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	0,3266

*** 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	0,0570	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 Spacc 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	0,98	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	0,30	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 2 [illegible] YRS	ITEM	DIFFICULTY p 20-69 YRS
1	[illegible] 2	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	0,60 ***
16	0,88 **	36	1,00
17	0,84 **	37	0,96
18	0,96	38	0,92 *
19	0,92 *	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 sampling. The resultant co-efficient of correlation for the Space Relations 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 spatial relations test :

- 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 all 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 position in space test:-

- 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 ch. ' 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.

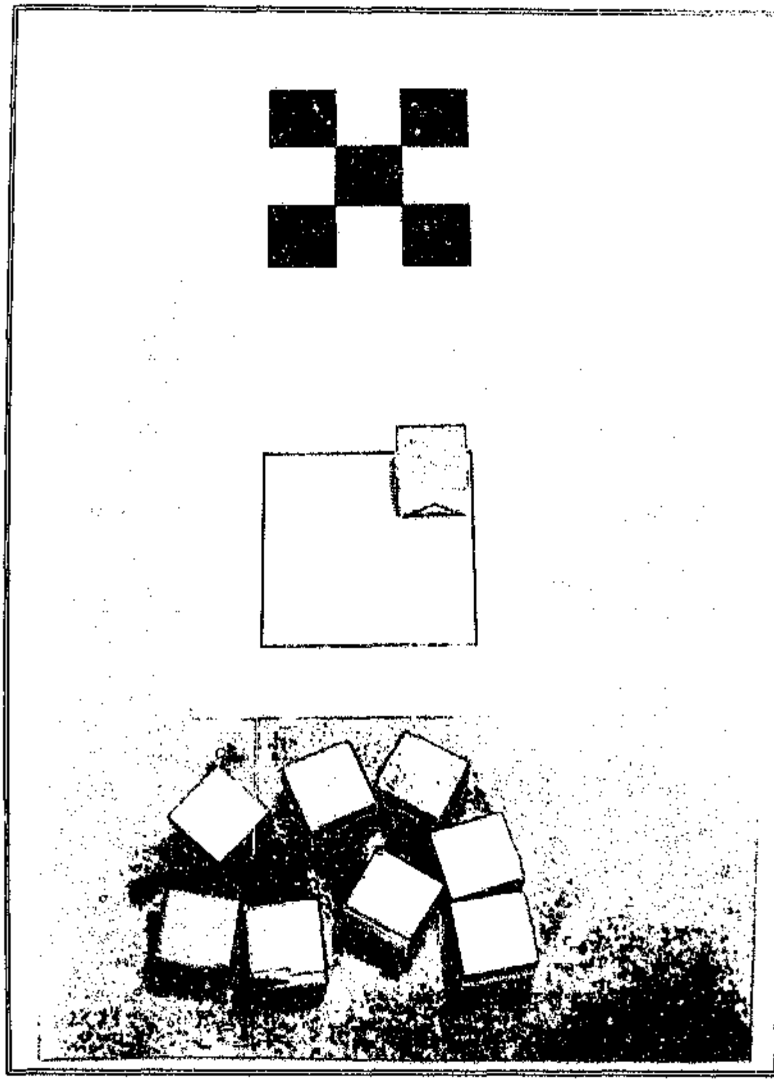

FIGURE 12.1 CUBE COPYING TEST - RIVERMEAD PERCEPTUAL ASSESSMENT BATTERY

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

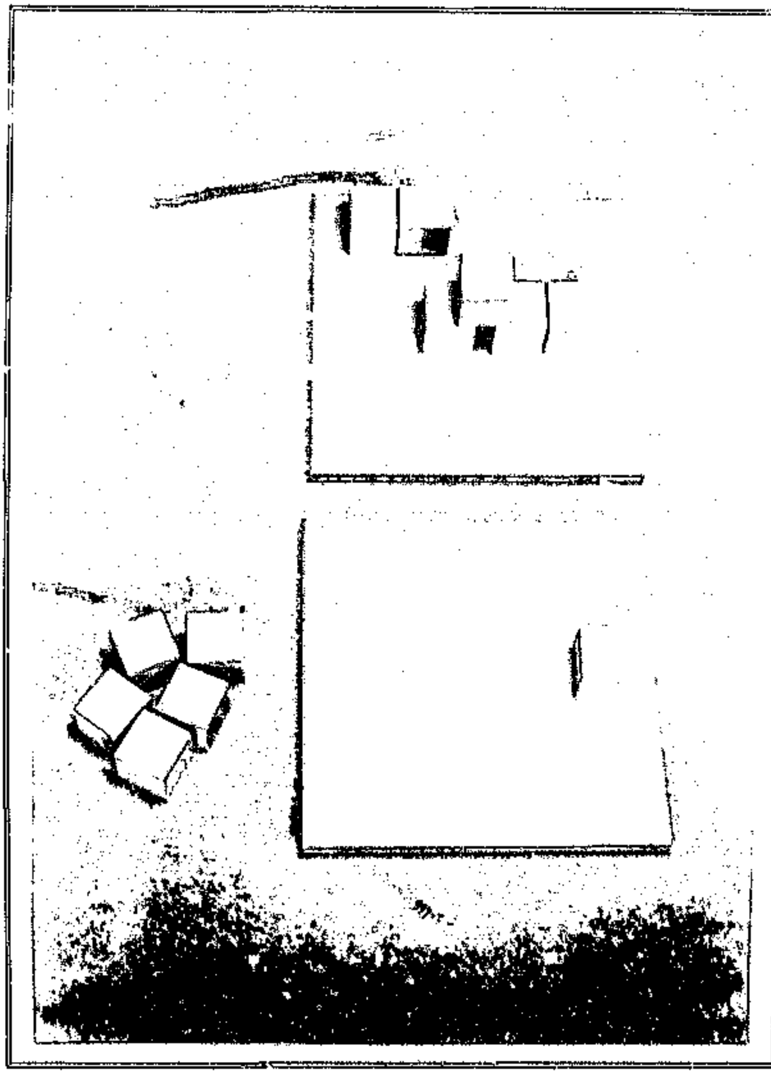

FIGURE 12.2 TRIAL ITEM - RIVERMEAD PERCEPTUAL BATTERY - THREE DIMENSIONAL BLOCK CONSTRUCTION

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 m~del 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	1,88	6-12
40-49	19	9,84	2,12	6-12
50-59	19	9,90	2,28	4-12
60-69	19	9,88	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 Score^ - 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
116	83,06	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	41-146
40-49	19	74,74	22,17	45-136
50-59	19	76,53	16,95	52-110
60-69	19	95,74	34,68	52-189
70-79	20	123,85	41,65	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 existe... ...ong some of the regrouped ages, percentiles were calculated ... t ie 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		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 reflected 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. Accurac' 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	64-276
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.

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 constructic . f 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 r..structuring of a design which does not result in a correct construction in the end (patient nine) (planning problem).

g) Poor and incorrect spa . .cement 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 to meet the requirements of construct validity. Mistakes made by 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 blocks, perseveration, the building of her construction on top of the model and difficulties in the correct spatial placement of blocks (Fig 12.9).

