores ($P = 0{,}0154$). As there were differences in the scores the separate variance t with the Bonferroni correction was used to determine where the differences were (Table 13.13).

It was interesting to note that there were no differences in the time scores obtained by the different age groups on the Stick Construction II test. None the less percentiles were calculated for the different age groups as previously described (see Appendix 2).

However, once the data were regrouped the above mentioned procedures were carried out and a difference was then found in the scores obtained by the twenty to thirty-nine year olds and the seventy to seventy-nine year olds at $P<0{,}05$ level of significance (Table 13.14).

TABLE 13.13 DIFFERENCES IN TIME SCORES - SCII

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,6401
20-29	40-49	0,2050
20-29	50-59	0,0911
20-29	60-69	0,0147
20-29	70-79	0,0046
30-39	40-49	0,4921
30-39	50-59	0,2184
30-39	60-69	0,0363
30-39	70-79	0,0071
40-49	50-59	0,4353
40-49	60-69	0,0713
40-49	70-79	0,0110
50-59	60-69	0,2668
50-59	70-79	0,0229
60-69	70-79	0,0851

TABLE 13.14 DIFFERENCES IN RE-GROUPED TIME SCORES - SCII

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-59	0,0788
20-39	60-69	0,0164
20-39	70-79	0,0054 *
40-59	60-69	0,1343
60-69	70-79	0,0851

* P< 0,05

13.4.5. Discussion Of The Results Of Testing Normal Individuals

It is noteworthy that neither the Stick Construction I test nor the Stick Construction II test reflected age related changes in the accuracy scores. The only other study in which the performance of the elderly was compared with that of the younger age groups was carried out by Farver and Farver (1982) when they used the Stick Construction tasks from the Boston Diagnostic Aphasia Examination (Goodglass and Kaplan 1972). Farver and Farver found a statistically significant age effect ($P < 0,05$) when testing normal subjects between the age of forty and eighty-nine years. This difference could be due to the fact that the Goodglass and Kaplan stick designs (See Figure 13.1) were more detailed and complicated than the figures used for this thesis; in addition persons older than seventy-nine years were included in their sample; and the subjects were required to reproduce the design from memory. The results were therefore not directly comparable.

13.5. RESULTS OF TESTING PATIENTS

13.5.1. Accuracy Scores - Stick Construction I Test

The mean and the standard deviation was calculated for the age groups between which no statistical difference was found when testing normal subjects. In the case of the Stick Construction I Test, all age groups were combined (Table 13.15).

TABLE 13.15 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - SCI

AGE	n	MEAN	SD	RANGE
20-79 yrs	24	9,04	2,81	0-13

The number of patients evaluated on this test was less than the total due to the discharge of patients before testing was completed (six) and the fact that three patients suffered from severe problems which rendered them untestable.

13.5.2. Accuracy Scores - Stick Construction II Test

The mean and the standard deviation was calculated as mentioned above. The results are seen in Table 13.16.

TABLE 13.16 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - SCII

AGE	n	MEAN	SD	RANGE
20-79 yrs	24	6,04	9,92	0-11

13.5.3. Time Scores - Stick Construction I Test

The means and the standard deviations were calculated for the time scores obtained by the age groups between which no statistical difference was found in the scores obtained by the normal group i.e. the twenty to fifty-nine year olds and the sixty to seventy-nine year olds (Table 13.17).

TABLE 13.17 MEAN AND SD FOR TIME SCORES FOR PATIENTS - SCI

AGE	n	MEAN/SECS	SD	RANGE
20-59 yrs	15	116,64	36,58	76-218
60-79 yrs	9	201,80	137,14	103-501

13.5.4. Time Scores - Stick Construction II Test

The means and the standard deviations (Table 13.18) were calculated for the Stick Construction II test as described in item 13.5.3.

TABLE 13.18 MEAN AND SD FOR TIME SCORES FOR PATIENTS - SCII

AGES	n	MEAN/SECS	SD	RANGE
20-59 yrs	15	138,5	49,45	66-247
60-79 yrs	9	201,20	88,24	101-383

13.5.5. Comparison Between Patients and Normal Subjects

13.5.5.1. Accuracy Scores - Stick Construction I and II Tests

Comparisons were made between the normal group as a whole and the entire group of patients as no age related differences were found in the normal group's accuracy scores. Levine's test for variances between the groups was found to be significant (P = 0,0000) in both instances, therefore, the Welch Test was used to establish what the differences were (Table 13.19).

TABLE 13.19 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - SCI AND SCII

TEST	AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
SC I	20-79	0,0000	0,0009 ***
SC II	20-79	0,0000	0,0000 ***

*** P < 0,001

As can be seen from Table 13.19, both tests showed highly significant differences between the scores obtained by the normal group and those obtained by the patients.

13.5.5.2. Time Scores - Stick Construction I and II Tests

Comparisons made between the different age groups showed a difference in the scores obtained by the normal group of subjects and the patients. The results are reflected in Table 13.20.

TABLE 13.20 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - SCI AND SCII

TEST	AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
SC I	20-59 60-79	0,0139 0,0001	0,0002 *** 0,0233 *
SC II	20-59 60-79	0,0000 0,0195	0,0013 ** 0,0078 **

* P < 0,05 ** P < 0,01 *** P < 0,001

13.5.6. Analysis Of Tests Executed By Patients

13.5.6.1. Stick Construction I Test

a) In the age group twenty to thirty-nine years two patients received scores below the fifteenth percentile, therefore, demonstrating a problem with spatial relations, three scored above and two were discharged prior to completing the test.

b) In the forty to fifty-nine year age group only three patients received scores above the fifteenth percentile. Six patients demonstrated problems as they received scores below the fifteenth percentile and one was discharged prior to completing the test.

c) In the age group sixty to sixty-nine years, four patients received scores above the fifteenth percentile, four below and two were discharged.

d) In the seventy to seventy-nine year age group, the two patients who completed the test obtained scores below the fifteenth percentile and therefore demonstrated problems, three were untestable and one was discharged prior to completing the test.

13.5.6.2. Stick Construction II Test

a) In the age group twenty to thirty-nine years, three patients received scores above the fifteenth percentile, two below therefore demonstrating a problem with position in space and one was discharged prior to completing the test.

b) In the forty to fifty-nine year age group, six patients received scores at or above the fifteenth percentile, three below and one was discharged.

c) In the sixty to sixty-nine year age group, three patients received scores above the fifteenth percentile, five below and two were discharged.

d) The two patients who completed the test in the seventy to seventy-nine year age group demonstrated problems with position in space as they received scores below the fifteenth percentile, three were untestable and one discharged.

As in the Block Construction tests, the patients appeared to find the construction of stick designs from the card (Stick Construction II test) more difficult than the Stick Construction I test. Only two patients received a score that was higher in the Stick Construction II test than in the Stick Construction I test. In two cases the score was identical and in nineteen cases the score for test I was greater than in test II. This was to be expected as the Stick Construction II tasks involved some space visualisation which is thought to be a higher level space task than concrete spatial relations.

It can be seen from Table 13.20 that the level of significance of the differences between the patient and the normal group varied between P < 0,05 and P < 0,001.

13.5.7. Problems That May Be Diagnosed From Performance On The Stick Construction Tests

As in the Block Construction tests, an analysis of the mistakes made by the patients was more difficult than that of paper and pencil tasks. An attempt was made to draw the patient's constructions, while the patient was in the process of constructing the design. The Stick Constructions were easier to record than those of block design as they were easier to draw quickly.

The following types of errors were recorded in both Stick Construction tests :-

a) There was much trial and error i.e. the patient placed a stick and continued to move it until he was satisfied with its placement. This was seen in the performance of patient sixteen. It is the candidate's opinion that this could be a sign of apraxia. However in the case of patient sixteen an intention tremor was the cause.

