



# Kent Academic Repository

Guest, Richard M. (1999) *The diagnosis of visuo-spatial neglect through the computer-based analysis of hand-executed drawing tasks*. Doctor of Philosophy (PhD) thesis, University of Kent.

## Downloaded from

<https://kar.kent.ac.uk/94386/> The University of Kent's Academic Repository KAR

## The version of record is available from

<https://doi.org/10.22024/UniKent/01.02.94386>

## This document version

UNSPECIFIED

## DOI for this version

## Licence for this version

CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

## Additional information

This thesis has been digitised by EThOS, the British Library digitisation service, for purposes of preservation and dissemination. It was uploaded to KAR on 25 April 2022 in order to hold its content and record within University of Kent systems. It is available Open Access using a Creative Commons Attribution, Non-commercial, No Derivatives (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) licence so that the thesis and its author, can benefit from opportunities for increased readership and citation. This was done in line with University of Kent policies (<https://www.kent.ac.uk/is/strategy/docs/Kent%20Open%20Access%20policy.pdf>). If you ...

## Versions of research works

### Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

### Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal**, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

### Enquiries

If you have questions about this document contact [ResearchSupport@kent.ac.uk](mailto:ResearchSupport@kent.ac.uk). Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

The Diagnosis of Visuo-Spatial Neglect  
Through the Computer-Based Analysis of  
Hand-Executed Drawing Tasks

Richard M. Guest

September 1999

A Thesis submitted to the  
University of Kent at Canterbury  
in the subject of Electronic Engineering  
for the degree of Doctor of Philosophy

## Abstract

This thesis documents the investigation of a technique for the computer-based assessment of visuo-spatial neglect for use within a population of stroke patients. Analysing the hand-drawn responses from a battery of neuropsychological tasks, a series of automated feature extraction routines have been implemented to accurately and consistently assess performance in a novel way, leading to a diagnostic indication of neglect severity.

An investigation into the reliability of existing neglect assessment methods highlights the ambiguity in interpretation of marking criteria and the inaccuracy introduced due to human error in score calculation. The implemented feature extraction routines overcome these problems by algorithmically applying identical criteria to all test responses.

The results of a clinically-based trial using the developed system show that significant performance differences can be identified both using conventional *static* features (the outcome of the test response) and novel *dynamic* time-based constructional features which establish previously unmeasured performance characteristics of neglect-based response while increasing the sensitivity of the detection of neglect. The correlation between the static features and existing assessments of neglect verify the ability of the computer-based battery to detect neglect.

A feasibility study into the automated classification of feature measurements indicates the sensitivity of the individual tasks to detect neglect performance and shows that it is possible to classify responses by the analysis of the principal features extracted from test responses.

## Acknowledgements

In reflecting on the work documented in this thesis, I would particularly like to thank my supervisor, Mike Fairhurst, for his continual and untiring support and guidance throughout the research. Your advice and commitment has been invaluable.

The multidisciplinary nature of this research has required expert advice and interpretation from a wide variety of sources. The guidance and assistance that I have received from Jonathan Potter, Nick Donnelly, Tony Deighton, Mahool Patel, Simon Hanaford and Gill Gower over the course of the research has been greatly appreciated. I would also like to thank the staff (especially the Occupational Therapy Departments) and patients at Nunnery Fields Hospital, Canterbury, Queen Victoria Memorial Hospital, Herne Bay and Queen Elizabeth the Queen Mother Hospital, Margate.

I would like to thank the members of the Computer Vision and Image Processing Group at UKC for their ideas, support and encouragement, particularly Chris Allgrove, Will Cobbah and Fuad Rahman.

Finally, I would like to thank Tanya who has listened to me talk about nothing other than this thesis for the past five months. This must have been very dull.

Richard M. Guest.

Canterbury, September 1999.

# Table of Contents

<b>Abstract.....</b>	<b>i</b>
<b>Acknowledgements.....</b>	<b>ii</b>
<b>List of Figures.....</b>	<b>ix</b>
<b>List of Tables.....</b>	<b>xii</b>
<b>Chapter 1 - Introduction and Background.....</b>	<b>1</b>
1.1 General Introduction.....	1
1.2 Visuo-Spatial Neglect and Rehabilitation.....	2
1.3 Computer Based Testing and Evaluation.....	4
1.4 Thesis Structure.....	6
<b>Chapter 2 - Computer Based Diagnosis of Visuo-Spatial Neglect</b>	
<b>- A Review.....</b>	<b>8</b>
2.1 Introduction.....	8
2.2 Stroke and Visuo-Spatial Neglect.....	9
2.2.1 Effects on Everyday Activities.....	10
2.2.2 Rehabilitation of Neglect.....	11
2.3 Testing for Neglect.....	12
2.3.1 Copying and Drawing Tasks.....	13
2.3.2 Cancellation and Visual Search Tasks.....	17
2.3.3 Line Bisection Tasks.....	21
2.3.4 Test Sensitivity and Neglect Test Batteries.....	24
2.4 Computer Based Neuropsychological Testing.....	25
2.4.1 Static Features.....	28
2.4.2 Dynamic Features.....	28
2.5 Pattern Recognition Techniques.....	30
2.5.1 Principal Component Analysis.....	31
2.5.2 Bayesian Statistical Classification.....	32
2.5.3 Cluster Analysis.....	33
2.5.4 Kohonen Self-Organising Map.....	35
2.5.5 Adaptive Resonance Theory.....	36

2.6	Design Objectives.....	38
2.7	Summary.....	39
<b>Chapter 3 - Reliability of Behavioural Inattention Test.....</b>		<b>41</b>
3.1	Introduction.....	41
3.2	The Behavioural Inattention Test.....	41
3.2.1	Conventional Subtasks of BIT.....	42
3.2.1.1	Line Cancellation.....	43
3.2.1.2	Letter Cancellation.....	43
3.2.1.3	Star Cancellation.....	43
3.2.1.4	Figure and Shape Copying.....	45
3.2.1.5	Line Bisection.....	46
3.2.1.6	Representational Drawing.....	46
3.2.2	Battery Score Interpretation.....	47
3.3	Interrater Methodology.....	49
3.3.1	Assessment of Result.....	50
3.3.1.1	Percentage of Agreement ( $P_0$ ) and Agreement by Chance ( $P_C$ ).....	51
3.3.1.2	Cohen's Kappa ( $\kappa$ ).....	53
3.3.1.3	Kendall's Coefficient of Concordance ( $W$ ).....	55
3.3.1.4	Cronbach's Alpha ( $\alpha$ ).....	55
3.3.1.5	Intraclass Correlation Coefficient (ICC).....	56
3.4	BIT Agreement Results.....	57
3.4.1	Subtask Agreement.....	59
3.4.2	Drawing Subtask Agreement.....	60
3.4.3	Patient Response Agreement.....	62
3.5	Assessment Timings.....	67
3.6	Summary.....	68
<b>Chapter 4 - Experimental Infrastructure for Pen Based Data Capture.....</b>		<b>70</b>
4.1	Introduction.....	70
4.2	Handwritten Data Capture Requirements.....	70
4.2.1	Data Capture Peripheral.....	71