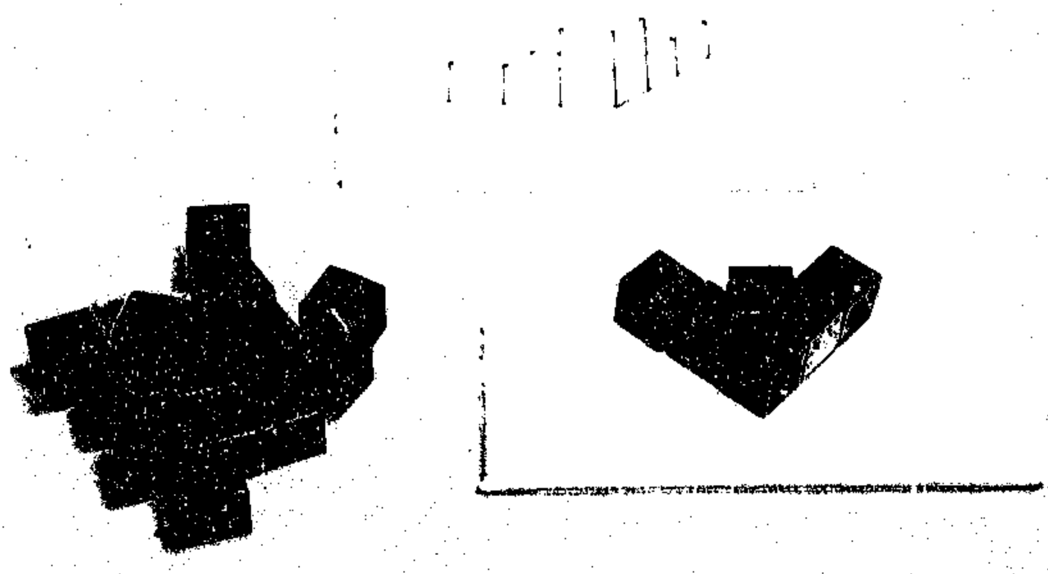

FIGURE 12.3 CONSTRUCTION ROTATED

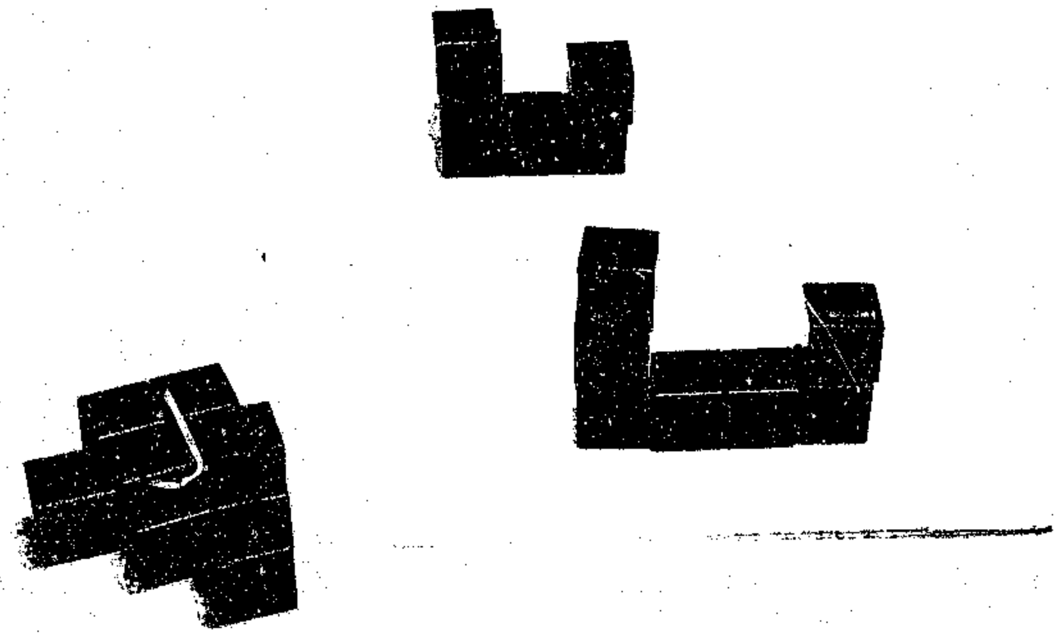

FIGURE 12.4 POOR SPATIAL PLACEMENT

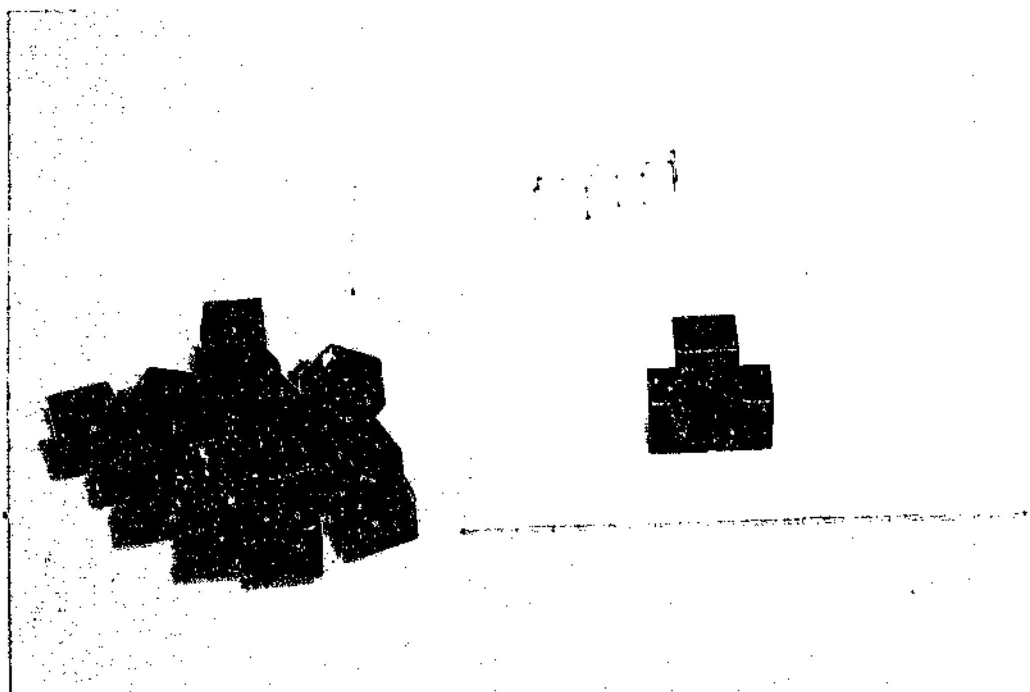

FIGURE 12.5 INCORRECT SPATIAL PLACEMENT

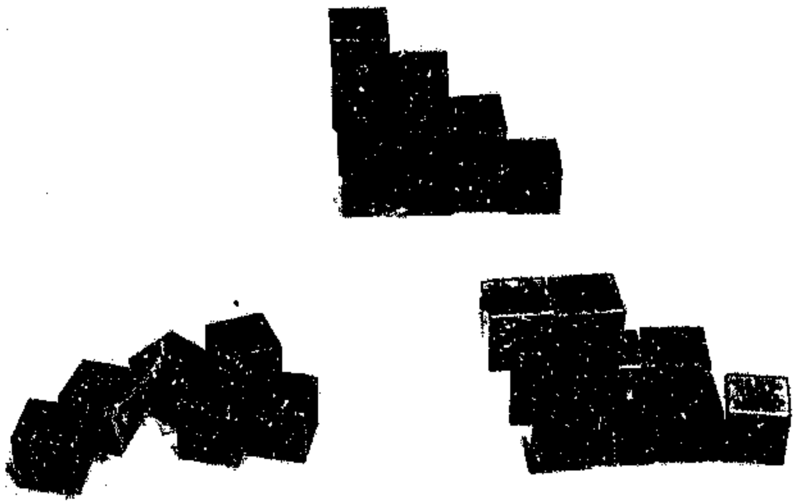

FIGURE 12.6 INCORRECT SPATIAL PLACEMENT

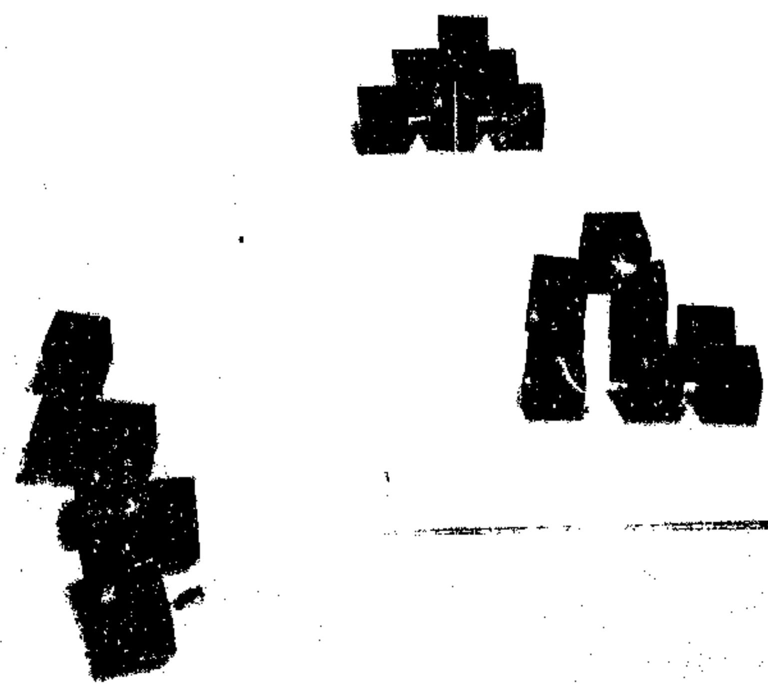