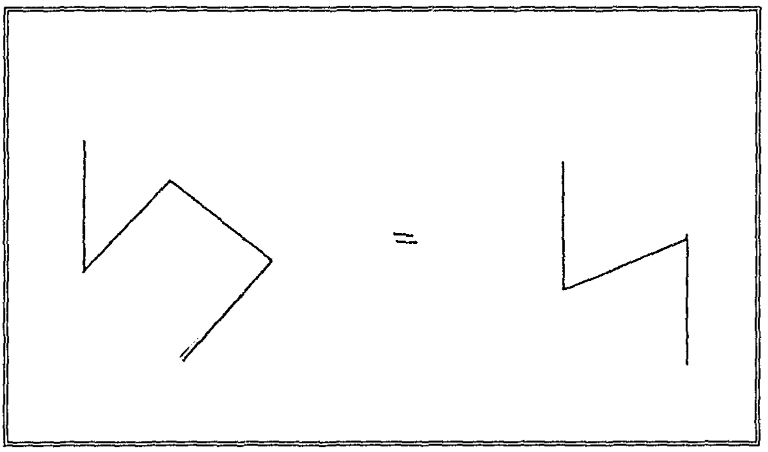


FIGURE 13.3 LEAVING OUT PARTS OF DESIGN

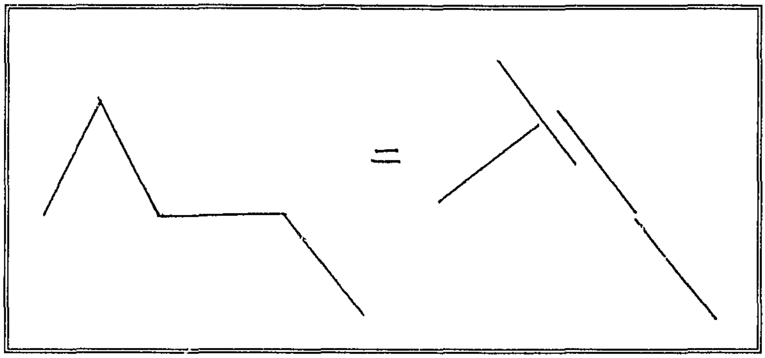


FIGURE 13.4 SPATIAL MISPLACEMENT

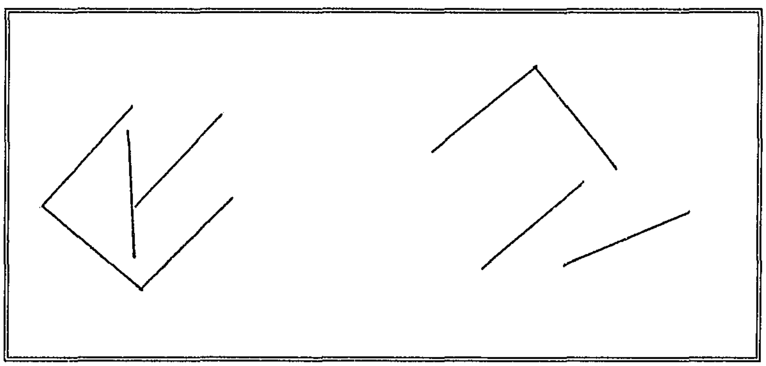


FIGURE 13.5 SPATIAL MISPLACEMENT

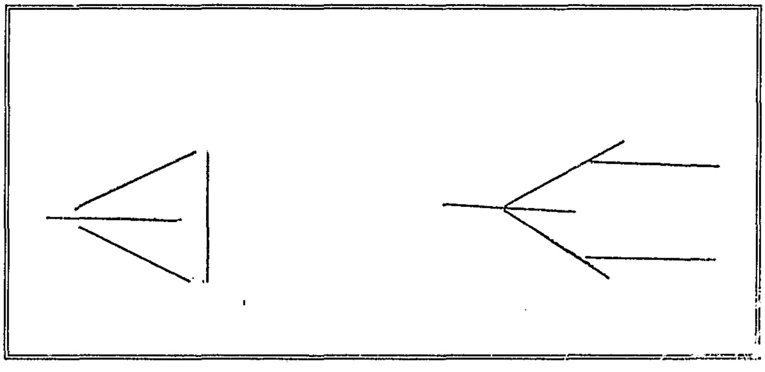


FIGURE 13.6 SPATIAL MISPLACEMENT

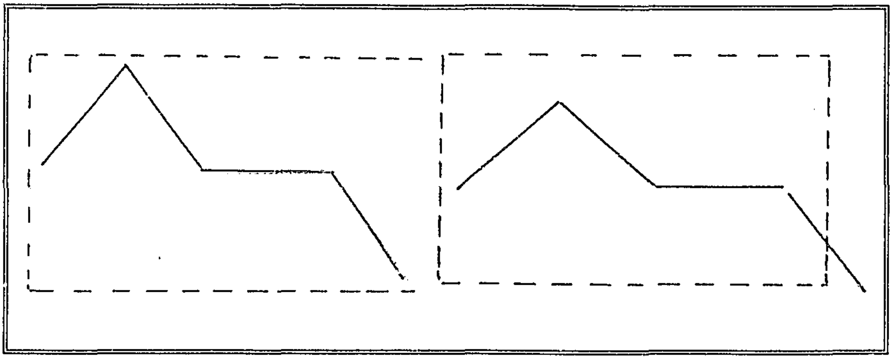


FIGURE 13.7 INABILITY TO USE ALLOCATED SPACE FOR CONSTRUCTION

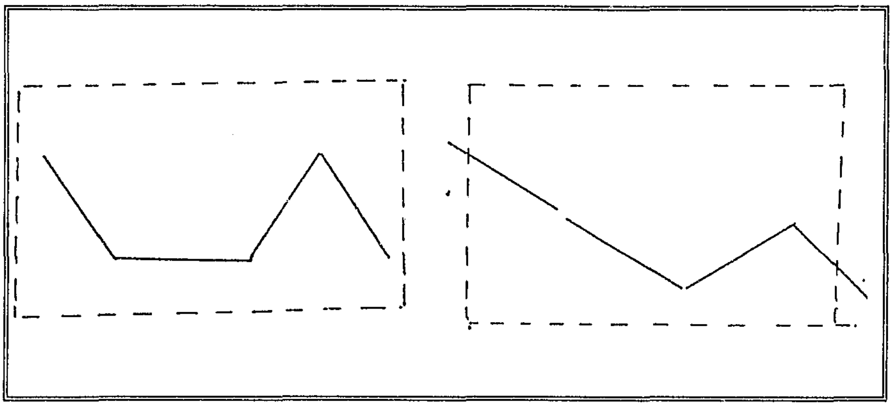


FIGURE 13.8 SEPARATION OF STICKS

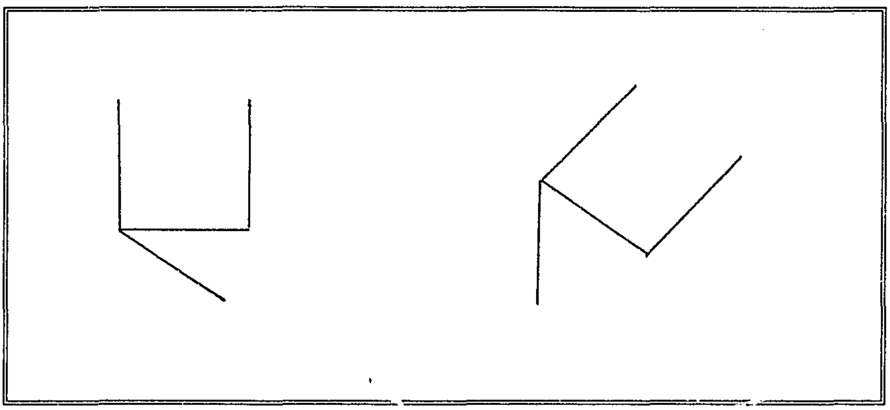


FIGURE 13.9 ROTATION OF DESIGN

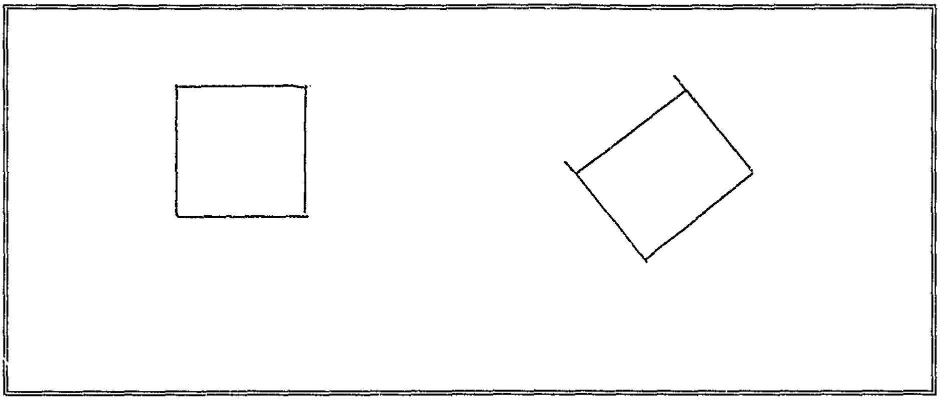


FIGURE 13.10 ROTATION OF DESIGN

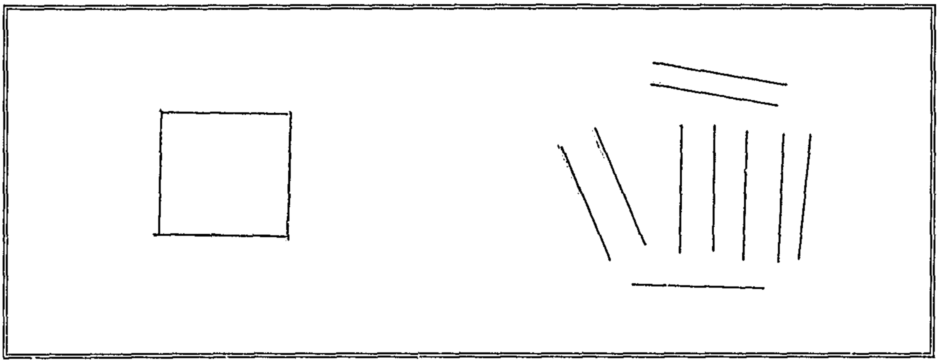


FIGURE 13.11 PERSEVERATION

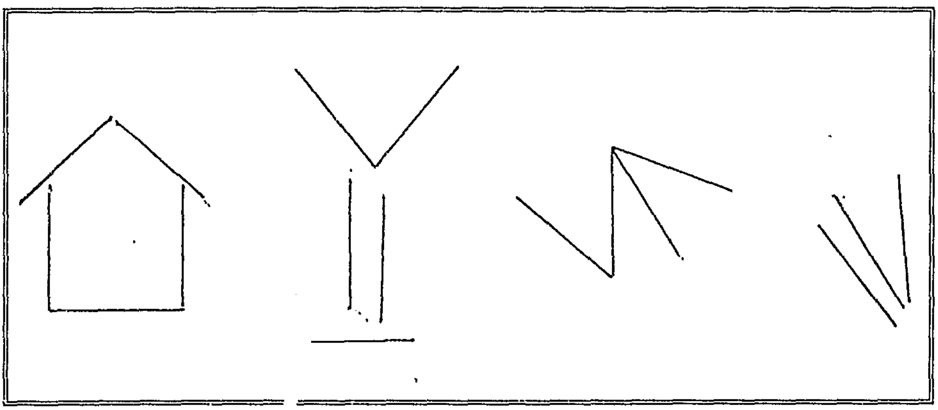


FIGURE 13.12 CONSTANTLY CHANGING DESIGN

b) Leaving out parts of a design (patient twenty-seven) (see figure 13.3). This could be an indication of visual inattention. According to Critchley (1953) it is also a symptom of constructional apraxia.

c) Spatial misplacement of sticks which ranged from gross misplacement seen in patients with severe spatial problems (patient two and ten) (Figures 13.4, 13.5, 13.6) to the less severe spatial inability to place the sticks within the bounds of the felt construction board which was of the same size as the stimulus card (patients fifteen, seventeen, and twenty) (Figures 13.7 and 13.8). Other mild spatial deficits were seen in the separation of the sticks.

d) Rotation of designs which Critchley labelled as a symptom of constructional apraxia and which also demonstrated spatial problems (patient twenty-seven) Figures 13.9 and 13.10.

e) Perseveration was very easily observed in the performance of patient nine (Figure 13.11). According to the candidate this is a strong indicator of apraxia.

f) Constantly changing the construction could also be a sign of apraxia (Figure 13.12).

The designs depicted in Figures 13.3 to 13.12 were taken from the drawings made by the examiner of each patient's construction. The figure on the left represents the model while the figure on the right depicts the patients' construction.

Stick Construction tasks appeaedr to be tasks in which spatial problems were easily identified.

13.6. Item Analysis

13.6.1. Difficulty Index

The difficulty index of the items in both the Stick Construction tests can be seen in Table 13.21. The scores from the normal group of subjects was used for this purpose.

It seemed that the items of the Stick Construction II test (i.e. Stick Construction from a design printed on a card) were slightly more difficult than those of the Stick Construction I test. Re-ordering of the items for the Stick Construction I test should be done in the following way:- four, one, two, six, five, three and for the Stick Construction II test, six, three, four, five, two and one.

TABLE 13.21 DIFFICULTY INDEX - SCI AND SCII

ITEM	DIFFICULTY p 20 - 79 YRS	
	SC I	SC II
1	0,95	0,64
2	0,92	0,71
3	0,77	0,80
4	0,96	0,79
5	0,82	0,75
6	0,91	0,81

The difficulty index for the responses obtained from the patient group was also calculated (Table 13.22).

The Pearson Chi-square Test was used to calculate whether there were statistical differences between the normal subjects and the patient group in the responses obtained for each item. As can be observed in Table 13.22 all of the items in the two tests showed a significant difference ($P < 0,001$) between the scores obtained by the patient group and those obtained by the normal group.

TABLE 13.22 DIFFICULTY INDEX - PATIENT GROUP - SCI AND SCII

ITEM	DIFFICULTY P 20 79 Yrs	
	SC I	SC II
1	0,71 ***	0,17 ***
2	0,67 ***	0,29 ***
3	0,46 ***	0,29 ***
4	0,79 **	0,42 ***
5	0,46 ***	0,42 ***
6	0,75 *	0,42 ***

* $P < 0,05$ ** $P < 0,01$ *** $P < 0,001$

The difficulty index of the patient scores on both tests indicated that the re-ordering of the items would be similar to the order of difficulty for normal subjects.

13.6.2. Item Discrimination

It was difficult to determine the discriminatory level of each item as the difficulty p of each item was high and cut points could not be established as in the case of the Block Construction tests.

13.7. TEST RELIABILITY

13.7.1. Split-Half Reliability

This determines the internal consistency of the test. The resultant co-efficients of correlation for the Stick Construction tests, age group twenty years to seventy-nine years, can be seen in Table 13.23. The Split-Half reliability co-efficients were r = 0,56 for the Stick Construction I test and r = 0,48 for the Stick Construction II test. This demonstrated a low degree of correlation between the two halves of the test.

TABLE 13.23 RELIABILITY CO-EFFICIENTS - SCI AND SCII

TEST	AGE	SPLIT-HALF	KUDER- RICHARDSON
SC I	20-79	0,56	0,56
SC II	20-79	0,48	0,80 *

* P < 0,05

13.7.2. Kuder-Richardson Formula 20

Using the Kuder-Richardson Formula 20, a co-efficient of r = 0,56 was obtained for the Stick Construction I test. This was the same co-efficient as that obtained in the Split-half calculation. The Kuder-Richardson formula 20 gave a much higher co-efficient of correlation i.e. r = 0,80 for the Stik Construction II test than did the Split-half correlation.

None of the reliability co-efficients except for the Kuder-Richardson reliability co-efficient for the Stick Construction II test showed statistical significance and this co-efficient was only significant at the P < 0,05 level of significance. The Stick Construction tests therefore did not prove to have internal consistency.

13.7.3. Inter-Rater Reliability

The Friedman two-way analysis of variance showed that there was no statistical difference between either the accuracy or time scores obtained by three examiners on each of the six individuals tested. In the Stick Construction I test P = 0,8825 for the accuracy scores and P = 0,1398 for the time scores. In the Stick Construction II test P = 0,3114 for the accuracy scores and P = 0,4169 for the time scores.

13.8. TEST VALIDITY

Two types of validity were determined.

13.8.1. Content Validity

Content validity data were provided by fifteen occupational therapists (group one) who rated the tests on a 0-5 scale, according to how well they thought the tests measured the various defined skills.

Both tests were only rated as a "better than average" assessment of space relations for which only the Stick Construction I test was compiled. The stick Construction II test was intended to measure position in space but was rated low for this skill by the therapists.

TABLE 13.24 CONTENT VALIDITY - SCI AND SCII

PERCEPTUAL SKILL	SC I Mean Rating	SC II Mean Rating
Form as Mass	0,7	0,7
Overall Contour	2,0	2,0
Sub-Units of Form	1,6	1,7
Perceptual Constancy	0,9	1,0
Directionality	3,0 **	2,8
Space Relations	3,3 **	3,5**
Position in Space	1,3	1,4
Figure Ground	0,6	0,7
Colour	0,0	0,0
Praxis	required	required

** a better than average assessment of skill

13.8.2. Construct Validity

Three types of construct validity were considered.

13.8.2.1. Age Differentiation

Discussion already presented has indicated that some perceptual skills deteriorate with age. Both Stick Construction tests ere not sensitive to these changes and therefore they were not a valid measure of age differentials.