4.2.2	Sampling Rates and Spatial Resolution.....	73
4.2.3	Additional Pen Based Features.....	74
4.2.4	Data Capture in the Clinical Environment.....	74
4.3	System Infrastructure.....	75
4.4	Wacom Tablet and Portable Computer.....	77
4.5	Data Transmission Protocol and Storage.....	78
4.6	Pre-processing of Raw Test Response Data.....	82
4.6.1	Offset Calculation.....	83
4.6.2	Quadratic Interpolation of Data.....	85
4.6.3	Gaussian Low-Pass Filtering.....	87
4.7	Feature Extraction.....	91
4.8	Summary.....	94
<b>Chapter 5 - Task Definitions and Feature Extraction.....</b>		<b>96</b>
5.1	Introduction.....	96
5.2	Test Battery.....	96
5.2.1	Point Location Task.....	97
5.2.2	Line Bisection Task.....	98
5.2.3	Cancellation Tasks.....	99
5.2.4	Figure Completion Tasks.....	101
5.2.5	Figure Drawing Tasks.....	102
5.2.6	Drawing Profile Tasks.....	103
5.3	Performance Features.....	108
5.3.1	Generic Task Features.....	108
5.3.1.1	Overall Task Execution Time.....	108
5.3.1.2	Pen Contact/Pen Movement Ratio.....	109
5.3.1.3	Mean and Peak Pressure.....	109
5.3.1.4	Pen X/Y Tilt Standard Deviation.....	110
5.3.1.5	Pen Lifts within Drawing.....	111
5.3.2	Point Location Features.....	111
5.3.2.1	Target Distance Error.....	111
5.3.2.2	Quadrant Comparison.....	112
5.3.3	Line Bisection Features.....	112
5.3.3.1	Deviation from Actual Midpoint.....	113

5.3.3.2	Direction of Bisection.....	116
5.3.3.3	Quadrant and Line Length Comparison.....	116
5.3.4	Cancellation Features.....	116
5.3.4.1	Total Number of Cancellations.....	117
5.3.4.2	Number of Cancellations per Quadrant.....	118
5.3.4.3	Time per Cancellation.....	119
5.3.4.4	Processing Time per Quadrant.....	119
5.3.4.5	Processing Time per Cancellation in Quadrant.....	120
5.3.4.6	Sequence Starting Location (and Quadrant).....	121
5.3.4.7	Sequence Analysis.....	121
5.3.4.8	Duplications.....	124
5.3.4.9	Path Crossings.....	125
5.3.4.10	Inter-Cancellation Timings and Regression Analysis...	127
5.3.5	Generic Drawing Based Features.....	130
5.3.5.1	Segmentation and Component Count.....	130
5.3.5.2	Component Drawing, Premovement and Movement Times.....	136
5.3.5.3	Component Length and Euclidean Component Distance.....	136
5.3.5.4	Pen Velocity.....	137
5.3.5.5	Velocity Profiling.....	138
5.3.5.6	Pen Acceleration.....	139
5.3.5.7	Pen Hesitation Percentage.....	140
5.3.5.8	Image Width and Height.....	140
5.3.5.9	Line Curvature Deviation.....	141
5.3.5.10	Starting Position in Drawing.....	145
5.3.5.11	Component Corner Formations and Intersection Angle.....	146
5.3.5.12	Spatial Deviation.....	147
5.3.6	Drawing Task Specific Feature Extraction.....	148
5.3.6.1	Figure Completion Features.....	148
5.3.6.2	Figure Copying Features.....	150
5.3.6.3	Drawing Profile Features.....	152
5.4	Summary.....	153

<b>Chapter 6 - Analysis of Diagnostic Tasks.....</b>	<b>155</b>
6.1 Introduction.....	155
6.2 Trial Demographics.....	156
6.3 Data Analysis.....	160
6.4 Point Location Results.....	160
6.5 Line Bisection Results.....	162
6.6 Cancellation Results.....	167
6.6.1 OX1 Results.....	167
6.6.2 OX2 Results.....	173
6.6.3 Albert's Cancellation Task.....	175
6.7 Figure Completion Results.....	179
6.7.1 Overlay 1 (Diamond) BIT based patient groupings.....	179
6.7.2 Overlay 2 (Diamond) BIT based patient groupings.....	180
6.7.3 Overlay 1/Overlay 2 Result Ratio.....	182
6.7.4 Overlay 3 (Man) BIT based patient groupings.....	183
6.7.5 Overlay 4 (Man) BIT based patient groupings.....	184
6.7.6 Overlay 3/Overlay 4 Result Ratio.....	185
6.7.7 Overlay 5 (House) BIT based patient groupings.....	186
6.7.8 Overlay 6 (House) BIT based patient groupings.....	187
6.7.9 Overlay 5/Overlay 6 Result Ratio.....	187
6.7.10 Overlay 1 (Diamond) grade-based patient groupings.....	188
6.7.11 Overlay 2 (Diamond) grade-based patient groupings.....	188
6.7.12 Overlay 1/Overlay 2 grade-based result ratio.....	189
6.7.13 Overlay 3 (Man) grade-based patient groupings.....	190
6.7.14 Overlay 4 (Man) grade-based patient groupings.....	190
6.7.15 Overlay 3/Overlay 4 grade-based result ratio.....	190
6.7.16 Overlay 5 (House) grade-based patient groupings.....	191
6.7.17 Overlay 6 (House) grade-based patient groupings.....	192
6.7.18 Overlay 5/Overlay 6 grade-based result ratio.....	193
6.8 Figure Copying Results.....	193
6.8.1 Square BIT based patient groupings.....	193
6.8.2 Cross BIT based patient groupings.....	194
6.8.3 Star BIT based patient groupings.....	195

6.8.4	Cube BIT based patient groupings.....	197
6.8.5	Square grade-based patient groupings.....	198
6.8.6	Cross grade-based patient groupings.....	198
6.8.7	Star grade-based patient groupings.....	199
6.8.8	Cube grade-based patient groupings.....	199
6.9	Drawing from Memory Results.....	200
6.9.1	Square BIT based patient groupings.....	200
6.9.2	Cube BIT based patient groupings.....	201
6.9.3	Square grade-based patient groupings.....	202
6.9.4	Cube grade-based patient groupings.....	202
6.10	Shape Complexity.....	203
6.11	Kinematic profiling results.....	206
6.12	Feature Space Analysis and Classification.....	209
6.12.1	Classification Aims and Methodology.....	210
6.12.2	Principal Features.....	211
6.12.3	Bayes' Statistical Classification.....	211
6.12.4	Cluster Analysis.....	212
6.12.5	Kohonen Self Organising Map.....	215
6.12.6	Adaptive Resonance Theory.....	216
6.12.7	Classifier Summary.....	219
6.13	Summary.....	221
	<b>Chapter 7 – Summary and Conclusions.....</b>	<b>224</b>
7.1	Introduction.....	224
7.2	Summary of Research Programme.....	224
7.3	Suggestions for Further Research.....	228
	<b>References.....</b>	<b>232</b>
	<b>Appendix A - Stroke Test Subject Details.....</b>	<b>244</b>
	<b>Appendix B - Computer based assessment script used by therapists</b>	
	<b>administering test.....</b>	<b>248</b>

## List of Figures

Figure 1.1 : Rehabilitation feedback loop.....	3
Figure 1.2 : Computer based test system schematic.....	6
Figure 2.1 : Clock drawing and writing from a neglect patient.....	11
Figure 2.2 : Cross figure copying task.....	14
Figure 2.3 : Drawing dysfunction criteria.....	15
Figure 2.4 : Velocity profiles of an RCVA neglect subject.....	18
Figure 2.5 : Albert's cancellation task.....	19
Figure 2.6 : Number of omissions in Albert's cancellation task.....	21
Figure 2.7 : Hypothetical performance of (A) mild, (B) moderate and (C) severe neglect patients over two line lengths.....	23
Figure 2.8 : Two dimensional Euclidean distance cluster analysis.....	34
Figure 2.9 : 4 x 4 Kohonen map with 3 input and output nodes.....	35
Figure 2.10 : ART data flow schematic.....	37
Figure 3.1 : Letter cancellation task overlay.....	44
Figure 3.2 : Star cancellation task overlay.....	44
Figure 3.3 : Figure copying models.....	45
Figure 3.4 : Shape copying models.....	45
Figure 3.5 : Line bisection overlay.....	46
Figure 3.6 : Figure and shape copying models and representational drawing responses.....	47
Figure 3.7 : Distribution of marks with the BIT conventional battery.....	48
Figure 4.1 : 3 phase test system schematic.....	76
Figure 4.2 : Test apparatus.....	78
Figure 4.3 : Example test response file.....	81
Figure 4.4 : Test response file directory structure.....	83
Figure 4.5 : Offset calibration overlay.....	84
Figure 4.6 : Quadratic interpolation of line points.....	85
Figure 4.7 : Interpolation example.....	87
Figure 4.8 : Gaussian response profile ( $\sigma = 1$ ).....	88
Figure 4.9 : One dimensional Gaussian convolution mask (width = 2, $\sigma = 1$ ).....	89