FIGURE 12.7 PROBLEMS OF SPACE RELATIONS - GRASPING SPACE

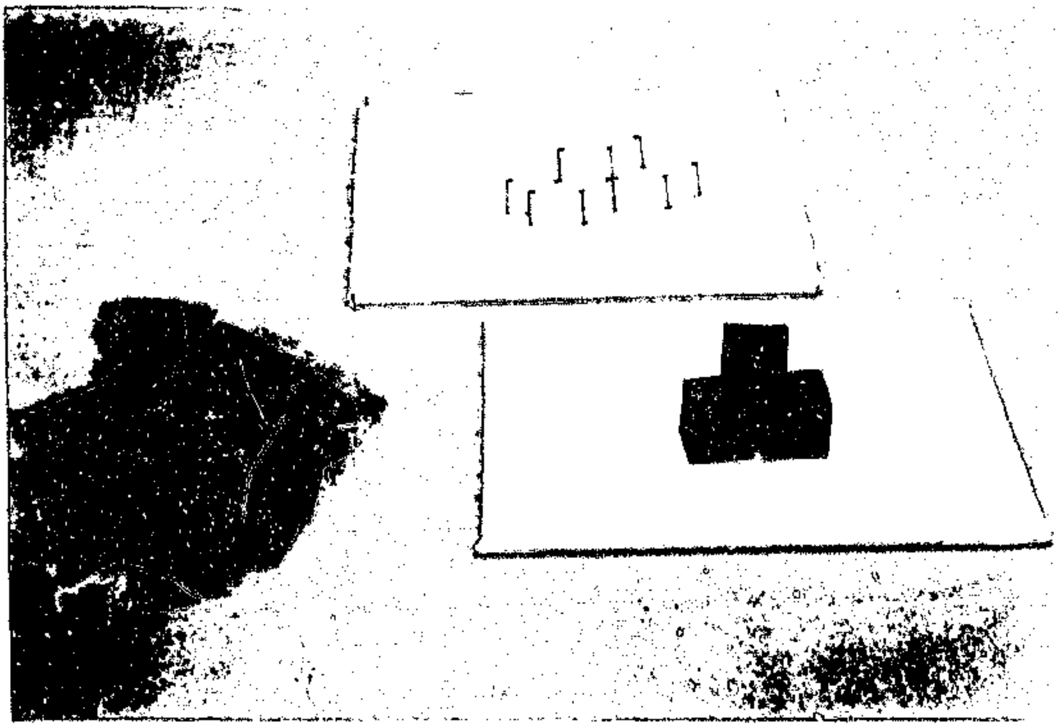

FIGURE 12.8 PROBLEMS OF POSITION IN SPACE - ACTION SPACE

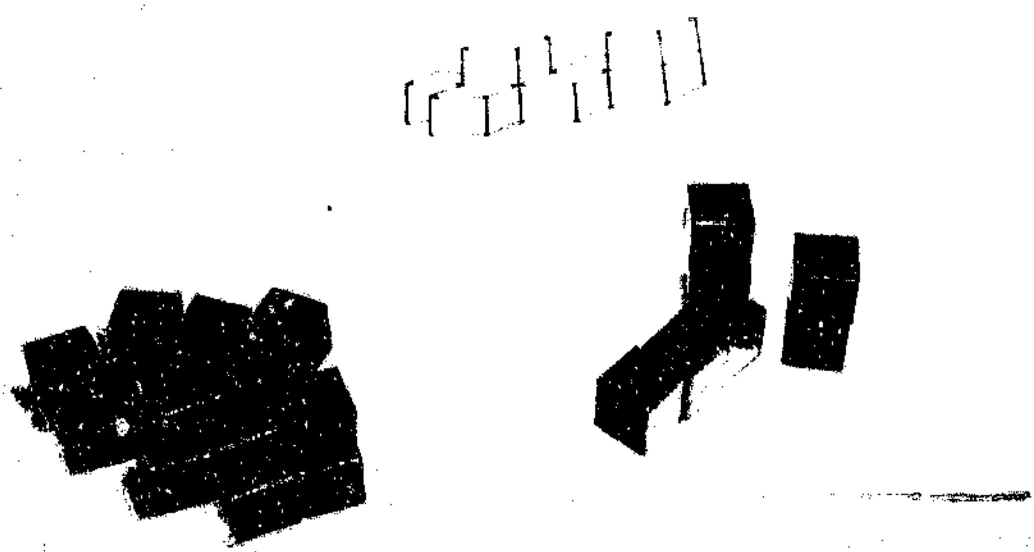

FIGURE 12.9 PROBLEMS OF POSITION IN SPACE - ACTION SPACE

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	0,64

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	0,66 ***