13.8.2.2. Correlation With Other Tests In The Battery

The Stick Construction I test correlated best with the Dot-Line test (r = 0,51). This corresponded with the therapists rating which indicated that the test demanded directional ability (Chapter Twenty One).

The Stick Construction II test correlated with the Design Construction test (r = 0,57) (which was formulated to evaluate spatial relations) and the Dot-Line test (r = 0,53), (see Table 13.25).

The correlation between the two Stick Construction tests was r = 0,47. This was a fairly low degree of correlation considering the nature of the two tests.

A factor analysis was carried out to determine whether the individual tests had factors related to each other (Table 13.26).

The Stick Construction I test had a low loading for Factor I which appeared to measure spatial ability. There was no loading on Factor II which seemed to be a visual discrimination or figure-ground factor (see Table 13.26). The highest loading was on Factor III. The other tests with a loading on this factor seemed to need either colour recognition or matching (Colour Matching test r = 0,81 and Design Construction r = 0,40). However, as all loadings of 0,50 or less should be ignored, it was apparent that the Stick Construction I test did not clearly belong to any of the factors identified in the V-P-M battery. The Stick Construction II test loaded for factor II (r = 0,60) which appeared to be a figure-ground visual discrimination factor.

TABLE 13.25 CORRELATION WITH OTHER TESTS IN THE BATTERY - SCI AND SCII

TEST	SC I	SC II
Arrows	0,14	0,37
VMI	0,36	0,42
Block Construction I	0,37	0,46
Block Construction II	0,30	0,27
Position in Space	0,36	0,23
Space Relations	0,21	0,39
Design Construction	0,47	0,57
Dot-Line	0,51	0,53
Figure-Ground	0,36	0,40
Perceptual Constancy	0,37	0,46
Part-Whole	0,47	0,36
Colour	0,21	0,21

TABLE 13.26 SORTED ROTATED FACTOR LOADINGS - SCI AND SCII

TEST	FACTOR I	FACTOR II	FACTOR III
Stick Construction 1	0,46	0,0	0,53
Stick Construction II	0,0	0,60	0,46

As the Stick Construction tests did not clearly load in appropriate factor categories, careful consideration must be given to their retention in future versions of the test.

13.8.2.3. Results of Testing Patients

An almost certainly significant difference was found between the scores obtained by the patients and the normal subjects on both the Stick Construction I test (P<0,001) and the Stick Construction II test (P <0,001). It would seem, therefore, that both tests were valid measures of some type of V-P-M dysfunction in C.V.A. patients. The question remains, however, as to exactly what that function is. The analysis ofpatients test results' also showed that the test was an effective way of making a clinical diagnosis of some types of V-P-M dysfunction e,g. apraxia.

13.9. CONCLUSIONS

Taking all the foregoing into account it would seem that the Stick Construction tests should be dropped from the V-P-M battery for the following reasons:-

- The reliability co-efficients were low.
- The tests were not sensitive to age related changes in perceptual function.
- They did not appear to load with an appropriate perceptual factor.
- They correlated with other tests in the battery i.e. Dot-Line test, which appeared to be a measure of a similar function.
- They did not have high content validity as evaluated by the fifteen occupational therapists.

It was apparent, that despite the fact that differences between normal and patient performance were highly significant and that performance on the tests were an excellent way of clinically demonstrating some spatial perceptual problems, the psychometric characteristics are not robust enough to justify the use of the Stick Construction tests.

Future research should investigate the value of the Stick Construction tests described by Goodglass and Kaplan (1972) as a measure of two dimensional constructional ability as this test contains more complicated figures than that of the V-P-M Battery.

CHAPTER 14

THE DESIGN CONSTRUCTION TEST

14.1. INTRODUCTION

Depth perception is used to interpret the three dimensional world and is needed to enable us to move around successfully. It is difficult to evaluate depth perception as large pieces of equipment are needed to test this skill successfully. The Key Stone Telebinocular Apparatus was developed to meet this need but is no longer extensively used. Another means of evaluating depth perception is through block construction tasks which are three dimensional in nature. In order to solve some of the problems of testing depth perception, a two dimensional task based on one of the laws of perspective i.e. overlapping figures was constructed. Overlapping figures cue the individual to those figures that are located in the foreground as compared to those located in the background and are therefore further away. The Design Construction test was, therefore, developed using this principle of perspective to evaluate spatial ability. Depth perception was classified as part of spatial relations and belongs to the subjective group of space constructs.

A description of the Design Construction test is to be found in Chapter Four Section 4.6.14. and Appendix 1 Section 2.10.5.

14.2. REVIEW OF SIMILAR TESTS

Benson and Barton (1970) formulated the Token Pattern Test which consists of two designs *"constructed on 8 1/2" by 11" sheets of paper from small chips of red plexiglass of different shapes"* (Benson and Barton 1970:30). In this test the subject is required to make an exact copy of the examiner's model. No normative data were provided on this test and a limited study of the performance of eight patients was described.

The LOTCA (Itzkovich et al 1990) offers a sub-test similar to the design construction test of the V-P-M battery in which the subject is required to copy a design depicted on a card with corresponding solid plastic shapes. No overlapping or underlapping is, however, possible in this test. Normative data were obtained on fifty-five normal individuals, but though sensitive to problems in patients, the test did not demonstrate age related changes.

14.3. Reasons For Compiling A New Test

As the candidate did not find any test which required the usage of overlapping figures as a means of evaluating depth perception, it was thought useful to provide a task in addition to block construction, to evaluate this perceptual skill. Thus a test was constructed which required the subject to construct a design using thin plastic coloured shapes.

14.4. Results Of Testing Normal Individuals

14.4.1. Accuracy Scores

The means and the standard deviations were calculated for the normal group as a whole (Table 14.1) and for each age group individually (Table 14.2).

TABLE 14.1 MEAN AND SD OF ACCURACY SCORES -DC

n	MAX SCORE	MEAN	SD	RANGE
118	12	10,41	2,06	2-12

TABLE 14.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - DC

Age	n	Mean	SD	RANGE
20-29	20	10,95	1,23	8-12
30-39	21	11,10	1,04	8-12
40-49	19	11,26	0,93	9-12
50-59	19	10,68	1,64	6-12
60-69	19	9,84	2,54	2-12
70-79	20	8,60	2,95	2-12

The Pearson Chi-square test was used to determine if any differences existed in the scores (Table 14.3). A cut point of ten was used for this purpose.

TABLE 14.3 COMPARISON OF ACCURACY SCORES - DC

AGE	20-29	30-39	40-49	50-59	60-69	70-79	TOTAL
Less than 10	30,0	14,3	21,1	36,8	47,4	60,0	34,7
Greater than 10	70,0	85,7	78,9	63,2	52,6	40,0	65,3

The Pearson Chi-square Value = 12,65 P<0,05

As a significant difference was found in the scores obtained by the different age groups, further analysis was carried out to determine where these differences were to be found as follows:

a) The Pearson Chi-square Test revealed that there were no differences in the scores obtained for the age groups twenty to twenty-nine, thirty to thirty-nine, forty to forty-nine and fifty to fifty-nine (P=0,3763). These scores were, therefore, combined for further analysis.

b) Fisher's Exact test showed that there were no differences in the scores obtained by the sixty to sixty-nine and seventy to seventy-nine year olds (P = 0,5273). These scores were, therefore, combined for further analysis.

c) Fisher's Exact test demonstrated a significant difference between the scores for the re-grouped age categories (Table 14.4).

   For the sake of continuity, percentiles were calculated for the age groups already used for this purpose i.e. twenty to thirty-nine, forty to fifty-nine, sixty to sixty-nine and seventy to seventy-nine (see Appendix 2).

TABLE 14.4 COMPARISON OF ACCURACY SCORES 20-59 AND 60-79 YEARS - DC

AGE	20-59	60-79	TOTAL
Less than 10	25,3	53,8	34,7
Greater than 10	74,7	46,2	65,3

Fisher's Exact Test P = 0,0037 $P<0,01$

It can be seen from Table 14.4 that there was a significant decrease in the scores obtained by the subjects over sixty years of age. The null-hypothesis that decreased V-P-M function will not be found with advancing age was rejected at the 0,01% level of significance.

14.4.2. Time Scores

The means and the standard deviations were calculated for the time scores for the group as a whole (Table 14.5) and for each group separately (Table 14.6).

TABLE 14.5 MEAN AND SD OF TIME SCORES - DC

n	MEAN/SECS	SD	RANGE
118	119,1	41,15	53-255

TABLE 14.6 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - DC

AGE	n	MEAN/SECS	SD	RANGE
20-29	20	87,85	25,54	58-141
30-39	21	98,86	26,64	61-146
40-49	19	110,47	24,38	53-157
50-59	19	114,42	34,69	69-202
60-69	19	143,89	44,46	78-255
70-79	20	160,70	37,95	111-254

Levine's test for variances between the groups established that there was not a significant difference in the variances (P=0,3860) therefore the F test was used to investigate if there was a difference between the mean time scores obtained by each group. A significant difference was found (P=0,0000) therefore the pooled variance t with the Bonferroni correction for multiple comparisons was used to find out where these differences were (Table 14.7).

The data were re-grouped on the basis of the above findings and the procedure repeated as above. Since Levine's test for equal variances was not significant (P=0,2501) the F test was used to determine any differences. The re-grouped ages were found to differ significantly from one another (P=0,0000) therefore the pooled variance t with the Bonferroni correction for multiple comparisons was used to determine exactly where the differences lay (Table 14.8).

Although no differences existed between the twenty to thirty-nine year olds and the forty to fifty-nine year olds, the scores were kept separate for the compilation of percentiles (see Appendix 2). This was in line with data compiled on the other tests. As a progressive increase in the time taken to complete the test was apparent with advancing age the null hypothesis was rejected.

14.4.3. Discussion of the Results of Testing Normal Individuals

The Design Construction test was compiled to provide additional knowledge of spatial concepts and to measure the possible decrease in spatial ability with advancing age. There were no other tests comparable to the Design Construction test of the V-P-M battery in the literature, therefore, the age related changes found in this test could not be compared with any other tests. The Design Construction test, in line with the overall findings of this study, demonstrated a decrease in scores obtained by the subjects older than sixty years.

TABLE 14.7 DIFFERENCES IN TIME SCORES - DC

AGE		POOLED VARIANCE t P VALUE
20-29	30-39	0,2879
20-29	40-49	0,0345
20-29	50-59	0,0144
20-29	60-69	0,0000 ***
20-29	70-79	0,0000 ***
30-39	40-49	0,2685
30-39	50-59	0,1490
30-39	60-69	0,0000 ***
30-39	70-79	0,0000 ***
40-49	50-59	0,7140
40-49	60-69	0,0023 *
40-49	70-79	0,0000 **
50-59	60-69	0,0069
50-59	70-79	0,0000 ***
60-69	70-79	0,1146

* P<0,05 ** P<0,01 *** P<0,0001

TABLE 14.8 DIFFERENCES IN RE-GROUPED TIME SCORES - DC

AGE		POOLED VARIANCE t P Value
20-39	40-59	0,0118
20-39	60-69	0,0000 ***
20-39	70-79	0,0000 ***
40-59	60-69	0,0009 **
40-59	70-79	0,0000 ***
60-69	70-79	0,1144

** P<0,01 *** P< 0,001

Again the types of mistakes made by the normal elderly individual were similar to those made by the patients, i.e. there were incorrect spacing of the pieces so that they did not fit onto the felt test board, incorrect orientation of the pieces and a great deal of trial and error *used to obtain the correct orientation of the triangles in item two.* Problems were seen in the spatial arrangement of the pieces in relation to each other and with respect to the correct overlapping and underlapping of the pieces.

14.5. RESULTS OF TESTING PATIENTS

14.5.1. Accuracy Scores

The means and the standard deviations were calculated for the age groups between which no statistical difference was found in the scores when testing the normal individual i.e. the twenty to fifty-nine year olds and the sixty to seventy-nine year olds (see Table 14.9). Four patients in the twenty to fifty-nine year old group were discharged prior to completing the test. In the sixty to seventy-nine year old group five patients were untestable, one of these due to severe visual neglect. Two patients were discharged prior to testing.

TABLE 14.9 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - DC

AGE	n	MEAN	SD	RANGE
20-59	14	8,14	2,85	2-12
60-79	8	5,5	3,96	0-10

14.5.2. Time Scores

The means and the standard deviations were calculated for the time scores in the age groups within which no statistical difference was found in the normal subjects i.e. twenty to thirty-nine years, forty to sixty-nine years and seventy to seventy-nine years (Table 14.10). Unfortunately no patients in the seventy to seventy-nine year category were subjected to evaluation by means of this test.