Figure 4.10 : Gaussian low-pass cut-off frequencies.....	90
Figure 4.11 : Gaussian filtration example.....	91
Figure 4.12 : Total drawing time feature extraction flow chart.....	93
Figure 4.13 : X and Y coordinate profiles from a drawing from memory square task.....	94
Figure 4.14 : Stroke order in a square drawing from memory task.....	95
Figure 5.1 : Point location task overlay 1.....	97
Figure 5.2 : Line bisection task overlay 1.....	99
Figure 5.3 : OX cancellation task overlay 1.....	100
Figure 5.4 : OX cancellation task overlay 2.....	101
Figure 5.5 : Figure completion representational drawings.....	102
Figure 5.6 : Figure copying and drawing from memory models.....	102
Figure 5.7 : Line drawing task.....	105
Figure 5.8 : Line drawing task with distractors.....	105
Figure 5.9 : Line drawing task with distractors and pseudo-random ordering...	106
Figure 5.10 : Dot joining square task.....	106
Figure 5.11 : Visual field square drawing task.....	107
Figure 5.12 : Clock-face line drawing task.....	107
Figure 5.13 : Horizontal midpoint deviation calculation.....	113
Figure 5.14 : Interpolated crossing point.....	114
Figure 5.15 : Interpolation calculation points.....	115
Figure 5.16 : Target areas for cancellation tasks.....	118
Figure 5.17 : Raster and snake cancellation traversal method.....	121
Figure 5.18 : Location identification numbers (OX1 task).....	122
Figure 5.19 : Sequence correlation error example.....	123
Figure 5.20 : Barrel shift correctness of match operation.....	124
Figure 5.21 : Path with single crossing point.....	125
Figure 5.22 : Cancellation task assessment - timing definitions.....	128
Figure 5.23 : Timing regression for overall timing.....	129
Figure 5.24 : Velocity profiling example 1.....	132
Figure 5.25 : Velocity profiling example 2.....	133
Figure 5.26 : Original drawing.....	134
Figure 5.27 : 1/3 point axis locations.....	134
Figure 5.28 : 9 point sector definitions.....	135

Figure 5.29 : Exploded view with segment labels.....	136
Figure 5.30 : Velocity profile model of single line component.....	138
Figure 5.31 : Best line fit to drawn segment.....	142
Figure 5.32 : Error distance sample points.....	143
Figure 5.33 : Distance between best fit line and drawn line.....	143
Figure 5.34 : Starting position calculation.....	145
Figure 5.35 : Corner formation distance error.....	146
Figure 5.36 : Intersection angle calculation.....	147
Figure 5.37 : Component analysis of figure completion task.....	150
Figure 5.38 : Component analysis of figure drawing tasks.....	152
Figure 5.39 : Visual field based square drawing task labels.....	153
Figure 6.1 : Cumulative BIT scored obtained by neglect group.....	158
Figure 6.2 : Cumulative BIT scored obtained by stroke control group.....	158
Figure 6.3 : Neglect group shape complexity vs exclusion %.....	204
Figure 6.4 : Stroke control group shape complexity vs exclusion %.....	204
Figure 6.5 : Age matched control group shape complexity vs exclusion %.....	205
Figure 6.6 : Bayesian classifier performance.....	213
Figure 6.7 : K-means classifier performance.....	213
Figure 6.8 : Euclidean cluster classifier performance.....	214
Figure 6.9 : 20 x 16 Kohonen activation map, control groups.....	217
Figure 6.10 : 20 x 16 Kohonen activation map, neglect group.....	217
Figure 6.11 : Kohonen activation centres.....	218
Figure 6.12 : Kohonen activation centres, neglect sub-groups.....	218
Figure 6.13 : BIT score vs ART classification pattern.....	219
Figure 6.14 : Mean feature vector output patterns identified by the ART.....	220
Figure 7.1 : Data capture within a hospital environment.....	228
Figure 7.2 : Classifier architecture configurations.....	230

## List of Tables

Table 2.1 : Lesion location and task specific deficit.....	24
Table 2.2 : Test sensitivities of Rivermead BIT subtasks.....	25
Table 3.1 : Conventional subtask marks and cut-off scores.....	49
Table 3.2 : Two assessor agreement in figure and shape copying task.....	51
Table 3.3 : Fleiss Kappa statistic interpretation.....	54
Table 3.4 : Mean BIT results and marking variation.....	58
Table 3.5 : Subtask agreement for 10 sets of BIT responses.....	59
Table 3.6 : Drawing subtask agreement between assessors for 20 responses.....	61
Table 3.7 : Drawing object agreement between assessors for 20 responses.....	61
Table 3.8 : Test subjects ranked by disagreement on standard conventional task marking scheme.....	62
Table 3.9 : Test subjects ranked by disagreement on drawing tasks.....	63
Table 3.10 : Clockface drawing responses.....	64
Table 3.11 : 'Man' drawing responses.....	65
Table 3.12 : Butterfly drawing responses.....	66
Table 3.13 : Assessment and marking times for the conventional and behavioural BIT batteries.....	67
Table 4.1 : Wacom IVe data components.....	79
Table 4.2a : Initial decimal packet contents.....	79
Table 4.2b : Binary representation of packet and bit identification.....	80
Table 4.2c : Reconstructed data items.....	81
Table 4.3 : Task identification numbers.....	82
Table 5.1 : Position of target dot in point location task.....	98
Table 5.2 : Position of target line in line bisection task.....	98
Table 5.3 : Figure completion half image sizes.....	102
Table 5.4 : Figure copying shape sizes.....	103
Table 5.5 : Targets per quadrant in cancellation task.....	118
Table 5.6 : Archetypal sequences defined for OX1 cancellation task.....	122
Table 5.7 : Automatic component labelling.....	135
Table 5.8 : Figure completion components.....	149
Table 5.9 : Figure completion intersection and spatial measurement points.....	149

Table 5.10 : Figure copying components.....	151
Table 5.11 : Figure copying intersection and spatial measurement points.....	151
Table 6.1 : Neglect grade-based analysis groupings.....	157
Table 6.2 : Stroke control grade-based analysis groupings.....	159
Table 6.3 : Geriatric age analysis groupings.....	159
Table 6.4 : Percentage of each test group included in point location analysis....	161
Table 6.5 : Mean error distances (mm) to centre of target dot.....	161
Table 6.6 : Mean movement timing (sec) between dots.....	162
Table 6.7 : Horizontal bisection error direction from midpoint.....	162
Table 6.8 : Horizontal bisection error over 8 overlays for BIT defined groups...	162
Table 6.9 : Horizontal bisection error by line lengths for BIT defined groups....	163
Table 6.10 : Horizontal bisection error by quadrant for BIT defined groups.....	164
Table 6.11 : Horizontal bisection error by quadrant for 140mm bisection lines for BIT defined groups.....	164
Table 6.12 : Horizontal bisection error across 8 overlay (both line lengths) and for 140mm lines using severity marking scheme defined groups.....	165
Table 6.13 : Horizontal bisection error by quadrant for 140mm bisection lines using severity marking scheme defined groups.....	166
Table 6.14 : Significant BIT grouped OX1 feature results.....	168
Table 6.15 : OX1 task cancellation sequence starting target position.....	169
Table 6.16 : OX1 task cancellation sequence starting quadrant.....	169
Table 6.17 : OX1 task best matched model cancellation sequence.....	170
Table 6.18 : OX1 task best matched model cancellation sequence type.....	170
Table 6.19 : Significant grade-based grouped OX1 feature results.....	171
Table 6.20 : OX1 task cancellation sequence starting target position - grade-based grouping.....	172
Table 6.21 : OX1 task best matched model cancellation sequence - grade-based grouping.....	173
Table 6.22 : Significant BIT grouped OX2 feature results.....	174
Table 6.23 : Significant grade-based grouped OX2 feature results.....	175
Table 6.24 : Significant BIT grouped Albert's feature results.....	177
Table 6.25 : Albert's task cancellation sequence starting quadrant.....	176