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	3,68 **	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

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   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

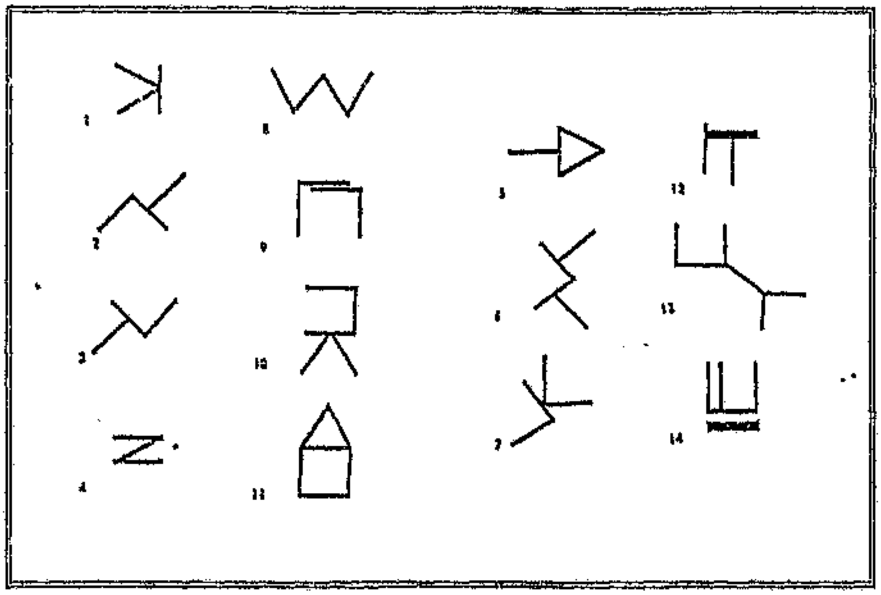

FIGURE 13.1 THE GOODGLASS AND KAPLAN STICK CONSTRUCTION TEST

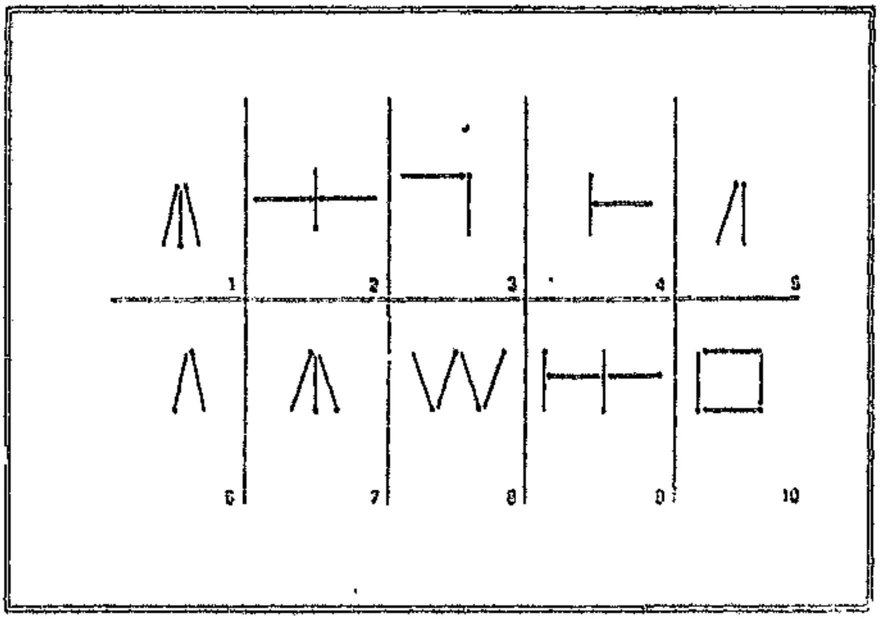

FIGURE 13.2 THE BUTTERS AND BARTON STICK CONSTRUCTION TEST

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	0,61	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	1,36	7-12
50-59	19	10,26	1,45	8-12
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 a·· groups therefore the null hypothesis which stated that decreased ..ility 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 = 0,4053