TABLE 14.10 MEAN AND SD FOR TIME SCORES FOR PATIENTS - DC

AGE	n	MEAN/SECS	SD	RANGE
20-39	4	209,75	94,72	10-6295
40-69	17	196,82	51,71	1469-305
70-79	0	-	-	-

14.5.3. Comparison between Patients and Normal Subjects

14.5.3.1. Accuracy Scores

The means of the accuracy scores obtained by the patients were compared with those obtained by the normal subjects to determine whether there was a statistical difference between the two sets of scores.

Levine's test for equal variance proved to be significant in the case of the twenty to fifty-nine year olds (P=0,0000) therefore the Welch Test was used to establish whether there were differences in the means calculated for the two groups (Table 14.11).

In the case of the sixty to seventy-nine year olds (in reality the sixty year olds only) Levine's test was not significant (P=0,0235) therefore the t test was used to determine the differences between the means (Table 14.11).

It can be seen from Table 14.11 that a difference existed between the mean scores obtained by the patients and the normal groups at the P<0,01 level of significance in both age categories.

TABLE 14.11 DIFFERENCES IN ACCURACY SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - DC

AGE	LEVINE'S TEST P VALUE	t TEST P VALUE
20-59	0,0000	0,0023 **
60-79	0,1235	0,0028 **

** P<0,01

14.5.3.2. Time

As with the accuracy scores, the means of the time scores for the normal individuals and the patients were compared to determine whether there were any statistical differences. In the case of the twenty to thirty-nine year olds and the forty to sixty-nine year olds Levine's test of equal variances was significant (P = 0,0000 and P = 0,0278 respectively). Therefore the Welch Test was used to determine the differences in the means (Table 14.12).

TABLE 14.12 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - DC

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20-39	0,0000	0,0929
40-69	0,0278	0,000 ***
70-79	-	-

*** P <0,0001

It was interesting to note that no statistical difference existed in the time scores obtained by the normal and the patient groups in the age range of twenty to thirty-nine years.

14.5.4. Analysis Of Tests Executed By Patients

An analysis of the scores of the patients showed the following:-

a) In the twenty to thirty-nine age group only one patient obtained a score below the fifteenth percentile and therefore demonstrated a problem with space perception as evaluated by this test. Three patients were discharged prior to being submitted to this test.

b) In the age group forty to fifty-nine years, five patients obtained scores below the fifteenth percentile, one was discharged prior to completing the test and four performed within normal limits.

c) In the age group sixty to sixty-nine years, four patients had defective scores i.e. below the fifteenth percentile, four performed within the normal limit and two were discharged prior to testing.

d) Of the patients available for testing in the seventy to seventy-nine year old group, two were untestable by means of this test although they had both participated in the two preceding construction tests. This was probably due to a combination of factors which included fatigue. Two patients were untestable on all tests in the battery, while another patient was untestable on all but the simplest matching tasks. One was discharged prior to attempting this test.

14.5.5. Problems That May Be Diagnosed From Performance On The Design Construction Test

a) Space Relationships. The mistakes made were similar to those described for the normal subjects, however, the errors were more obvious. Space relationship problems were as follows:-

- Pieces of the design had the incorrect orientation to one another other as seen in the construction made by patient twenty-one.

- Constructions were also very untidy with pieces separated or the overlaps were too deep - patient ten.

b) Position in Space. These problems were demonstrated through incorrect orientation of the design pieces, (patient twenty-one) and rotation of the entire design (patient two and patient twenty-one).

c) Apraxia. Apraxia was evident in the design constructed by patients seven and nine in that the pieces were re-ordered and built on top of the model.

d) Disregard for left space. This was seen in patient nine.

Although spatial problems and apraxia were clearly evident, performance on this test did not demonstrate problems other than those already discussed in the Block and Stick Construction tests. In addition problems were exacerbated by the difficulty experienced in handling the pieces which were made of thin cardboard covered with very thin plastic. The pieces were thus flexible and the small pieces were difficult for the hemiplegic patients to manoeuvre with one hand.

14.6. Item Analysis

14.6.1. Item Difficulty Index

The difficulty index was calculated for the normal group in the age groups between which no statistical differences were found in the scores obtained i.e. twenty to fifty-nine year olds and sixty to seventy-nine year olds (Table 14.13).

TABLE 14.13 DIFFICULTY INDEX - DC

ITEM	DIFFICULTY p 20-59 YRS	DIFFICULTY p 60-79 YRS
1	0,96	1,0
2	0,60	0,50
3	0,95	0,79
4	0,84	0,79
5	0,97	0,92
6	0,99	0,92

It appears from Table 14.13 that the items in order of difficulty were one, six, five, four, three and two.

The difficulty index for the responses obtained from the patient group was also calculated (Table 14.14)

TABLE 14.14 DIFFICULTY INDEX - PATIENT GROUP - DC

ITEM NO.	DIFFICULTY p 20-59 YRS	DIFFICULTY p 60-79 YRS
1	0,57 ***	0,50
2	0,21 ***	0,38 *
3	0,50 ***	0,25
4	0,57	0,50 **
5	0,57 ***	0,38 *
6	0,64	0,25

* $P<0,05$ ** $P<0,01$ *** $P<0,001$

The Pearson Chi-square test was used to calculate whether there were statistical differences between the normal and the patient group in the responses obtained on each item. This was done to help to determine the final test form (see Table 14.14).

It was interesting to note that the items differentiated between the normal and the patient performance in the younger age group better than they did in the older age group. This might have been due to the decreased ability in the normal older individual, as well as to the fact that there were no scores from the seventy to seventy-nine year old group.

All the items, except for item six, should be retained if the test were to become part of the battery, but re-ordering of the items as mentioned above should take place.

14.6.2. Item Discrimination

Two ways were used to determine the discriminatory level of each item.

14.6.2.1. Testing the Significance of the Difference between Two Proportions

The responses obtained by those normal individuals whose overall scores were in the upper 27% were compared with those whose overall scores were in the lower 27%. A cut point of one was used for the Fisher's Exact calculation in making this comparison (Table 14.15).

All items that showed a significant difference between the upper 27% and the lower 27% should be retained as they discriminated between those subjects who could do the test and those who could not. Thus, according to the results depicted in Table 14.15, all items should be retained although they might not have discriminated in the younger age group, they did so in the older age group.

TABLE 14.15 ITEM DISCRIMINATION - DC

ITEM	20-59 YRS	60-79 YRS
1	0,1069	0,0325 *
2	0,0000 ***	0,0007 ***
3	0.0207 *	0,0007 **
4	0,0002 **	0,0108 *
5	0,2317	0,0325 *
6	0,0013 **	0,0001 ***

* P<0,05 ** P<0,01 *** P<0,0001

14.6.2.2. Correlational Approach

This approach showed the relationship of the individual responses to the total test score (Table 14.16).

TABLE 14.16 BI-SERIAL CO-EFFICIENT OF CORRELATION BETWEEN ITEM AND TOTAL SCORE - DC

20 - 59 YRS				60-69 YRS		
Item	Upper 27%	Lower 27%	r Value	Upper 27%	Lower 27%	r Value
1	1,0	0,81	0,51 ***	1,0	0,56	0,72 ***
2	1,0	0,24	0,84 ***	0,80	0,0	0,85 ***
3	1,0	0,74	0,56 ***	0,80	0,0	0,85 ***
4	1,0	0,48	0,73 ***	1,0	0,44	0,80 ***
5	1,0	0,86	0,50 ***	1,0	0,56	0,70 ***
6	1,0	0,57	0,70 ***	1,0	0,11	0,89 ***

*** P<0,001

The correlational approach indicated that all items were discriminatory and should therefore be retained in the final test form.

14.7. TEST RELIABILITY

14.7.1. Split-Half Reliability

This determined the internal consistency of the test. The resulting coefficient of correlation for the age group twenty to fifty-nine was r=0,43 and for the age group sixty to seventy-nine, r=0,89. It seemed, therefore, that the reliability of the design construction test was lower when the young age group was considered (See Table 14.17).

14.7.2. Kuder-Richardson Formula 20

This method of calculating reliability coefficients determined the inter-item consistency of the test. The co-efficient of correlation for the age group twenty to fifty-nine years was r=0,79 and for age group seventy to seventy-nine years r=1,0 which was an exact correlation of each item with the total test score (Table 14.17).

TABLE 14.17 RELIABILITY CO-EFFICIENTS - DC

AGE	SPLIT-HALF	KUDER-RICHARDSON
20-59	0,43 ***	0,79 ***
60-79	0,89 ***	1,00 ***

*** P < 0,001

As can be seen from Table 14.17 all the reliability co-efficients were significant at the P<0,001 level of significance.

14.7.3. Inter-Rater Reliability

The Friedman two-way analysis of variance showed that there was no statistical difference between either the accuracy (P=0,6065) or time (P=0,6065) scores obtained by the three examiners on each of the six individuals tested. The test, therefore, showed inter-rater reliability.

14.8. TEST VALIDITY

To determine the validity of the Design Construction test as a measure of spatial relationships, two types of test validity were investigated.

14.8.1. Content Validity

Content validity was provided by the fifteen occupational therapists (population one) who rated the test on a 0-5 scale. Using this method the design construction test was shown to be a very much better than the average method of evaluating spatial relationships (Table 14.18) and was, therefore, considered to be a valid method of evaluating this perceptual skill.

TABLE 14.18 CONTENT VALIDITY - DC

PERCEPTUAL SKILL	MEAN RATING
Form as mass	1,4
Overall Contour	1,3
Sub-Units of Form	1,5
Perceptual Constancy	1,3
Colour	2,5
Directionality	1,9
Space Relations	3,7 ***
Position in Space	1,9
Figure-Ground	1,7
Praxis	Required

*** A very much better than average assessment.

It also appeared to be an average assessment of colour perception. This was due to the different colours used in the designs.

14.8.2. Construct Validity

14.8.2.1. Age Differentiation

Discussion already presented indicated that some *perceptual skills showed a progressive decrease with age.* The Design Construction test demonstrated a decrease in the ability to do the test from the age of approximately sixty years. This test could therefore be said to be a valid method of evaluating age related changes.

14.8.2.2. *Correlation with other tests in the Battery*

The highest correlations of the Design Construction test were with the two drawing tests and the Block Construction I test. The lowest correlations were with the Colour Matching test, the Arrows test and the Block Construction II test (Table 14.19). The reason for the low correlation with the Block Construction II test was not clear as both tests were constructed to measure the same skill.

A factor analysis was carried out to determine whether other tests contained factors related to the Design Construction test.

TABLE 14.19 CORRELATION WITH OTHER TESTS IN THE BATTERY - DC

	r VALUE		r VALUE
VMI	0,61	Position in Space	0,52
Part-Whole	0,56	Block Construction I	0,64
Perceptual Constancy	0,56	Block Construction II	0,39
Figure-Ground	0,59	Stick Construction I	0,47
Arrows	0,32	Stick Construction II	0,57
Dot-Line	0,63	Colour	0,23
Space Relations	0,56		

The tests that appeared to measure the same factor as the *Design Construction test were the Spatial Relations test, the* Arrows test, the Stick Construction II test, the Perceptual Constancy test, the Dot-Line test and the Figure-Ground test (Table 14.20).

TABLE 14.20 SORTED ROTATED FACTOR LOADINGS - DC

TEST	FACTOR II
Dot-Line Test	0,52
Figure-Ground	0,54
Space Relations	0,78
Arrows	0,75
Stick Construction II	0,60
Design Construction	0,57
Perceptual Constancy	0,52

The tests in Factor II, other than the Arrows test, the Perceptual Constancy test and the Figure-Ground test were constructed to measure spatial concepts. It seemed, however, that these tests also needed visual discriminative ability. The loading of the Design Construction test for Factor II might be due to the discriminative ability needed to *determine the overlapping of the shapes correctly.* The Design Construction test also loaded for Factor I (0,51) which was the space factor. The Design Construction test, *therefore, did nor clearly measure one perceptual construct.*

14.8.2.3. Results of Testing Patients

The Design Construction test was shown to be discriminative of perceptual problems in patients, a significant difference between the scores of the normal group and the patients being obtained (twenty to fifty-nine years $P<0,01$, sixty to seventy-nine years $P<0,01$).