Table 6.26 : Albert's task cancellation sequence starting quadrant - grade-based grouping.....	176
Table 6.27 : Significant performance based grouping Albert's feature results....	178
Table 6.28 : Percentage of group excluded from analysis, figure completion, overlay 1 (diamond).....	180
Table 6.29 : Significant BIT grouped overlay 1 (diamond) figure completion feature results.....	180
Table 6.30 : Percentage of group excluded from analysis, figure completion, overlay 2 (diamond).....	181
Table 6.31 : Significant BIT grouped overlay 2 (diamond) figure completion feature results.....	181
Table 6.32 : Significant BIT grouped overlay 1/overlay 2 figure completion feature results ratios (diamond).....	182
Table 6.33 : Percentage of group excluded from analysis, figure completion, overlay 3 (man).....	183
Table 6.34 : Significant BIT grouped overlay 3 (man) figure completion feature results.....	184
Table 6.35 : Percentage of group excluded from analysis, figure completion, overlay 4 (man).....	184
Table 6.36 : Significant BIT grouped overlay 4 (man) figure completion feature results.....	185
Table 6.37 : Percentage of group excluded from analysis, figure completion, overlay 5 (house).....	186
Table 6.38 : Significant BIT grouped overlay 5 (house) figure completion feature results.....	186
Table 6.39 : Percentage of group excluded from analysis, figure completion, overlay 6 (house).....	187
Table 6.40 : Significant BIT grouped overlay 6 (house) figure completion feature results.....	187
Table 6.41 : Significant grade-based grouped overlay 1 (diamond) figure completion feature results.....	188
Table 6.42 : Significant grade-based grouped overlay 2 (diamond) figure completion feature results.....	189

Table 6.43 : Significant grade-based grouped overlay 1/overlay 2 (diamond) figure completion feature results ratios.....	189
Table 6.44 : Significant grade-based grouped overlay 3 (man) figure completion feature results.....	190
Table 6.45 : Significant grade-based grouped overlay 4 (man) figure completion feature results.....	191
Table 6.46 : Significant performance grouped overlay 3/overlay 4 (man) figure completion feature results ratios.....	191
Table 6.47 : Significant performance grouped overlay 5 (house) figure completion feature results.....	192
Table 6.48 : Significant performance grouped overlay 6 (house) figure completion feature results.....	192
Table 6.49 : Significant performance grouped overlay 5/overlay 6 (house) figure completion feature results ratios.....	193
Table 6.50 : Percentage of group excluded from analysis, figure copying, square.....	194
Table 6.51 : Significant BIT grouped square figure copying feature results.....	194
Table 6.52 : Percentage of group excluded from analysis, figure copying, cross.....	195
Table 6.53 : Significant BIT grouped cross figure copying feature results.....	195
Table 6.54 : Percentage of group excluded from analysis, figure copying, star.....	196
Table 6.55 : Significant BIT grouped star figure copying feature results.....	196
Table 6.56 : Percentage of group excluded from analysis, figure copying, cube.....	197
Table 6.57 : Significant BIT grouped cube figure copying feature results.....	197
Table 6.58 : Significant grade-based grouped square figure copying feature results.....	198
Table 6.59 : Significant grade-based grouped cross figure copying feature results.....	198
Table 6.60 : Significant grade-based grouped star figure copying feature results.....	199
Table 6.61 : Significant grade-based grouped cube figure copying feature results.....	199

Table 6.62 : Percentage of group excluded from analysis, drawing from memory, square.....	200
Table 6.63 : Significant BIT grouped square drawing from memory feature results.....	201
Table 6.64 : Percentage of group excluded from analysis, drawing from memory, cube.....	201
Table 6.65 : Significant BIT grouped cube drawing from memory feature results.....	202
Table 6.66 : Significant grade-based grouped square drawing from memory feature results.....	202
Table 6.67 : Significant grade-based grouped cube drawing from memory feature results.....	203
Table 6.68 : Overlay 1 timing based measurements.....	206
Table 6.69 : Overlay 2 timing based measurements.....	206
Table 6.70 : Overlay 3 timing based measurements.....	207
Table 6.71 : Overlays 6 and 7 timing based measurements.....	207
Table 6.72 : Overlay 8 timing based measurements.....	208
Table 6.73 : Overlay 1 velocity based measurements.....	208
Table 6.74 : Overlay 2 velocity based measurements.....	208
Table 6.75 : Overlay 3 velocity based measurements.....	208
Table 6.76 : Overlays 6 and 7 velocity based measurements.....	209
Table 6.77 : Overlay 8 velocity based measurements.....	209
Table 6.78 : Principal features from computer based test battery.....	212
Table 6.79 : Kohonen network size effects on classification rate.....	216
Table 6.80 : ART classification patterns.....	219

# Chapter 1

## Introduction and Background

### 1.1 General Introduction

This thesis documents the investigation of the specification, development and evaluation of a computer-based assessment battery for the detection of visuo-spatial neglect in stroke patients. Extracting features measuring performance based on the pen movements made during a series of implemented 'pencil-and-paper' neuropsychological tasks, the accuracy and consistency of diagnosis compared with existing assessment metrics is increased through a standardised algorithmic approach to assessment. The understanding of the clinical condition is extended through the use of a range of novel time-based and constructional features which provide information hitherto unavailable for assessment. The utilisation of pattern recognition techniques introduce cross-feature multidimensional analysis enabling an examination of performance interaction both within and between individual tasks leading to automated diagnosis.

The impetus for the research and development of a new assessment of neglect is driven by two main factors. Firstly, research into stroke-related illness is a priority area in the UK Department of Health's Research and Development Plan [1], specifically the measurement and evaluation of rehabilitation outcome of stroke patients. Rehabilitation programmes involve the long-term care of patients, making them expensive to fund, therefore the careful evaluation of schemes will establish the most cost-beneficial treatment methods while the development of better techniques can also offer cost benefits. Equally, from the point of view of "quality of life", an effective rehabilitation program means a quicker recovery and a return to normal everyday activities for the patient. Accurate measurement of conditions and recovery progress is very important in facilitating this evaluation.

The second impetus is derived from an engineering perspective. Recent studies [2][3] have indicated that the computer-based algorithmic assessment of handwriting and pen-based movements can increase the accuracy and the diagnostic understanding behind a variety of clinical conditions. In particular, the extraction of pen movement and sequencing features

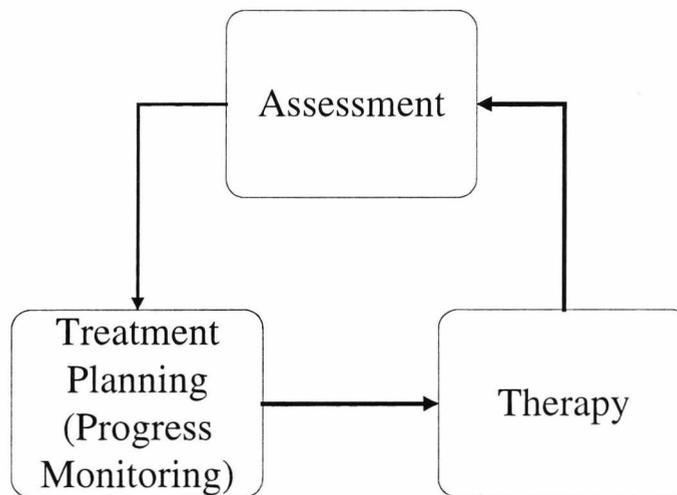
provide data which has not been available in traditional assessments of the tasks. These developments have been matched by the expansion of the use of computer technology within the area of healthcare for the monitoring, diagnosis and storage of performance data and patient history.