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 [illegible] SCORES FOR EACH AGE GROUP - SCI

AGES	n	MEAN/SECS	SD	RANGE
20-29	20	65,2	21,09	38-114
30-39	21	64,43	21,24	37-131
40-49	19	64,90	15,60	19-89
50-59	19	68,58	17,[illegible]	39-100
60-69	19	84,21	25[illegible]	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	0,0095
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	0,2868
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 Ex mination (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).

<u>**TABLE 13.15 MEAN AND SD FOR ACCURACY SCORES FOR PATIENTS - SCI**</u>

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	36,68	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.

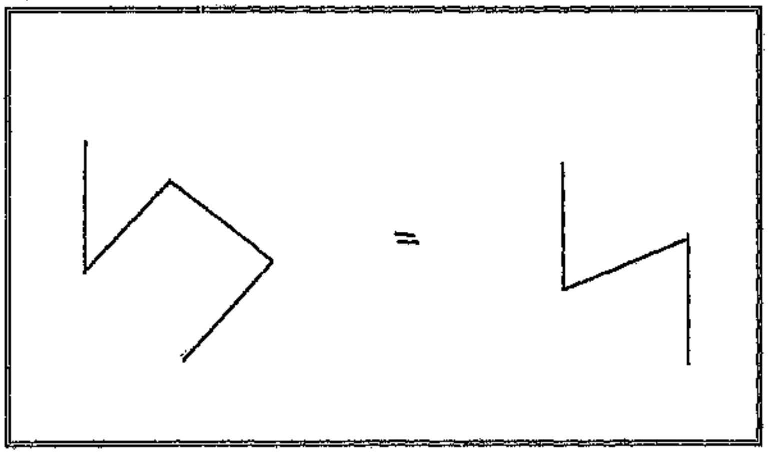

FIGURE 13.3 LEAVING OUT PARTS OF DESIGN

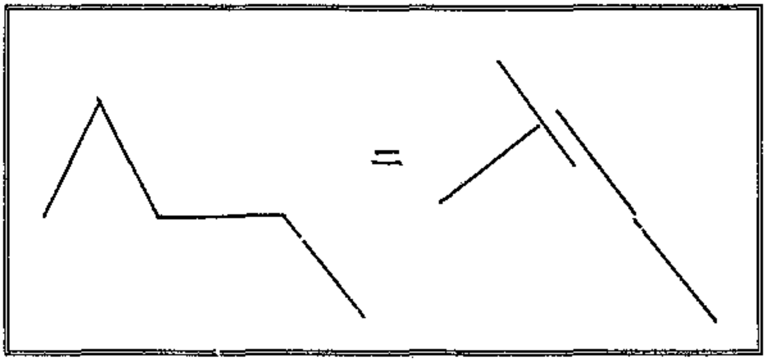

FIGURE 13.4 SPATIAL MISPLACEMENT

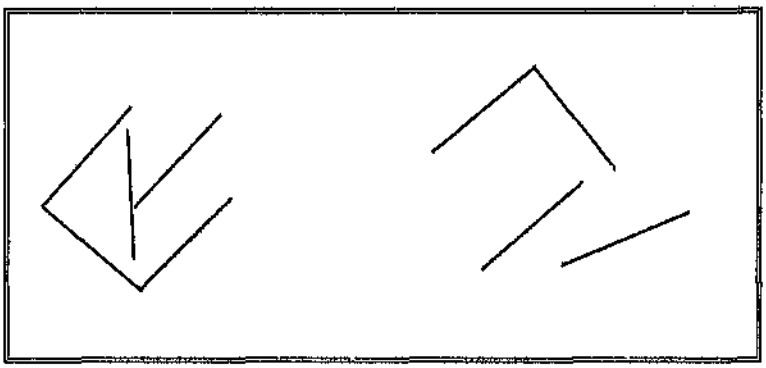

FIGURE 13.5 SPATIAL MISPLACEMENT

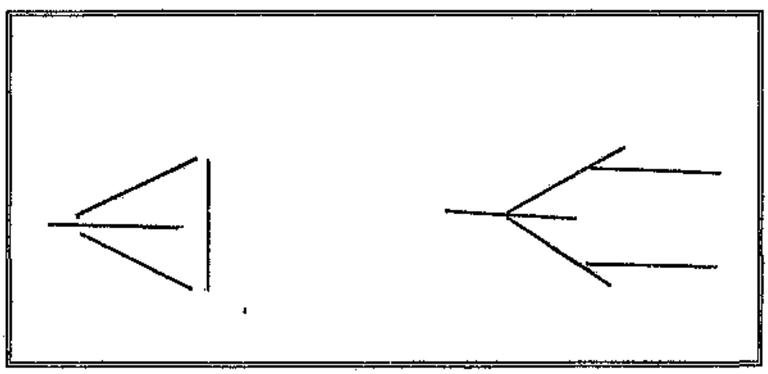

FIGURE 13.6 SPATIAL MISPLACEMENT

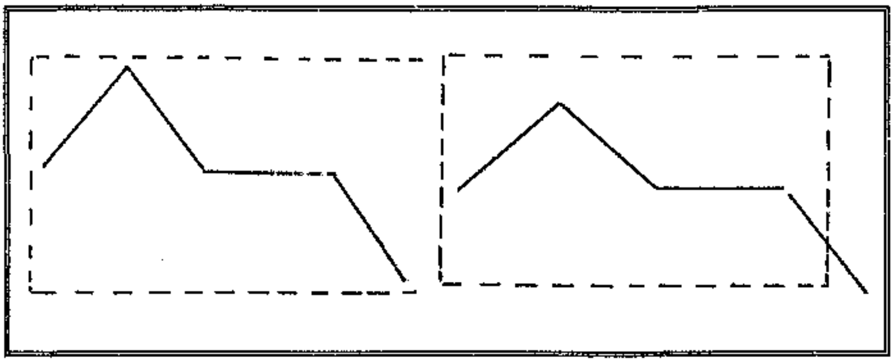

FIGURE 13.7 INABILITY TO USE ALLOCATED SPACE FOR CONSTRUCTION

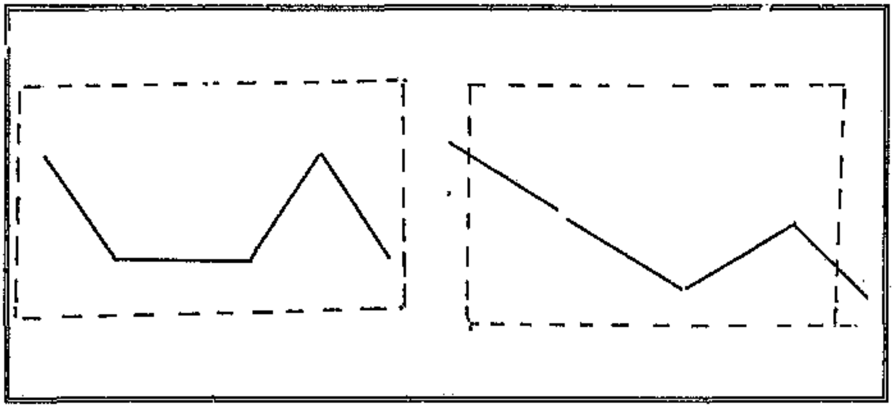

FIGURE 13.8 SEPARATION OF STICKS

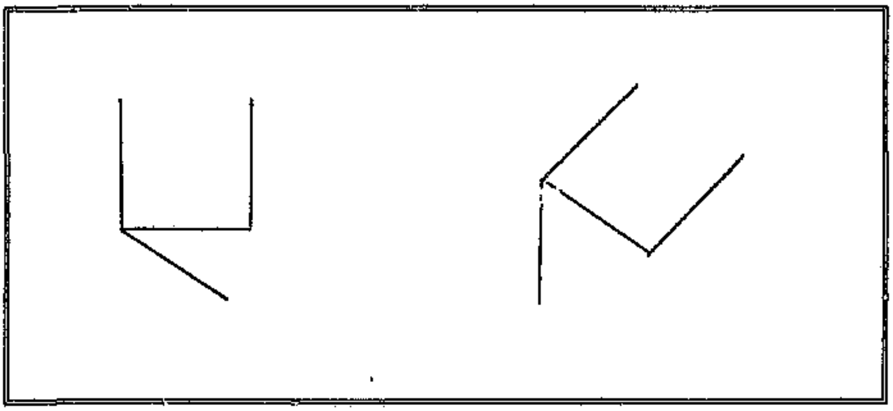

FIGURE 13.9 ROTATION OF DESIGN

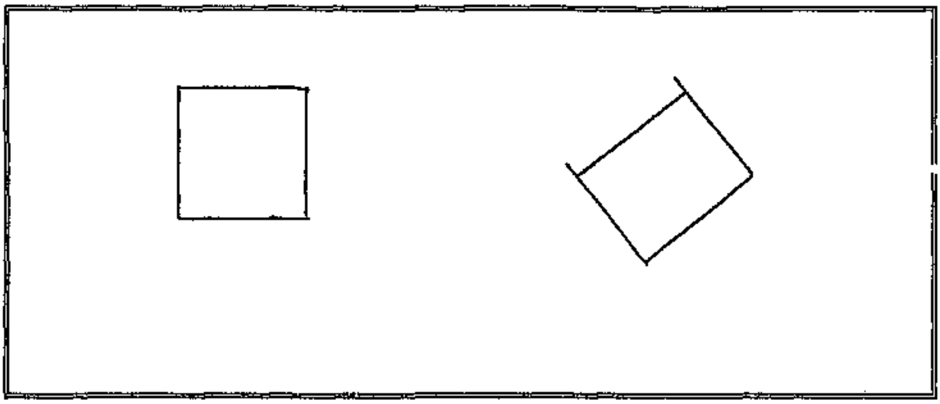

FIGURE 13.10 ROTATION OF DESIGN

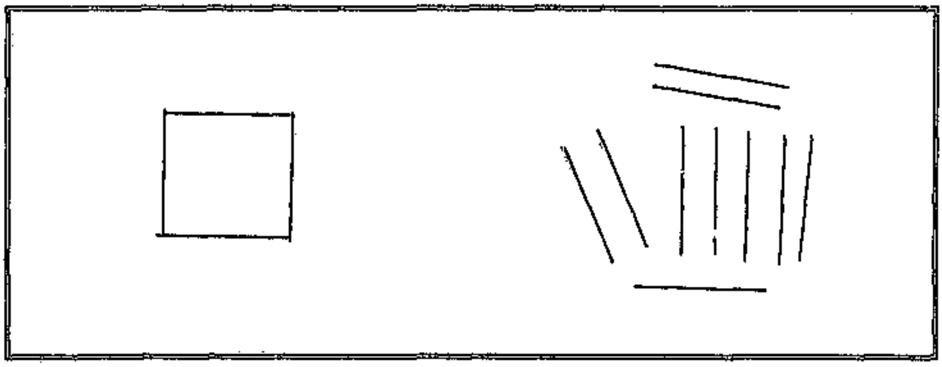

FIGURE 13.11 PERSEVERATION

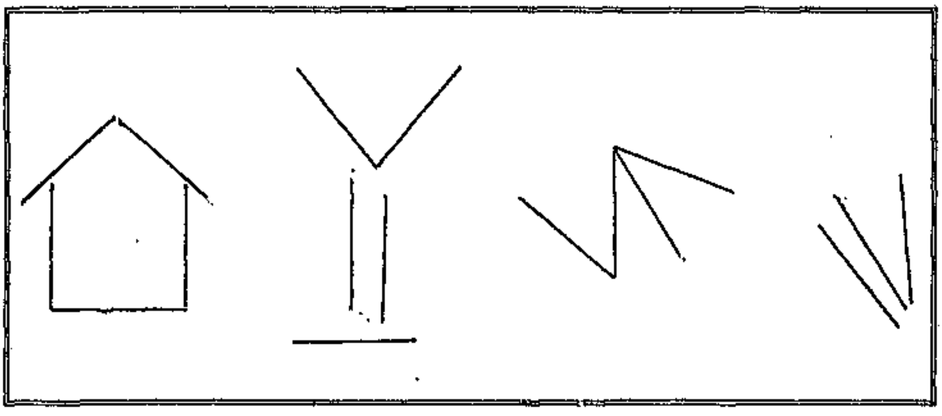

FIGURE 13.12 CONSTANTLY CHANGING DESIGN