14.9. CONCLUSIONS

In summary the following has been found on the Design Cr ... truction test:-

- Although easy to administer, the subjects found the pieces difficult to handle, a fact which could have led to some of the mistakes made in constructing the designs;
- The test was sensitive to changes in V-P-M function with advancing age and could differentiate between the problems of ageing versus the problems caused by C.V.A. There was a demonstrable difference between the performance of the normal group of individuals and the patients;
- The items differentiated between those subjects who could do the test and those who could not and the difficulty level of the items was variable;
- Reliability coefficients were at acceptable levels;
- The Design Construction test was according to the occupational therapists also a valid measure of the skill for which it was constructed. It was also a valid means of demonstrating perceptual dysfunction in C.V.A. patients;
- The test seemed to correlate with most other tests in the battery and had a higher loading for Factor II which appeared to be a visual discrimination factor, it however also loaded for Factor I which was the space factor, therefore its content has not been clearly demonstrated;

Consideration must be given to all the space tests and whether there is a need for the variety of construction tests compiled for the V-P-M battery. It appeared that the type of spatial problems that could be diagnosed through the Design Construction test could be just as easily diagnosed through the Stick and Block Construction tests. The ability to overlap figures did not feature as a major problem in testing both the normal and the patient group and did not necessitate the retention of this test in the battery.

Although the psychometric characteristics were fairly robust for this test, it might need to be withdrawn from the battery due to the type of pieces out of which the test is constructed and the resultant difficulty in reproducing the pieces in future versions of the test.

CHAPTER 15

THE COLOUR MATCHING TEST

15.1. INTRODUCTION

The perception of colour gives interest and vitality to the perceived world. Without perception of the myriad hues within the colour spectrum, the world would be dull and uninteresting. Apart from this, colour assists us to perceive depth as well as to distinguish an object from its background. It is important to determine a patient's ability to distinguish colour as defective colour perception will mitigate against the use of colour cues to help patients to perform when poor spatial function makes activities, such as dressing difficult.

A description of the Colour Matching test is to be found in Chapter Four Section 4.6.15. and Appendix 1 Section 2.11.

15.2. REVIEW OF SIMILAR TESTS

Two tests for colour blindness have been formulated (Hardy et al 1955 and Ishihara 1964). The stimulus materials for both these tests consisted of cards printed with different coloured dots forming recognisable figures against a background of contrasting coloured dots. Although these are tests for colour blindness they have been used in assessing brain damaged patients. De Renzi and Spinnler (1967) combined the Ishihara plates' with tests of pointing, naming, colour drawing and verbal memory for colour, in a battery. for evaluating impaired colour perception in brain impaired patients. The Ishihara plates' were felt to be unsuitable for use in the V-P-M Battery as they required visual closure skills and figure-ground perception as well as colour recognition skills.

Other tests of colour perception require the subject to arrange coloured paper chips (Farnsworth 1943) or coloured plastic chips (Whiting et al 1985) according to hue.

The Colour Matching sub-test of the RPAB (Whiting et al 1985) is suitable for use with C.V.A. patients, as it is the only test available that offers normative data. However, the test was published after this study was commenced. The Farnsworth test was unobtainable as it was published over fifty years ago.

Goldstein and Sheerer (1953) used skeins of wool to evaluate the ability to differentiate colour hue and brightness. However, the Goldstein and Sheerer test was primarily designed to measure abstract and concrete thought processes. The subject, when completing the test, is required to explain his reasons for matching certain colours. Although Kertesz (1979) used this test on his patients to evaluate colour agnosia, the test was thought to be unsuitable for the purposes required of this battery due to the type of skill that it was designed to measure.

Varney (Lezak 1983) developed a set of twenty-four line drawings of familiar objects each of which was accompanied by samples of four different colours. The subject is required to point to the colour appropriate to the object. This test was administered to one hundred normal subjects only four of whom failed to identify the colours of at least twenty objects correctly.

15.3. Reasons For Compiling A New Test

15.3.1. As tests such as the Ishihara Plates' and the Goldstein and Sheerer skeins of wool were found to be unsuitable for the V-P-M Battery, it was necessary to formulate a standardised test of matching colour hues.

15.3.2. The Colour Matching sub-test of the Rivermead Perceptual Assessment Battery (Whiting et al 1985), required the subject to manipulate plastic chips. This was found to be difficult for apraxic patients. The Colour Matching test of the V-P-M Battery described in this thesis does not require motor output apart from pointing. This action does not seem to be affected by apraxia and makes execution of the test easier.

15.4. RESULTS OF TESTING NORMAL INDIVIDUALS

15.4.1. Accuracy Scores

The means and the ''ndard deviations were calculated for the normal group as a whole (Table 15.1) and for each group separately (Table 15.2).

TABLE 15.1 MEAN AND SD OF ACCURACY SCORES - COL

n	MAX SCORE	MEAN	SD	RANGE
118	16	14,67	1,29	10-16

TABLE 15.2 MEAN AND SD OF ACCURACY SCORES FOR EACH AGE GROUP - COL

AGE	n	MEAN	SD	RANGE
20-29	20	14,85	1,14	13-16
30-39	21	14,57	1,72	11-16
40-49	19	15,21	0,92	13-16
50-59	19	14,95	1,13	12-16
60-69	19	14,0	1,53	10-16
70-79	20	14,45	0,83	13-16

The Pearson Chi-square test showed that there was no statistical difference in the performance of the groups when the percentage obtaining a score greater than the cut point of fourteen was compared to the percentage obtaining a score of less than fourteen (Table 15.3). Therefore the null hypothesis of no significant difference between the age groups was retained in this instance.

Percentiles were calculated for each age group within the defined parameters of age that were established on the results of other tests (Appendix 2) for the sake of uniformity.

TABLE 15.3 COMPARISON OF ACCURACY SCORES - COL

AGE	20-29	30-39	40-49	50-59	60-69	70-70	TOTAL
Less than 14	40	33,3	21,7	26,3	52,6	55,0	38,1
Greater than 14	60	66,7	78,9	73,7	47,4	45,0	61,9

Pearson Chi-square value = 0,1668 not significant

15.4.2. Time Scores

The means and the standard deviations were calculated for the time scores for the group as a whole (Table 15.4) and for each group separately (Table 15.5).

TABLE 15.4 MEAN AND SD OF TIME SCORES - COL

n	MEAN/SECS	SD	RANGE
118	46,41	23,63	12-152

TABLE 15.5 MEAN AND SD OF TIME SCORES FOR EACH AGE GROUP - COL

AGE	n	MEAN/SEC	SD	RANGE
20-29	20	31,5	12,84	17-65
30-39	21	33,95	13,82	17-72
40-49	19	45,16	19,70	18-97
50-59	19	48,16	14,96	27-86
60-69	19	55,58	32,79	27-152
70-79	20	65,20	24,68	34-134

Levine's test for equal variance established that there was a significant difference in the variances among the groups (P=0,0170) and the Welch Test established that there was a significant difference in the means obtained for the time scores for the different age groups (P<0,001).

The separate variance t was therefore used with the Bonferroni correction for multiple comparisons, to establish where the differences were to be found (Table 15.6).

TABLE 15.6 DIFFERENCES IN TIME SCORES - COL

AGE		SEPARATE VARIANCE t P VALUE
20-29	30-39	0,5593
20-29	40-49	0,0160
20-29	50-59	0,0007 *
20-29	60-69	0,0065
20-29	70-79	0,0000 ***
30-39	40-49	0,0474
30-39	50-59	0,0036
30-39	60-69	0,0135
30-39	70-79	0,0000 ***
40-49	50-59	0,6006
40-49	60-69	0,2445
40-49	70-79	0,0080
50-59	60-69	0,3779
50-59	70-79	0,0133
60-69	70-79	0,3098

* P<0,05 ** P<0,01 *** P<0,001

It was noteworthy that although a probably significant difference in the performance time existed between the twenty to twenty-nine and the fifty to fifty-nine year olds, no difference was demonstrated between the twenty to twenty-nine and sixty to sixty-nine olds. It was possible, therefore, to combine the time scores as per the already established groupings. A similar analysis was then carried out on the regrouped data to determine whether differences existed in the newly formed age groups (see Table 15.7).

TABLE 15.7 DIFFERENCES IN RE-GROUPED TIME SCORES - COL

AGE		SEPARATE VARIANCE t P VALUE
20-39	40-59	0,0002 **
20-39	60-69	0,0081 *
20-39	70-79	0,0000 ***
40-59	60-69	0,2780
40-59	70-79	0,0056 **
60-69	70-79	0,3098

* <P 0,05 ** P<0,01 *** P<0,001

The new age categories thus formed showed no difference in the length of time taken to complete the test between the forty to fifty-nine year olds and the sixty to sixty-nine year olds. There was also no difference found between the performance of the sixty to sixty-nine year olds and the seventy to seventy-nine year olds.

Time scores were converted to percentiles which can be found in Appendix 2.

15.4.3. Discussion Of The Results Of Testing Normal Individuals

Although the research into changes in perceptual function with age is fairly extensive, not much work had been done on the changes experienced in colour perception with advancing age. Botwinick (1981: 156) said that *"the elderly do not seem to have special problems seeing colours, although some colours, not all, are seen less clearly and vividly".* He said that blues and violets (shorter wave lengths) were seen less vividly, but that yellows and reds (longer wave lengths) were seen clearly. This was contrasted to the findings of Birren et al (1959) who found a startling decline after the age of thirty years in the ability to distinguish colours, with blue and green discrimination being most affected. Birren et al (1959) have also reported that with increasing age, colour had less influence on the perception of an object than did form.

The results obtained by the candidate for colour matching did not demonstrate any differences between the older and younger age groups in the ability to distinguish different colours. The mistakes that were made also did not involve blue/green confusion, but were related to the different hues of the same colour. This confirmed the results found by Whiting et al (1985) who also did not find a decrease in colour perception with advancing age.

15.5. Results Of Testing Patients

15.5.1. Accuracy Scores

The means and the standard deviations were calculated for all age groups combined (see Table 15.8) as no difference was found in the performance levels of the various ages in the normal group.

TABLE 15.8 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - COL

Age	n	MEAN	SD	RANGE
20-79	32	12,44	2,94	5 - 16

15.5.2. Time Scores

The means and the standard deviations were calculated for the age groups twenty to thirty-nine years, forty to sixty-nine years and seventy to seventy-nine years (see table 15.9) as these were the age groups within which no statistical difference occurred in the scores obtained by the normal group.

TABLE 15.9 MEAN AND SD FOR TIME SCORES FOR PATIENTS - COL

AGE	n	MEAN/SECS	SD	RANGE
20-39	7	97,14	102,17	28-277
40-69	20	85,90	46,90	22-199
70-79	5	187,20	116,98	78-379

15.5.3. Comparison Between Patients And Normal Subjects

15.5.3.1. Accuracy Scores

No statistical difference was found between the scores obtained by the patients and those obtained by the normal individuals, although several patients obtained scores below the fifteenth percentile.

15.5.3.2. Time Scores

The time scores obtained by the patients in the age categories twenty to thirty-nine, forty to sixty-nine and seventy to seventy-nine were compared with those scores obtained by the normal individuals in those age groups.

Levine's Test for equal variance was found to be significant in all age categories, however, the Welch Test showed that there was no significant difference between the scores obtained by the normal and the patient group in the age groups twenty to thirty-nine years and seventy to seventy-nine years, a significant difference was found only in the age group forty to sixty-nine years (Table 15.10).

TABLE 15.10 DIFFERENCES IN TIME SCORES OBTAINED BY NORMAL SUBJECTS AND PATIENTS - COL

AGE	LEVINE'S TEST P VALUE	WELCH TEST P VALUE
20 - 39	0,0000	0,1557
40 - 69	0,0030	0,0033 **
70 - 79	0,0000	0,1694

** P<0,01

15.5.4. Analysis Of Tests Executed By Patients

An analysis of the scores obtained by the patients showed that seventeen of the patients obtained scores that could be judged defective as they fell below the fifteenth percentile. Ten of these failures were due to left neglect, and seven were due to mis-judgement of the finer hues of certain colours. Only one patient demonstrated a true (colour agnosia) achromatopsia (see Chapter Eighteen). One patient was unable to complete this test along with all the other tests in the battery due to poor concentration and severe hemianopia.

15.5.5. Problems That May Be Diagnosed From Performance On The Colour Matching Test

a) Neglect of a Visual Field: As in all other tests described in the V-P-M battery, patients who had a severe left neglect of the visual field performed poorly on this test as none of the colours located on the left half of the multiple choice plate were correctly matched (see patient number twenty-eight). The Colour Matching test was not the only test in the battery that could be used for the diagnosis of neglect of a visual field.

b) Colour agnosia or achromatopsia is a defect in the perception of colours and is discussed in Chapter Eighteen. The Colour

Matching test of the V-P-M battery was a suitable test for the diagnosis of achromatospia particularly as language is not required for its execution. (The diagnosis of achromatopsia was made in the case of patient number seven).

15.6. ITEM ANALYSIS

As no significant differences occurred in the accuracy scores obtained by the various age groups, the scores for all the subjects in the normal group (i.e. group three) were combined for the completion of the item analysis, validity and reliability studies.