The research reported in this thesis is primarily concerned with the extraction, computation and classification of features from a series of pen-based neuropsychological tests. The research programme draws on both the theoretical basis and practical experience of a number of disciplines: Computer-based Image Processing, Psychology, Geriatric Medicine and Occupational Therapy. Each of these areas has contributed to the choice of the techniques implemented and interpretation of results obtained from the studies and experiments reported. While the use of a cross-disciplinary team is invaluable for input and guidance in all areas of the research, the wide range of supporting knowledge also presents a problem in reviewing the literature and interpreting the data obtained from trials and studies. The work reported in this thesis has therefore deliberately taken an engineering-based view of the system development and the application of computer-based techniques and algorithms to provide a solution.

## 1.2 Visuo-Spatial Neglect and Rehabilitation

The condition of visuo-spatial neglect relates to a dysfunction caused by brain damage [4]. The main effect of the condition is to cause subjects to fail to respond to stimuli in the visual field on the opposite side to the location of the lesion. Traditional testing [5] has exploited this effect by measuring the identification of objects within the visual field. Diagnosis of the condition is critical for the selection of a rehabilitation process specially devised to compensate for the effects of neglect. Inadequate detection of neglect at an early stage of therapy will result in performance deficit from the patient during rehabilitation and a failure to respond to treatment, prolonging the time-scale required for recovery and hence increasing associated costs.

Accurate assessment is important throughout the rehabilitation process to enable planning, monitoring and modification of treatment methods administered to individual patients. Figure 1.1 shows the basic feedback model of assessment, treatment planning and therapy. Assessment inaccuracy or the application of an unsuitable treatment programme at any stage of this model will cause a degradation of (and/or an extension to) the rehabilitation process.



*Figure 1.1 : Rehabilitation feedback loop*

Neglect is assessed by observations and formal testing procedures undertaken by a range of healthcare staff. Functional assessments obtained by Occupational Therapists establish deficiencies in everyday activities such as washing, dressing and eating [6]. Clinical assessment can determine the extent of neglect by observation of objects around a room and limb acknowledgement [5].

The focus of this research is a series of standard pencil and paper based tests which can be used to *quantify* performance. These tasks involve the completion or drawing of a task printed on a sheet of paper which is placed directly in front of the patient. Using a pencil or pen, typical tasks involve the cancelling of printed targets or the drawing of simple geometric shapes. The responses of these tasks are then evaluated by therapists or trained assessors. Due to the simplicity of the required apparatus, these tests can be used in a variety of confined hospital environments, for example while sitting up in bed or whilst seated in a wheelchair.

While the performance effects of neglect subjects completing these neuropsychological tests are well documented [4], a number of potential problems exist with the testing and assessment methodology:

- 
- **Fatigue** - test batteries often cause tiredness within patients which in turn causes modification of test performance. This effect is particularly prevalent in a geriatric population. Test administration and assessment are also affected by therapist fatigue and sometimes by complacency caused by overfamiliarisation with the testing procedure.
  - **Subjective Assessment** – with ambiguous assessment guidelines, patient performance is unstandardised between both therapists and subjects introducing repeatability errors in cross-patient comparison and performance monitoring over time.
  - **Resource Intensive** – the administration and assessment of tests involves the utilisation of a trained therapist. Although task specific, existing tests for visuo-spatial neglect typically take up to 3 hours to produce an assessment.
  - **Accuracy** – Even with strictly defined marking criteria, test responses are still subject to human error in assessment, such as score miscalculation and criteria application which can affect the performance rating from a particular task.
  - **Distress to Patients** – the processes of being tested, and recognising poor performance, can cause distress and feelings of frustration for a patient.

Performance and assessment inconsistencies caused by one or a combination of these problems will often prevent an accurate measurement of the patient's ability and, in terms of the rehabilitation feedback loop, mask the extent of a patient's progress. The research described in this study includes an evaluation of the current pencil and paper assessment techniques to determine the level of agreement between test evaluations.

### 1.3 Computer Based Testing and Evaluation

In implementing a computer based capture and analysis system, the first consideration must concern the operational requirements. The aims of using computer based data capture and analysis techniques for the examination of test responses are threefold:

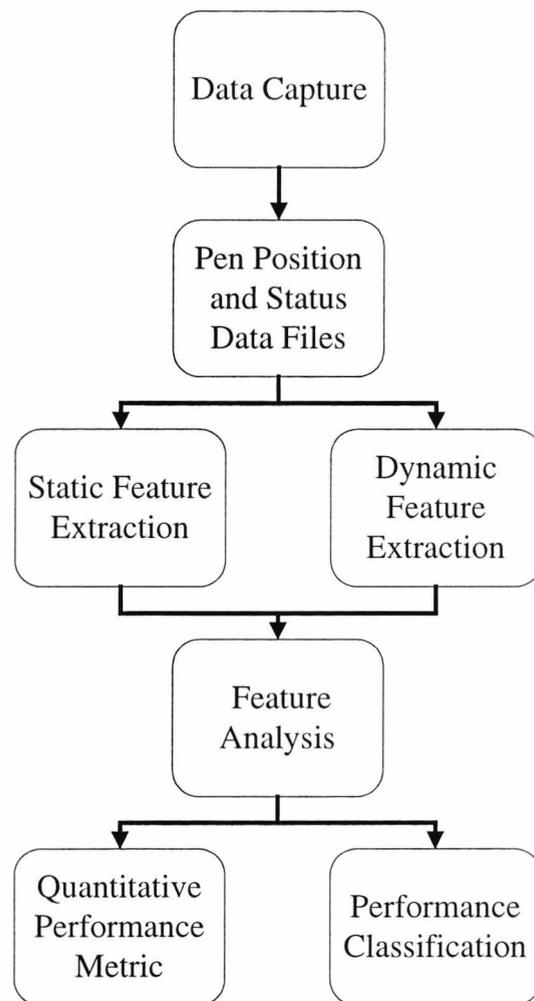
- Improve accuracy and assessment consistency – the subjectivity and inconsistency described in the previous section can be removed by assessing performance with respect to a predetermined set of rules and criteria.

- The extraction and diagnostic analysis of new dynamic features – Two distinct types of features are able to be extracted from drawing response: *Static* features relate to the outcome of the drawing (i.e. measurements taken from the completed drawings, for example total distance drawn and size of drawing) whereas *Dynamic* features measure timing and constructional ordering of drawings, for example the total time taken to draw and the order in which the sides of a geometric shape were drawn. These features are not available within conventional assessment methodologies and may provide additional diagnostic features and performance measures.
- Performance classification and quantification – the use of pattern recognition techniques may identify performance *similarities* between responses from groups of patients and hence provide automatic classification and diagnosis of an unknown test response.

Capturing data using a computer peripheral such as a graphic tablet allows a test subject to interface with the computer without modification of the standard pencil and paper based test methodology. Other benefits of the computer-based analysis carried out here include the ability to store raw pen data (such as coordinates), features and classifications and consistently monitor performance over time. As the test response from a patient is stored as a list of pen coordinates, this data can be replayed and reassessed without the need to retest with the test subject. Most importantly, additional features can be extracted from a test response at any stage of the study simply by replaying the stream of coordinates.

Figure 1.2 shows the standard implementation schematic for the capture and analysis of pen based features adopted here [7]. Pen position data is captured from the input device and stored as a stream of coordinates. Traditional static features are then extracted (with increased accuracy) alongside the novel dynamic based features. From both sets of features, performance based classification can be obtained using standard pattern recognition and data clustering techniques [8].

The final stages of this schematic show the generation of a classification and performance metric. The final outcome must also be verified against an existing test battery and a clinical evaluation of a test subject to ascertain the performance and reliability of the devised system.