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 intende ! 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 those 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 = 12,66 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	34,89	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[illegible]	0,0000 ***
40-59	[illegible]	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. T'

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	10-6296
40-59	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

14.6.3.2. 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	0,98	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 ***	0,60
2	0,21 ***	0,38 *
3	0,50 ***	0,25
4	0,67	0,60 **
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 [illegible] 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	0,0326 *
2	0,0000 ***	0,0007 ***
3	0,0207 *	0,0007 **
4	0,0[illegible] **	0,0108 *
5	0,2317	0,0326 *
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	0,65	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	0,63
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	0,67
Dot-Line	0,63	Colour	0,23
Space Relations	0,66		

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	0,76
Stick Construction II	0,60
Design Construction	0,57
Perceptual Constancy	0,52

The tests in Facto. ier 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 (0,61) 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 Construction 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 standard 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	12-15
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	31,6	12,84	17-65
30-39	21	33,95	13,82	17-72
40-49	19	45,16	19,70	13-97
50-59	19	48,16	14,66	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	0,6593
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	30-69	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	86,90	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 thos 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 tha. . e 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 ^f 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 - 78 YRS
1	0.0001 ***
2	0.1508
3	-
4	0.0199 **
5	0.1509
6	0.0309 *
7	0.3135
8	0.0012
9	-
10	0.0000 ***
11	-
12	0.0760
13	0.0000 ***
14	0.0109 **
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	0,84	0.30 **
6	1.00	0,68	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	0,61 ***

** 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 Rotations	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	0,40
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 ce in visual-perceptual-motor syndromes. Apraxia is one of the syndromes which could result from a cerebro-vascular accident. The differe · · nes of apraxia are, therefore, described and the results of the tests u· 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 purpose vement by an individual who has normal primary motor skills i.e. strength s 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 anipulate 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.