15.6.1. Item Difficulty Index

The difficulty index for the sixteen items in the Colour Matching test can be seen in Table 15.11.

TABLE 15.11 DIFFICULTY INDEX - COL

COLOUR	ITEM	DIFFICULTY p 20-79 YRS	COLOUR	ITEM	DIFFICULTY p 20-79YRS
Dark Olive Green	1	0,83	Beige	9	1,00
Pale Blue	2	0,98	Dark Green	10	0,74
Pink	3	1,00	Orange	11	1,00
Dark Red	4	0,92	Chocolate Brown	12	0,96
Blue	5	0,98	Apple Green	13	0,71
Burnt Orange	6	0,96	Blue	14	0,93
Dark Blue	7	0,98	Orange Red	15	0,94
Olive Green	8	0,86	Cinnamon Brown	16	0,92

The items in order of difficulty were: (three, nine, eleven), (two, seven), five, (twelve, six), fourteen, sixteen, four, eight, one, ten, thirteen.

Items <u>three, nine and eleven</u> do not have other similar colours of a slightly different hue in the test, they are therefore easy to distinguish. Items <u>two, seven and fourteen</u> are all blues and there is some difficulty in distinguishing them, particularly the colour falling in the middle of the range i.e. item fourteen. Item six (burnt orange) is often confused with the cinnamon brown colour of item sixteen. Item fifteen, the orange/red colour is often confused with item four which is the dark red colour. Items <u>one, eight, ten and thirteen</u> are all greens of different hues and these proved to be the most difficult to distinguish. The items containing the dark and light blue greens were transposed as was the dark green with the apple green.

The difficulty index for the responses obtained from the patient group was also calculated (Table 15.12).

The items in order of difficulty for the patient group were as follows, (three, nine, twelve), seven, fifteen, (one, two, eleven, five), six, thirteen, (four, eight), fourteen, sixteen, ten.

<u>TABLE 15.12 DIFFICULTY INDEX - PATIENT GROUP - COL</u>

ITEM NO.	DIFFICULTY p 20-79 YRS	ITEM NO.	DIFFICULTY p. 20-79 YRS
1	0,84	9	0,97
2	0,84 *	10	0,41 ***
3	0,97	11	0,84 ***
4	0,63 ***	12	0,97
5	0,84 **	13	0,81
6	0,81 *	15	0,56 ***
7	0,94	15	0,88
8	0,63	16	0,53 ***

* P<0,05 ** P<0,01 *** P<0,001

Although left neglect affected the performance of the patients, it did not appear to have unduly influenced the difficulty level. For example item thirteen located in the extreme left column, was not found to be as difficult for the patients as for the normal group.

Fischer's Exact test was used to calculate whether there were statistical differences between the normal and the patient group in the responses obtained for each item. Table 15.12 shows those items that differentiated between patient and normal performance i.e. sixteen, fourteen, eleven, ten, four, two, six and eight. It was interesting to note that they were not always the more difficult items and only one item was located in the extreme left column. The main problems experienced by the patients on this test were due to the difficulty in distinguishing the subtle difference in various hues and not left neglect.

15.6.2. Item Discrimination

This is the extent to which the item discriminates between those subjects who can do well on the test and those who can not. Two ways of determining this within the normal group were used.

15.6.2.1. Testing the Significance of the Differences Between Two Proportions

The responses of the top twenty seven percent of the test results were compared with those of the lowest twenty seven percent (Table 15.13).

On this basis, items, one, four, six, eight, ten, thirteen, fourteen, fifteen and sixteen should be retained. This would include hues of green, blue and red rather than the orange colour of different hues.

15.6.2.2. Correlational Approach

This approach showed the relationship of the individual responses to the total test score. For results see Table 15.14.

Using this approach it was apparent that only those colours that were different in hue should be retained in the test.

Thus items, four, five, six, eight, ten, thirteen, fourteen, fifteen and sixteen which tested hues of blue, red, orange and green should be retained in the final test form. It was interesting to note that other tests of colour perception (Farnsworth 1943, Goldstein and Sheerer, 1953 and Whiting et al 1985) test for the ability to discriminate between hues. Although it is not stated in the literature, this may be the main type of dysfunction occurring after brain damage (see achromatopsia).

TABLE 15.13 ITEM DISCRIMINATION - COL.

ITEM	20 - 79 YRS
1	0,0001 ***
2	0,1508
3	-
4	0,0199 **
5	0,1508
6	0,0389 *
7	0,3135
8	0,0012
9	-
10	0,0000 ***
11	-
12	0,0760
13	0,0000 ***
14	0,0199 **
15	0,0199 **
16	0,0101 **

$^*P < 0,05$ $^{**}P < 0,01$ $^{***}P < 0,001$

TABLE 15.14 BI-SERIAL CO-EFFICIENT OF CORRELATION BETWEEN ITEM AND TOTAL SCORE - COL

ITEM	UPPER 27%	LOWER 27%	DISCRIMINATION
1	1,00	0,63	0,70 ***
2	1,00	0,94	0,30 **
3	-	-	-
4	1,00	0,84	0,50 ***
5	1,00	0,94	0,30 **
6	1,00	0,88	0,50 ***
7	1,00	0,97	0,10
8	1,00	0,72	0,60 ***
9	-	-	-
10	1,00	0,53	0,72 ***
11	-	-	-
12	1,00	0,91	0,41 ***
13	1,00	0,34	0,82 ***
14	1,00	0,84	0,50 ***
15	1,00	0,84	0,50 ***
16	1,00	0,81	0,51 ***

** $P < 0,01$ *** $P < 0,001$

15.7. TEST RELIABILITY

15.7.1. Split-Half Reliability

The result of the coefficient of correlation between the two halves of the test was r=0,44. This low correlation may be due to the fact that responses to individual colours were scattered i.e. each colour was not grouped within the test, thus the Split-half reliability is probably an inappropriate calculation for this type of test.

15.7.2. Kuder-Richardson Formula 20

Using this formula, the resulting coefficient of internal consistency was r = 0,38.

Thus it can be seen that the Colour Matching test does not have high internal consistency.

15.7.3. Inter-Rater Reliability

There were no statistical differences in the scores obtained by the three examiners on the six individuals tested (see population five) for both accuracy (P = 0,8465) and time (P = 0,9592) therefore inter-rater reliability was proved.

15.8. TEST VALIDITY

15.8.1. Content Validity

The occupational therapists (i.e. population one) rated the test as a very much better than average measure of colour perception ($\bar{x}$=4,7). It would seem, therefore, that the Colour Matching test is what it is purported to be i.e. a test of colour discrimination. It is important to note that four therapists did not give a score of five out of five for this test as a test of colour matching ability.

15.8.2. Construct Validity

15.8.2.1. Age Differentiation

The results on this test confirmed the findings in the literature i.e. colour perception did not appear to change with age.

15.8.2.2. Correlation with Other Tests in the V-P-M Battery

As can be seen from Table 15.15 there was a very low correlation between the Colour Matching test and the other tests in the V-P-M battery. This was to be expected as the Colour Matching test was constructed to measure colour discrimination only. The highest correlation was with the Design Construction test, the only other test where colour is used.

TABLE 15.15 CORRELATION WITH OTHER TESTS IN THE BATTERY - COL

	r VALUE
Arrows	0,01
VMI	0,14
Block Construction I	0,13
Block Construction II	0,00
Dot-Line	0,28
Design Construction	0,38
Figure/Ground	0,03
Perceptual Constancy	0,21
Position in Space	0,21
Part Whole	0,16
Stick Construction I	0,21
Stick Construction II	0,21
Spatial Relations	0,15

A factor analysis was carried out on all the tests in the battery to determine which tests had factors that were related. The Colour Matching test was categorised as belonging to Factor III, the only other test with a rating above five in this factor was the Stick Construction I test (Table 15.16). It was unclear why this test should have a loading for Factor III which appeared to be a colour factor. The loading for the Design Construction test which made use of colour cues was surprisingly low for this factor.

TABLE 15.16 SORTED ROTATED FACTOR LOADINGS - COL

TEST	FACTOR III
Stick Construction II	0,46
Design Construction	0,40
Colour Matching	0,81
Stick Construction I	0,53
Block Construction I	0,42

15.8.2.3. Results of testing patients

The test showed no statistical difference between the patient group and the normal group. It was, however, interesting that clinically it was found to be useful in that a case of achromatopsia was demonstrated.

15.9. CONCLUSIONS

In summary:-

- The colour matching test was easy to administer.
- It, however, did not demonstrate age related changes, nor was there a significant difference between the performance of the normal group and the patients. However, the problems of a patient with achromatopsia were clearly demonstrated.
- The item difficulty p was high, however eight of the items demonstrated differences between normal and patient performance. The item discrimination studies showed that only the items that required the subject to differentiate between different hues of the same colour discriminated between those who could do the test and those who could not.
- Reliability studies, apart from the inter-rater reliability, were poor probably due to the nature of the test.

- The test was shown to be a valid measure of colour perception in patients, although it was not a valid measure of age related changes.

- It did not correlate well with other tests in the battery and it fell into Factor III which was labelled by the candidate a colour discrimination factor.

In spite of its low reliability the test should remain a part of the V-P-M battery but it should be reconstructed to consist only of the items which determine the ability to distinguish between hues.

CHAPTER 16

APRAXIA

16.1. INTRODUCTION

One of the aims of this study was to demonstrate the manner in which the V-P-M test battery could be used to diagnose certain visual-perceptual-motor syndromes. Apraxia is one of the syndromes which could result from a cerebro-vascular accident. The different types of apraxia are, therefore, described and the results of the tests on patients discussed in order to demonstrate the salient features of a constructional apraxia.

16.2. DEFINITION OF APRAXIA

Before apraxia is discussed one needs to have a workable definition of it. Apraxia, therefore, in the context of this thesis may be described as an impairment of the ability to carry out purposeful movement by an individual who has normal primary motor skills i.e. strength, reflexes and co-ordination and has no marked sensory or intellectual disturbances (Geschwind 1965, De Ajuriaguerra and Tissot 1969, Hecaen 1981, Miller 1986). The terms apraxia and dyspraxia are often used interchangeably. Apraxia however, refers to the complete absence of conscious purposeful movement, whereas dyspraxia refers to dysfunction. In general apraxia has been used to refer to adult malfunction whereas dyspraxia has been used to refer to a similar disorder in childhood. Although dyspraxia is probably a more suitable term, apraxia is used in this thesis due to its wide use in the literature.

The term apraxia was first used by Steinthal in 1871, however, it was Liepman who received the credit for describing apraxia as a clinical entity (Hecaen 1981). Much discussion about the nature of apraxia has taken place between then and the present. However, major contributions to the classification of the different types of apraxia were made by Pick in 1905 who distinguished between ideomotor and ideational apraxia, and Kleist, who in 1912 was the first researcher to refer to the phenomenon known as a constructional apraxia. He initially referred to it as optic apraxia but in work done together with Strauss in

1934 it was changed to the term used today (Hecaen 1981). More recently workers such as De Ajuriaguerra and Tissot (1969), Heilman (1979), Ayres (1985), Hecaen and Rondot (1985) and others added substantially to the understanding of apraxia.

Over the years many other types of apraxia were described, for example, Luria (1980) enumerated the varieties of frontal, kinaesthetic, dynamic and optic spatial apraxia. Hecaen (1981) described melokinetic, unilateral limb, buccofacial and dressing apraxia. On reviewing the literature it seemed that ideomotor, ideational and constructional apraxia incorporated all the symptoms that were compatible with the definition already given. In contrast the other types mentioned were distinguished by the difficulty of defining the limits between them and primary motor (melokinetic apraxia) or sensory (kinaesthetic apraxia) defects and by their sequential character or unilateral aspect (unilateral limb apraxia). Dressing apraxia has the same set of symptoms as visuo-constructional disorders resulting from right hemisphere lesions (Hecaen and Rondot 1985) and, therefore, did not seem to exist as a distinct entity. Somato dyspraxia described by Ayres (1985) and Fisher et al (1991) had a sensory base to it as the name implied and therefore did not fit the definition already given.

The basis of understanding apraxia is the understanding of the concept of praxis. Praxias are movement combinations which are acquired. These movement combinations, once learned, become automatic and the patterns laid down as motor engrams in the brain. A simple example of a praxis is learning to catch a ball, i.e. holding out one's arms in supination rather than reflexly holding them over the face for protection. The laying down and re-use of motor engrams is essential to the occupational therapy process in which both old and new skills are used.