*Figure 1.2 : Computer based test system schematic*

## 1.4 Thesis Structure

Following this brief overall introduction, the thesis is divided into a further six chapters describing the following elements of the research programme:

Chapter 2 introduces the practical and theoretical background to the research, investigating the condition of visuo-spatial neglect in stroke patients. Existing assessment methods are examined alongside the physical symptoms of the condition. The application of computer-based assessment methods are discussed with reference to the detection of neglect and suitable pattern recognition techniques for the classification of features extracted from the tasks are examined. Finally, the design aims for the research are defined.

---

Chapter 3 presents the findings of an interrater reliability investigation of the current standard for assessment of neglect, the Rivermead Behavioural Inattention Test (BIT). The BIT infrastructure is described in detail and an assessment made on the agreement between raters by comparing overall battery score and sub-task score identifying which task produces the most disagreement and why this occurs. The extent of misclassification of drawing tasks is investigated.

Chapter 4 describes the implementation options for a pen based capture system. Hardware devices to facilitate data capture are investigated and sampling and storage requirements are presented along with pre-processing and feature extraction routines to examine the captured data.

Chapter 5 describes the implemented computer-based test battery and individual static and dynamic features that are extracted from the pen-based data. Defined algorithmically, each feature is examined in relation to known response characteristics defined in the literature.

After detailing the demographics of the test subjects participating in a trial of the computer-based test, Chapter 6 presents the results from each of the features and the diagnostic capabilities of each of the battery sub-tasks are established. The correlation of each feature result against the obtained BIT classification is also established. The principal features discriminating between test groups are identified. The chapter also presents the findings of a feasibility study to assess a series of automated pattern classification techniques using the significant features extracted from the computer-based test system. The choice of classifier architecture is investigated with respect to increasing the accuracy and diagnostic ability of the system.

Chapter 7 draws some conclusions about the research programme reported here and includes suggestion for further work.

## Chapter 2

# Computer Based Diagnosis of Visuo-Spatial Neglect: A Review

### 2.1 Introduction

This chapter presents an overview of the theoretical background to the research and defines the developmental aims and objectives. The condition of visuo-spatial neglect in stroke patients is investigated including the physical symptoms and the effects on everyday activities. The importance of correctly diagnosing the condition is highlighted by the range of specific rehabilitation strategies to compensate for neglect. Lacking a strict definition of severity, the condition is difficult to diagnose accurately leading to a large variation in reported incidence of neglect between studies. Current clinical methods designed to standardise assessment are discussed with relation to neuropsychological impairments. Expected outcomes from these tests are described.

The second strand of the review examines the role of computers in neuropsychological assessment and in particular, the possibilities for extracting two types of feature data using digitised handwritten responses. Firstly, by applying an algorithmic approach to assessment, test response drawings and markings (*static features*) can be measured more accurately, with greater resolution and with consistency. Secondly, with reference to the literature on pen kinematics and dynamics, the types of additional movement and time-based features (*dynamic features*) designed to emphasise the symptoms of neglect are assessed.

Finally in this chapter, a range of pattern recognition methodologies that can be used to classify a feature set extracted from the devised tests are investigated. The advantages of using a computer based system within a clinical environment are discussed with reference to rehabilitation performance and an increased understanding of the condition of neglect. From the review, a series of implementation objectives are defined.

## 2.2 Stroke and Visuo-Spatial Neglect

Stroke is caused by either a blockage of blood flow or, less frequently, the haemorrhaging or rupturing of blood vessels within an area of the brain. The level of oxygen reaching the affected area is then decreased and damage is sustained. Stroke is the third commonest cause of death in the UK, after ischaemic heart disease and cancer, with an approximate annual incidence in the UK of 2 per 1,000 of the population. The true incidence of stroke in the UK is not known as it is estimated that up to 25% of strokes and transient ischaemic attacks are not reported. As 67% of strokes occur in persons over the age of 65 then stroke is predominately related to care of the elderly [9].

Visuo-spatial neglect is a condition that may occur following a stroke or less commonly, head injury [10]. The main effect of neglect is to cause the patient to fail to respond to or report visual stimuli contralateral (opposite side) to the location of the cerebral lesion. Thus a patient with neglect resulting from a right hemisphere lesion will fail to respond to a stimuli placed to the left of his visual field [11]. Neglect is more commonly associated with a right hemisphere lesion (right Cerebral Vascular Accident - CVA) where the symptoms are more pronounced [12], although less frequent and less severe cases of neglect do result from a left hemisphere lesion [13]. Definition and understanding of the condition has grown over the past two decades in both the clinical and neuropsychological fields. Originally thought to be a disturbance in visuo-spatial processing, more recent research has concluded that neglect is a heterogeneous collection of dysfunction in areas of motor, sensory and intellectual performance [14].

Neglect is not an 'all-or-nothing' condition [15]; the severity of neglect varies depending on the location and volume of the lesion, but neglect is commonly associated with a lesion in the right hemisphere posterior parietal region. The incidence rate of neglect varies considerably mainly due to an inadequate definition of the condition and the lack of standardised testing procedures. Halligan and Robertson [14] present a review of reported incidences of neglect which varied between 12% and 95% of assessed stroke patients. These figures were obtained by using a range of assessment techniques, all with differing sensitivities to the condition and without strictly defined and standardised marking criteria. In several cases, tests that were not designed to identify neglect (such as Parkinson's based writing tasks) have been used clinically with the assumption that similar dysfunctions were being assessed [16].

### 2.2.1 Effects on Everyday Activities

Depending on the severity of the condition, neglect can have a debilitating effect on everyday activities. Tasks such as washing and dressing are affected as a patient with neglect may fail to acknowledge their contralateral limbs and respond to objects placed to one side of space [17]. Other tasks such as eating may result in the patient leaving food on one side of the plate. Neglect patients also have difficulty in reading from books, clocks, and watching television. Writing and drawing performance is affected, with neglect subjects compressing their left hand visual field into the right (intact) side of the drawing. Writing tends to be right justified on the page [18]. Figure 2.1 shows typical examples of these two effects. In Figure 2.1a a neglect patient was asked to draw a clock face. The compression of numerals to the right hand side of the clock face is evident. This right hand side bias can be seen in normal writing performance (Figure 2.1b) where the text is justified to the right margin.

Whilst these tasks may cause frustration or leave a patient unaware of their actions to one side of their visual field, navigational disorientation means that activities such as walking and crossing the road present new dangers to which a person is not able to respond. Once a patient becomes aware of the neglect, he will learn to compensate by relocating the centre of his field of vision. Overcompensation to the attentive (non-affected) side sometimes causes failure to respond to a person communicating within the contralateral field.

The accurate detection of the presence and severity of neglect is critical for two reasons. Firstly, it enables the correct rehabilitation schemes to be used within a hospital environment which can be specifically tailored to the needs of the neglect patient. Secondly, without detection, a patient may be placed in a situation or confronted with activities where it would be dangerous without specific recognition of the condition [19]. In many cases, these effects are only present for the first few weeks post-stroke, although inattention to extreme right sided stimuli may continue for several months. The average length of stay within a stroke unit is 3 months, so the effects will still be present when discharged from hospital and the patient is faced with normal everyday activities without the assistance of trained hospital staff.

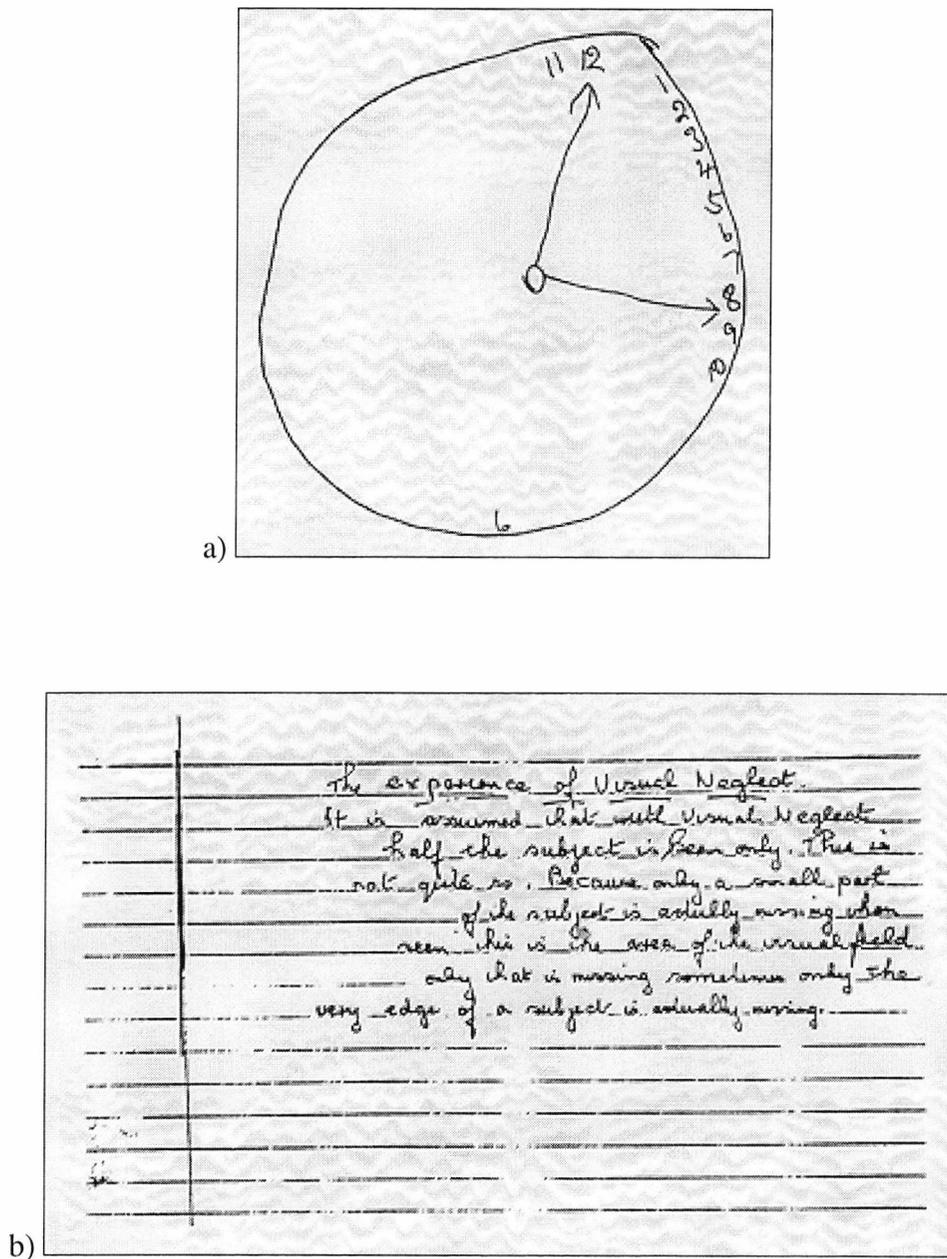


Figure 2.1 : Clock drawing and writing from a neglect patient

## 2.2.2 Rehabilitation of Neglect

The types of rehabilitation scheme used to treat neglect are implemented clinically across the range of therapy activities. Therapists attempt to make the patient aware of the condition and learn to compensate for the inattention by directing (or cueing) visual attention to objects placed in the inattentive field. Typical configurations for these rehabilitation exercises are the location of targets on a sheet of paper or objects within a workspace. Many of these tasks are

coupled with functional exercises administered by Occupational Therapists which require patients to locate and use everyday objects found in kitchen or bathroom activities. An overview of techniques used by therapists are presented in Ladavas, Menghini and Umilta [20], Robertson [21] and Lennon [22].

Length of treatment varies depending on severity of neglect and the recovery rate of an individual patient. Studies [23][24] indicate that the typical period for the condition to stabilise is six months and that an assessment of neglect stability should be determined by using a test battery on two separate occasions at least a month apart. This time estimate may however be affected by other dysfunctions a patient may have. The next section examines some of the standard techniques and tests used for the assessment of neglect.

### 2.3 Testing for Neglect

Many tests have been devised by clinicians, therapists and neuropsychologists to establish the presence and severity of neglect within a patient. Whilst there is little (or no) evidence of standardisation linking test performance to neglect severity across the entire range of devised tasks, the underlying assessment criteria is to establish and monitor performance differences between the left and right visual fields. Indeed, it is the case that some of the tasks are more sensitive to neglect detection than the others.

Halligan and Robertson [14] define four categories of tasks designed to test neglect:

- **Behavioural and Functional** - observation and assessment of everyday activities and object description in all areas of the visual field.
- **Drawing and Copying** – patients asked to copy and draw from memory geometric or representational shapes on a sheet of paper placed in the centre of their visual field.
- **Cancellation and Visual Search** – another test completed using a pencil and paper. Here the test subject has to locate and cancel (or mark) specific objects either with or without visual distractors. As the overlay is placed in the centre of their visual field, objects are located in both intact and inattentive fields for a neglect patient.

- **Line Bisection** – a further paper based test. Given a straight line of specific orientation, the test subject is required to mark (bisect) the midpoint of the line. Overlays may consist of a single or multiple lines.

Subtask configuration and typical performance characteristic of neglect patients using the pencil and paper based tasks are investigated in detail in the next sections.

### 2.3.1 Copying and Drawing Tasks

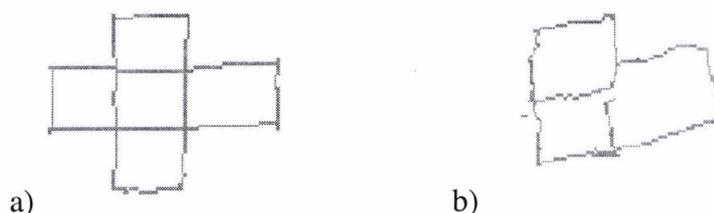
The diagnostic ability of using performance measurements from copying and drawing tasks to establish the presence of neglect has produced two conflicting opinions to the usefulness of the task. While studies such as Ericsson et al. [25] found that drawing performance decreased in proportion to levels of dysfunction, others have questioned the diagnostic properties of static features extracted from drawing tasks and have obtained results which do not correlate with other tests of neglect [26]. The use of dynamic constructional features from drawing tests [2] have, however, produced significant results for the assessment of neglect. Used individually or in conjunction with the static features, drawings do contain important clinical indicators.

Typical implementation of these tasks involve the copying [27] or drawing from memory [28] of a variety of simple geometric shapes (such as a square, star and diamond) or representational drawings (for example, a house, a man, and a tree). Modifications to this basic test methodology are a completion-based task, where half an image is presented to a test subject who is required to draw in the mirror image [29].

A general drawing characteristic with a right hemisphere neglect subject is the omission of left hand side components of the copied or drawn attempt [30] [31]. An example of this is shown in Figure 2.2.

Figure 2.2 (a) is a drawing from an age matched control subject. A neglect subject test attempt is displayed in Figure 2.2 (b) clearly showing component omission from the drawing. While the evidence of omission shown in the example demonstrates clear differences between test groups, this type of response is only evident in moderate to severe cases of neglect. In cases of very severe neglect, figure copying tasks that require the construction of geometric or

representational shapes that comprise of many components or are difficult to visualise (such as a three-dimensional cube) often produce an unrelated or no response from the test subject.



*Figure 2.2 : Cross figure copying task*

Apart from task complexity, a major deficit with the assessment of drawing based tasks is the subjectivity introduced through the absence of an objective marking scheme. In an attempt to rectify this Andrews et al. [32] examined drawing performance across a range of stroke subjects. In devising six conditions for drawing failure, categories for assessment could be referenced for marking guidelines. Again, application of these categories to real data is not clearly defined. This means that marking of individual responses are still reliant on the subjective judgement of the assessor as no standardisation between assessors is established.

Figure 2.3 shows examples of the six failure conditions. In drawings (a) to (e), a representational drawing of a house is required to be copied. Response (f) is a drawing of a man.