A movement combination is, of necessity spatial in character, i.e. the limb moves through space. A movement sequence, therefore, has two components. Firstly, there is a decision making component or cognitive component, which is the part of the process which decides on the sequence of the action and, therefore, the muscles which need to be involved to execute the movement. The second component is a perceptual one, in which the visual and spatial information that is necessary for the cognitive process, is provided. A spatial component is, therefore, an essential basic part of any movement pattern. De Ajuriaguerra and Tissot (1969) have described the development of praxias or motor engrams based on the Piagetian concept of the development of a space structure and have linked types of apraxia to each of the developmental levels.

16.3. Review Of The Neurological Basis Of Apraxia

Before examining the different types of apraxia it may be helpful to review the *neurological basis of praxias* (Geshwind 1965, Heilman 1979, Faglioni and Basso 1985).

It appears that it is the left parietal lobe that stores motor engrams for the control of all purposeful skilled movement. When a movement is to be performed impulses are generated in the parietal lobe and travel to the pre-motor cortex on the left and thence to the motor cortex. Impulses generated in this way activate the right side of the body, therefore, for skilled movement patterns to take place with the left side of the body, impulses must cross from the left pre-motor cortex to the right pre-motor cortex via the corpus callosum.

Movements can be initiated in several ways, for example, by commands, i.e. through the medium of language, by imitation or they can be self initiated.

Should a command or instruction be given to perform a certain series of movements such as "catch the ball" the auditory stimuli are first interpreted in the secondary auditory cortex. This area is connected to the pre-motor cortex via the parietal lobe where the relevant motor patterns are generated according to the verbal instructions given. Because the language centre and the motor planning centre are both located on the left side of the brain, it is conceivable that both defects can be seen together. Many apraxic patients are also aphasic and in this instance the language disorder is often difficult to distinguish from the apraxic disorder. However, it seems unlikely that impairment in language plays a role in impairing the ability to imitate gestures (De Renzi et al 1980). It is the carrying out of movement to verbal command that is affected in the left brain damaged individual. Thus, when a movement is self generated, i.e. is done as a response to thinking, the connection between language and movement is vital as thinking is done in language.

Geschwind (1965) attributed the fact that axial movements were rarely disturbed in apraxia to the fact that these movements were controlled mainly through the extra-pyramidal system which has its origin over the entire cortical mantel. The pyramidal system largely originates in the motor cortex and damage to the left parietal region prevents verbal commands from reaching the pyramidal system thus affecting fine movements generated through this system. The extra-pyramidal system, originating from Wernicke's area is not affected as it still has access to verbal command.

Should a movement be initiated through imitation i.e. initiated through a visual stimulus, a different route is utilised. This time the stimulus is generated in the visual cortex from where it is relayed via the left parietal lobe to the pre-motor centre. Should the visual stimulus have spatial connotations, the whole process becomes more complicated as the stimulus must be interpreted on the right side of the brain before crossing to left pre-motor cortex. Thus, lesions of the right side of the brain may affect the spatial component of movement planning.

Lesions of the left parietal area also cause some disturbances in the spatial ordering of movement but because this is the non-dominant hemisphere for spatial function, these disturbances are not as marked as when the lesion is on the right side. Performance is qualitatively impaired depending upon the hemisphere involved.

Warrington (1969) compared the drawings of patients with left sided and right sided lesions. Those with left sided lesions tended to simplify drawings while those with right sided lesions had problems with the proportion, articulation and spatial relationships of the drawings. This was due to the type of function disturbed (see 16.3.3.). The tests of the V-P-M battery will make it possible to investigate the function of the two sides of the brain in more detail.

This is a simplistic view of the neuro-anatomical basis of apraxia, however, it serves to help in the understanding of the mechanisms behind the various types.

16.4. Types Of Apraxia

16.4.1. Ideational Apraxia

Ideational Apraxia was defined by Solet (1975) and Poeck (1982) as the inability to perform a sequence of movements using a real object. Patients with ideational apraxia are also unable to perform correct motor patterns in response to language i.e. in response to command (Heilman 1979). In addition, ideational apraxia has been described as the inability to put the elements of a sequence of movements together in the required order although the individual elements can be executed correctly (De Ajuriaguerra and Tissot 1969).

The most well known example of this was the patient who could light a match but could not take the next step of putting the match to the candle. Ideational apraxia is thus, a disturbance of the basic ability to manipulate objects in a required sequence. The problem was ideational or conceptual, the patient knew what the end product was but not how to get there (Ayres 1985). Individual components of activities were perfectly performed but their relationship to each other was not. A patient would, therefore, be able to pour milk into a cup for tea but would not be able to stop pouring in time to leave space for the tea or able to take the next step of pouring the tea.

According to De Ajuriaguerra and Tissot (1969), ideational apraxia was a defect in Piaget's first stage of developing a space structure i.e. in the manipulation of objects in space or the practical stage

De Ajuriaguerra and Tissot (1969) stated that the manifestation of ideational apraxia was always bilateral and was often seen in the context of severe aphasia.

Miller (1986) provided an excellent error analysis of the problems common to the execution of tasks carried out by patients with an ideational apraxia i.e.:

- Elements occur in the wrong order.

- Sections of the sequence are omitted.

- Two or more elements are blended together e.g. patient lifts sugar towards a cup while at the same time making a stirring motion. The action is incomplete.

- The sequence continues beyond what is necessary e.g. person undresses instead of just taking off his jersey.

- Objects are used inappropriately e.g. tooth brush used to brush hair.

- Movements are made in the wrong plane of direction.

- Perseveration of movements occurs.

- Many attempts at the movement are made.

There are many examples of the potential danger to the patient when sequences of actions are not performed in the correct order e.g. turning on a gas stove. Filling the kettle and neglecting to light the stove. Occupational therapists are in the ideal position to evaluate the patient for ideational apraxia as everyday tasks are the most obvious testing medium. The Block, Stick and Design Construction tests of the V-P-M battery will demonstrate ideational apraxia if the manner in which the task is completed is carefully observed.

16.4.2. Ideomotor Apraxia

Ideomotor apraxia was defined by Solet (1975) as the impairment of symbolic or expressive gestures and/or the utilisation of an object in its absence. In this type the ideation was intact but the planning or programming had broken down (Ayres 1985). This defect in the space structure centred of one's own body (the second stage of Piaget's development of a space structure) is demonstrated in symbolic gestures such as waving goodbye, making the sign of the cross, etc.

Ideomotor apraxia has also been described as the inability to carry out individual elements of a sequence of acts while being able to carry out the whole sequence correctly. The level of ideation necessary for the whole sequence was, therefore, preserved (Hecaen 1981). Hecaen (1981) suggested that there was a progressive difficulty for the person who had an ideomotor apraxia from expressive gestures with more or less symbolic value e.g. wave goodbye to that of making gestures which were meaningless such as positions of the hand and body. This latter type of apraxia is usually not seen in spontaneous activity, therefore, it is looked for in special tests of verbal command or imitation.

The performance of a patient with ideomotor apraxia typically improves when he uses an object, as the use of an object limits the number of potential movements that he can perform as well as providing additional sensory input.

According to Kertesz and Ferro (1984) ideomotor apraxia seemed to be the result of localised left hemisphere lesions although it has also been found to a lesser extent in right brain damaged patients (De Renzi, et al 1986).

It seemed that ideational and ideomotor apraxia were not just different types but represented different levels of processing. Thus, the difference might be one of complexity with ideational forming the basic process upon which ideomotor was built. (Hecaen and Rondot 1985, De Renzi 1985).

Roy (1978) also saw ideational and ideomotor apraxia as a common disorder resulting from a common origin "*...the only means of differentiating them is in terms of degree of impairment, ideational apraxia represents a more severe case of ideomotor apraxia in which there is impaired praxis to not only command, as in ideomotor apraxia, but also to imitation and ultimately to object use*" (Roy 1978:197). Ideomotor apraxia is impaired planning of movement in response to command whereas ideational apraxia is a more severe form with both impairment of movement to command as well as impairment in the use of objects.

In right handers, gesture learning is under the control of the left hemisphere. However, once an action has become automatic, the input from the left hemisphere is no longer necessary. This explains why everyday activities which have become fully automatic through practice such as buttoning a shirt, knocking at a door, scratching one's head can be performed correctly by ideational and ideomotor apraxics alike (Faglioni and Basso 1985).

Ayres (1985) suggested that there were three processes, the first was ideation or conceptualisation, the second was planning or choosing a strategy of action and the third was the motor implementation of the action. Ayres isolated these three abilities as separate observable behaviours. The candidate felt, however, that the third part or motor execution stage could not be observed as intact, as Ayres suggested, in the presence of defects within the other two stages. Thus defects within the motor execution stage are usually due to inco-ordination or paresis.

From the foregoing it was apparent that ideomotor apraxia could occur without ideational but that there was some doubt as to whether ideational apraxia could occur without the concomitant presence of ideomotor apraxia (Hecaen and Albert 1978).

Lehmkuhl et al (1983) and Miller (1986) enumerated the error patterns seen in patients with ideomotor apraxia as follows:

- The body part is used as the object e.g. tea is stirred with the finger.

- Incorrect space is left for the absent object when pretending to use an object.

- As in ideational apraxia, the plane of movement is altered. Limbs may also not be positioned correctly which may be confused with akinesthesia.

- Similar movement may take place somewhere else in the body during the action. This is known as general enhancement or augmentation of the movement. Sometimes these movements are the only response.

- Noises or words may accompany the movement.

- Repetition (or perseveration) of a previous movement.

- Manipulation of a body part by another part i.e. the body part is pushed or pulled into position e.g. when the patient is asked to put out his tongue he tries to pull it out with his fingers.

Ideamotor apraxia produces an awkwardness of movement although it does not prevent everyday activities from being carried out. This may be incorrectly attributed to having to use the non-preferred hand or to inco-ordination and ataxia. According to Miller (1986) an unrecognised ideamotor apraxia could lead to false diagnosis elsewhere e.g. poor responses on pointing tasks as in the Form Matching and Colour Matching tests of the V-P-M battery leading to the diagnosis of problems in the skills measured by these tests.

As both ideational and ideamotor apraxia were seen as predominantly a sequelae of left brain dysfunction, they have not been analysed in the patients' test results used for this thesis.

16.4.3. Constructional Apraxia

Benton and Fogel (1962), De Renzi (1982) referred to any type of performance in which parts are put together or articulated to form a single entity or object as constructional praxis. This implies an organising activity in which spatial relations among the component parts must be accurately perceived if these parts are to be synthesised into the desired unity.

Constructional apraxia was, therefore, defined as a disorder of formative activities in which the spatial part of the task is disturbed. Critchley (1953) said that patients with constructional apraxia would have some type of defective orientation in space, a defect which might not emerge until a motor task was attempted within a visual sphere. Poeck (1985) stated that constructional apraxia was basically a spatial and not a motor disturbance. He went on to say *"... it is the motor or efferent aspect of visual-spatial disorientation which ... might not be a unitary symptom but rather the result of visual, proprioceptive or vestibular dysfunction due to parietal lobe damage"* (Poeck 1985:99).

According to Critchley (1953), Mayer-Gross specified that constructional apraxia involved the activity space within the sphere of the hands and fingers and appeared to be a problem of changing the hand from being a component of personal space into a tool manipulated within the extra-personal space. This would conform with Schilder's statement that apraxia might be due to a fault in the transposition of the conception of space to manual activity without the conception of space being in itself disordered (Critchley 1953).

Benton and Fogel (1962) pointed out that the perceptual and conceptual demands differed between the execution of two and three dimensional structures. Benton (1967) suggested that there were two types of constructional activity. Firstly design copying which involved precise graphic movements and therefore a higher degree of sensory-motor precision than the second type which incorporated block building and object manipulation. Ayres (1985) has labelled these graphic apraxia and constructional apraxia respectively.

Benton and Fogel (1962) found that there was a statistical difference between patient performance on two dimensional (copying tasks) and three dimensional (constructional tasks), $P<0,05$. They also found that patients could perform poorly on one task and not on the other.

Constructional apraxia due to right hemisphere lesions seemed to follow mostly from spatial disorders while that due to left lesions seemed to be related to a programming disorder (Warrington et al 1966, Hecaen and Assal 1970) and could be part of an ideational apraxia. Thus, constructional problems seen in the right brain damaged patient were spatial in nature i.e. they showed loss of perspective, fragmentation, poor spatial relationships and lack of symmetry. Whereas the problems seen in the left brain damaged patient occurred as a result of the inability to plan the order of movements and showed simplification of construction and drawing. For example angles were expanded and there was poor use of structure (Warrington et al 1966, Warrington 1969, Arena and Gainotti 1978, Lezak 1983).

Critchley (1953) enumerated the types of difficulty that might be seen in drawing and construction tasks in apraxic patients as follows:-

- the copy may be smaller than the original;
- lines drawn by the patient may be tremulous or wavy;
- there may be errors of articulation;
- the patient superimposes his copy on the model or very close to the model;
- directional changes may be seen in the lines and there may be reversals;
- there may be perseveration.

Gheorghita (1980) classified the drawing errors made by patients with a constructional apraxia into 5 categories i.e.

- errors in line matching;
- errors in design structures;

- unilateral neglect of space;
- closing in affect;
- lack of perspective.

These were all included in Critchley's (1953) list of mistakes except for unilateral neglect of left space which often appeared as a separate entity to apraxia.

Miller (1986) listed the factors that distinguish right brain damaged from left brain damaged individuals (see table 16.1). These factors clearly illustrated the spatial influence on the problem in right hemisphere lesions versus the planning problem seen in left hemisphere lesions.

TABLE 16.1 FEATURES OF DRAWING APRAXIA IN LESIONS OF THE RIGHT AND LEFT HEMISPHERES

RIGHT HEMISPHERE	LEFT HEMISPHERE
Loss of perspective	Poverty of detail with copying simplified e.g. cube as a square
Displacement and distortion of parts	Acute angles become right angles
Loss of spatial relationships between parts	Copies reduced in size
Piecemeal approach	Tendency to close in towards model or even overlap it
Fragmentation of the design	Perseveration
Poor recognition of errors	Action is slow and hesitant with repeated false starts
Presence of unwanted details	Patient may give up easily
Rotation of figures	Order of completion of task is abnormal
Execution is fluent and unhesitant and lines are bold and well formed	

It was apparent that the disorganisation of several factors in combination, could produce constructional apraxia (Benson and Barton 1970) such as problems with the analytical sequential, programming processes typical of left hemisphere function and the visuo-spatial, synthetic, global processes typical of right hemisphere function (Hecaen and Assal 1970, Warrington 1966). Ayres (1985) concluded that constructional praxis, while dependant upon motor processes for execution, was less a motor activity than it was an ideation, conceptualisation, space perception and planning activity. She stated that ideation was common to all praxias and an essential component of all task and movement execution.

16.5. Tests Of Apraxia

16.5.1. Ideomotor And Ideational Apraxia

The Solet test of apraxia (Solet 1975) is a comprehensive test for ideomotor and ideational apraxia which was drawn up by an occupational therapist. It included demonstrating the use of real objects, the use of an object in its absence and symbolic movements. Scoring criteria were presented to help to determine dysfunction but no normative data were available.

Luria described ways of asking the patient to copy movements (Christensen 1975) which could be used to determine the presence of ideomotor apraxia. Some standardisation of Luria's test has taken place in the Luria Nebraska Test (Golden 1981). Copying movements are also used in the praxis test of the LOTCA (Itzkovich et al 1990). Lehmkuhl and Poek (1981) chose tests that make use of several activities with multiple objects such as making a sandwich, for evaluating ideational apraxia. Brown (1972) and Goodglass and Kaplan (1972) described batteries for the evaluation of ideomotor apraxia. Any of the above tests could be used along with the V-P-M battery if ideational or ideomotor apraxia is suspected.

16.5.2. Constructional Apraxia

For the testing of Constructional Apraxia there are a variety of copy forms and construction type tests available, such as Benton's Three Dimensional Constructional Praxis Test, (Benton and Fogel 1962 and Benton et al 1983), Rey's Complex Figure Test (Lezak 1983 and

1995), the 3D Copying, Cube Copying and R/L Copy shapes of the Rivermead Perceptual Assessment Battery (Whiting et al 1985). All have some normative data for ages sixteen through to sixty-nine years. However, none of these tests distinguish between these various age groups.

Lezak (1983), Luria (1980) and Critchley (1953) described free drawing tests such as draw-a-daisy, clock and sun. These are non standardised although Lezak provided a scoring scale for drawing a bicycle. Itzkovich et al (1990) provided scores for drawing a clock in the LOTCA.

Many authors used copying designs (Critchley 1953, Warrington et al 1966, Warrington 1969, Hecaen and Assal 1970, Arena and Gainotti 1978, Benton et al 1983) which have no norms and patients are, therefore, evaluated against a control group for the diagnosis of constructional apraxia. Others used reproducing models using sticks (Crithcley 1953, Hecaen and Assal 1970) or blocks (Critchley 1953, Benton and Fogel 1962, Warrington 1969, Benton et al 1983). Although control subjects have been used in some instances no normative data against which patient performance could be compared have been provided.

In evaluating constructional apraxia, care should be taken not to use tests that also evaluate intelligence such as Koh's block designs as this made isolation of apraxia difficult (Arena and Gainotti 1978) and has led Poeck (1982) to query the existence of constructional apraxia as an entity. The complex block building test in the RPAB (Whiting et al 1985) which was similar to Koh's blocks and to the block construction sub-test of the WAIS was therefore unsuitable for the evaluation of apraxia. It was also worth noting that the method of test administration appeared to be an important variable in test performance (Fall 19[illegible]).

Fraser and Turton (1986) describe a battery of tests, the Cambridge Apraxia Battery, which was developed at Addinbrooks Hospital. This battery is similar to that described by Jongbloed et al (1986). These tests were designed to evaluate somato-sensory function, e.g. double tactile stimuli. These tests can, therefore, only be used to screen for sensory disturbances prior to evaluating apraxia. The authors did not make the distinction between a sensory based apraxia and other apraxis.

16.5.3. The V-P-M Battery

Discussion already presented indicated that there was some certainty that the space structure was disturbed in a constructional apraxia. As the type of activities selected for therapy is dependent upon determining the presence or absence of spatial problems as an entity as distinct from those associated with a constructional apraxia, the V-P-M test battery was constructed to be able to make this distinction. The test battery, therefore, included tasks which evaluate visual-spatial ability through construction and drawing as well as through matching type activities which evaluated spatial function without motor output.

16.6. Determining Constructional Apraxia From The Results Of The V-P-M Battery

The tests in this study were not intended for the evaluation of ideomotor and ideational apraxia although careful observation of test taking behaviour could reveal the presence of these two types of apraxia. Ideomotor apraxia in the form of gestures could be evaluated in the tasks requiring the patient to point to the correct response

In addition only patients with right hemisphere damage were selected for testing for this thesis (see Research Design, Chapter Two) thus ideational and ideomotor apraxia which seemed to be the result of left hemisphere dysfunction (Hecaen and Rondot 1985 and Faglioni and Basso 1985) would not be seen in the patient group selected. Discussion of patient results will, therefore, focus on constructional apraxia only.

The scores on motor enhanced tests (i.e. VMI, Dot-Line, Block, Stick and Design Construction) and the motor reduced tests spatial tests(Position in

Space and Spatial Relations) were analysed for twenty-five patients who had appropriate test results available to make a diagnosis.

The following was found:-

a) Only one patient (patient sixteen) performed poorly on the drawing tests and the construction tests yet within normal limits on spatial tests not requiring motor planning. This patient therefore met Shilder's criterion that apraxia was due to a fault in the transposition of the concept of space to manual activity without the concept of space in itself being disturbed.

b) Seven patients (one, two, nine, ten, fifteen, twenty-one and twenty-seven) performed poorly on all of the above mentioned tests. Two of these patients were said to be apraxic as well as having problems in the conception of space (i.e. patients number nine and twenty-seven) as they revealed some of the signs pointed out by Ctritchley (1953) such as perseveration. Patients number six and twenty-three were thought to manifest a drawing apraxia rather than a generic constructional apraxia thus confirming the point made by Benton and Fogel (1962).

c) Patient number twelve performed poorly on the two motor reduced spatial tasks but within normal limits on most of the construction and drawing tests. This patient could, therefore, be said to have a problem of space visualisation which improved with object utilisation.

d) Patient number nineteen appeared to have no problem at all as all his test results fell within normal limits.

If spatial problems were closely associated with, and formed part of constructional problems, patients with a constructional apraxia should perform poorly on both constructional and motor reduced tests. Thus, constructional apraxia could only be diagnosed through an analysis of the type of behaviour manifested during test performance, it could not be diagnosed through test scores alone. Two of the patients who did poorly on both types of tests will therefore be discussed.

Patient number nine, a sixty-four year old woman, suffered a C.V.A. involving the right middle cerebral artery. She was seen three months later when she was referred to occupational therapy. She was found to have a severe left hemiplegia, being at Brunnstrom stage two of recovery in the upper limb and three in the lower limb (Brunnstrom 1970). She also had a severe left spatial neglect due to a visual field defect. She was totally dependent in all forms of self care and could not participate in any form of recreational activity. Testing was extremely difficult and she performed poorly in every test that she completed, falling between the 1st-5th percentile in both motor and motor reduced tests. She was found to have not only a severe spatial problem but the problem was compounded by a severe constructional apraxia. This patient illustrated very clearly some of the characteristics of a constructional apraxia as enumerated by Critchley.

The VMI test showed severe perseveration. She crowded her drawings into the corner and her copy of the designs was smaller than the original (see Chapter 6). She also showed a characteristic which the candidate has noted in apraxic patients i.e. she waved her arm over the original design in an attempt to plan the movement before actually executing the drawing.

Block construction showed very clearly the patient's hesitation and constant changes in the construction. The first attempt had elements that showed that she could perceive something of the spatial characteristics, however, the constant re-ordering and changes of mind in the execution of the tasks produced a construction that did not resemble the original design. All activities that the patient engaged in were complicated by the patient's left sided neglect, a symptom which Gheorghita (1980) named as belonging to the syndrome of a constructional apraxia.

This patient performed well on the Solet Test of apraxia which evaluated the presence of ideomotor and ideational apraxia, this was expected in view of the fact that she had a right sided lesion. Her inability to dress and participate in other forms of activities of daily living was, therefore, due to her severe spatial and apraxic problems.

Patient number eight, a forty-nine year old man, suffered a right hemisphere C.V.A. one month prior to evaluation. This patient also had a severe left hemiplegia being at the same stage of recovery as the previous patient. He had a mild hemianopia and left neglect of the visual field and was dependent in some forms of self care. He performed poorly on the drawing tests and on two of the constructional tasks as well as the motor free spatial tasks. However, the mistakes seen in the drawing tests were due to visual neglect of the left space and to subtle mistakes in the accuracy of his copying.

The type of mistakes made by patient number eight did not show the characteristics of apraxia other than the omission of parts due to his left hemianopia and some anomaly of size and articulation, which according to Patterson and Zangwill (1944), were technical errors seen in drawings which could indicate disordered spatial thought.

It appeared, therefore, that constructional apraxia could only be diagnosed in most instances by analysing the type of behaviour manifested during test performance. It was difficult to diagnose through test scores alone.

A review of the types of behaviour manifested during the apraxic patients execution of construction and drawing tasks also showed that three of the features listed by Miller (1986) as being characteristic of left hemisphere patients were seen in the behaviour of the right hemisphere patients tested for this study i.e. reduction in size of copies, closed in effect and perseveration. The manifestation of differences between hemispheres is far less clear than different texts would lead one to believe and more work is required in this area. The V-P-M test will enable this work to proceed.

In evaluating the importance of apraxia vis-à-vis visual spatial problems without apraxia, it appeared that the patient who had a constructional apraxia was very much more handicapped. This handicap extended into all spheres of life and the patient was unable to execute simple activities of daily living.

16.7. Conclusions

Although Shilder as quoted by Critchley (1953) stated that constructional apraxia could be manifested without a concomitant disturbance in space perception, this is a fairly rare occurrence. The analysis of the patient's results have confirmed Benton and Fogel's (1962) finding that the perception of spatial relations among component parts is necessary if these parts are to be synthesised into the desired unity.

The fact that there was a lack of clarity between the many types of apraxia discussed in the literature could be overcome through the use of the classification suggested by De Ajuriaguerra and Tissot (1969). This classification helped to put in perspective the different types of apraxia in relation to the role that sensory input might have. The suggested classification is as follows:-

- Sensory-Kinaesthetic - in this type the patient would know the goal to be achieved and would describe its unfolding, but, as the name implies, there would be a sensory-kinaesthetic dysfunction preventing him from achieving the goal.

- Somato-Spatial Apraxia - this would be characterised by a disorganisation of the body in space giving rise to a disturbance of an action in space. This category would incorporate the previously discussed types of ideational and constructional apraxia and would present, according to the hemisphere disturbed, with different spatial problems.

- Symbolic Apraxia - this would be characterised by a disturbance of symbolic activity i.e. the previously named ideomotor apraxia.

The V-P-M battery seemed to be a useful tool in the diagnosis of a constructional apraxia and can be used to determine apraxia of either a graphic or constructional nature. In terms of De Ajuriaguerra and Tissot's classification it could be used to determine a somato- spatial apraxia.

Aspects of this Chapter were published in the British Journal of Occupational Therapy 1987, 50:7, 222-226 (see Appendix 5).