The defined categories are :

- a) disorganised drawing
- b) perseveration (multiple drawing of a single component)
- c) simultaneous agnosia
- d) overcopying
- e) unrelated activity
- f) visuo-spatial neglect

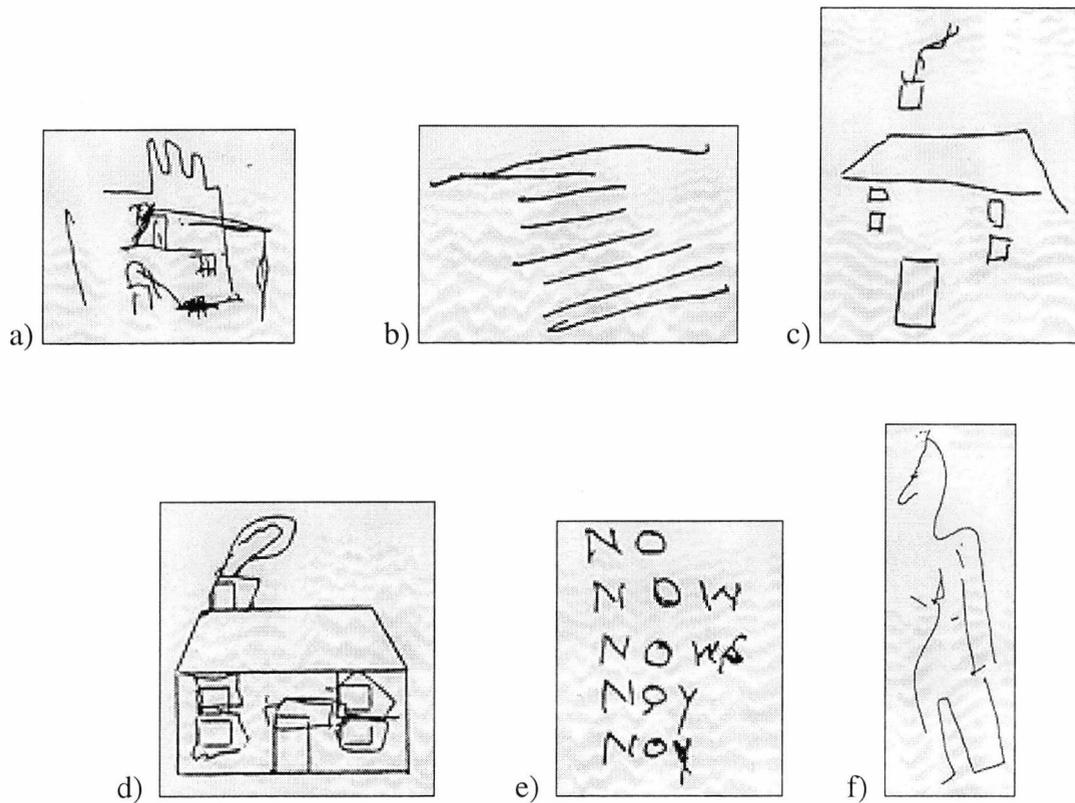


Figure 2.3 : Drawing dysfunction criteria (Andrews et al. [32])

Kirk and Kertesz [33] and Swindel et al. [34] both devised a series of low level assessment criteria for examination of drawing quality. In the former study, three sets of features were used to assess drawings:

- **Drawing accuracy** : drawing overlap, spatial relationship, drawing simplification, angle production, perseveration of lines, tremor within lines, perspective of three dimensional shapes.
- **Drawing positioning** : orientation of drawing, position drawn on page.
- **Item count within drawing** : components, angles, redrawn lines, lines crossing, lines joining, extra markings on the paper.

Swindel et al. used similar assessment criteria dividing features into qualitative (symmetry, components present) and quantitative (drawing size and spatial placement). Both scales relied on individual subjective judgement from an assessor.

Thurmond and Hancock [35] examine the effects of figural complexity with respect to drawing task response accuracy. The study concluded that as the complexity of the shape increased, the more difficult it was for a neglect subject to produce an accurate response, producing a better discrimination between subject groups. Ericsson et al. [25] found that complex shapes such as a cube or pentagon are the most sensitive to changes in cognition, whereas drawing a square or writing a sentence provide little discrimination. Accuracy of shape perception within all test subject groupings was improved when an outline of the shape to be copied was presented rather than a solid representation. Peru et al. [36] examined the ability of neglect patients when required to copy whole, half and chimeric shapes. The study showed that subjects based their drawing reference on the right components of shapes. When these were absent, inaccuracies in copying ability resulted within a neglect population.

One of the most widely used tests for analysis of visuo-spatial neglect and other cognitive dysfunction [37] conditions is the clock drawing task (Figure 2.1a). Several variations of this task exist, the diagnostic properties of which are investigated by DiPellegrino [38]. Using a single neglect case, DiPellegrino's study showed that when the test subject was required to draw a clock face and place the numerals in the correct positions *from memory*, then all twelve numerals were positioned to the right hand side of the dial between the 12 o'clock and 6 o'clock locations. The same effect occurred when *copying* an image of clock face. Both these tasks demonstrate the standard neglect performance modification on the left hand side of a drawing.

In the study of Alzheimer's patients performing the clock drawing task, Cahn [37] tested the stability of the test across a wide age range of control subjects. The results showed that performance did not deteriorate with age, thus indicating the task's suitability for diagnostic use within a geriatric population. However, in a cross-task comparison for the assessment of neglect, Ishiai et al. [26] found that clock drawing performance did not correlate with neglect severity identified by cancellation and bisection tasks. The conclusions of the study supports the use of other tasks such as cancellation and bisection, but demonstrated that the clock drawing task was not an accurate diagnostic tool. Other studies have used handwriting and drawing output from this task for the diagnosis of clinical conditions including Alzheimer's disease [39] and Parkinson's disease [40][41].

The constructional aspects of drawings have been examined through the use of computer based recording of pen movements. Smith and Fairhurst [42][43] examined the use of both static and dynamic features extracted from the drawings made by a range of test subjects

including children and hospital patients with dementia. By implementing a set of tools and feature extraction routines, both accuracy and consistency in static assessment were obtained. New dynamic or timing based features also extracted from the test response revealed differences between test populations which were previously unobtainable. This work was supplemented by Clar [44] and Higson [45] for the analysis of dyspraxia through an assessment of Beery Test [46] responses.

Kinematic profiling of movements made by right CVA neglect patients was explored by Mattingley et al. [2]. By obtaining a series of horizontal pen movements across a graphics tablet surface, the group found differences in severe neglect patients from a normal population when a pen movement was made from the right hand side of the page (their intact field) to the left (their inattentive field). By assessing the velocity profile, severe neglect patients drew more slowly and with a profile which was dissimilar from the standard bell-shaped velocity profile obtained when drawing a straight line [47]. The group were also slower to reach the peak velocity within the profile (indicating a longer acceleration phase). This peak velocity was lower than that produced by a normal population and the velocity profile contained more submovements indicating poor force control. Patients with mild neglect exhibited similar performance characteristics to a control population. Figure 2.4 shows two velocity profiles of pen movement from the Mattingley study. The first profile (a) is from a control patient while the second (b) is from a neglect subject. The difference in peak velocities, timings and profiles of leftward movement is apparent. Similar velocity profile results were found by Konczak and Karnath [48] examining the movement times to reach targets from a base position.

Kinematic algorithms and features for the assessment of pen based movement, are explored in detail in Section 2.4.2 , along with other time based dynamic features.

### 2.3.2 Cancellation and Visual Search Tasks

Many standardised implementations exist of cancellation and visual search tasks for the detection of neglect. One of the first developed, and widely regarded as a standard for assessing neglect, is the Albert's Cancellation Task [49]. In this task, a test subject must cancel 40 lines printed in a pseudo-random orientation on a single sheet of paper (Figure 2.5). Other tests include the star cancellation task [50] which introduces distractors (letters and large stars) amongst the cancellation targets (smaller stars).

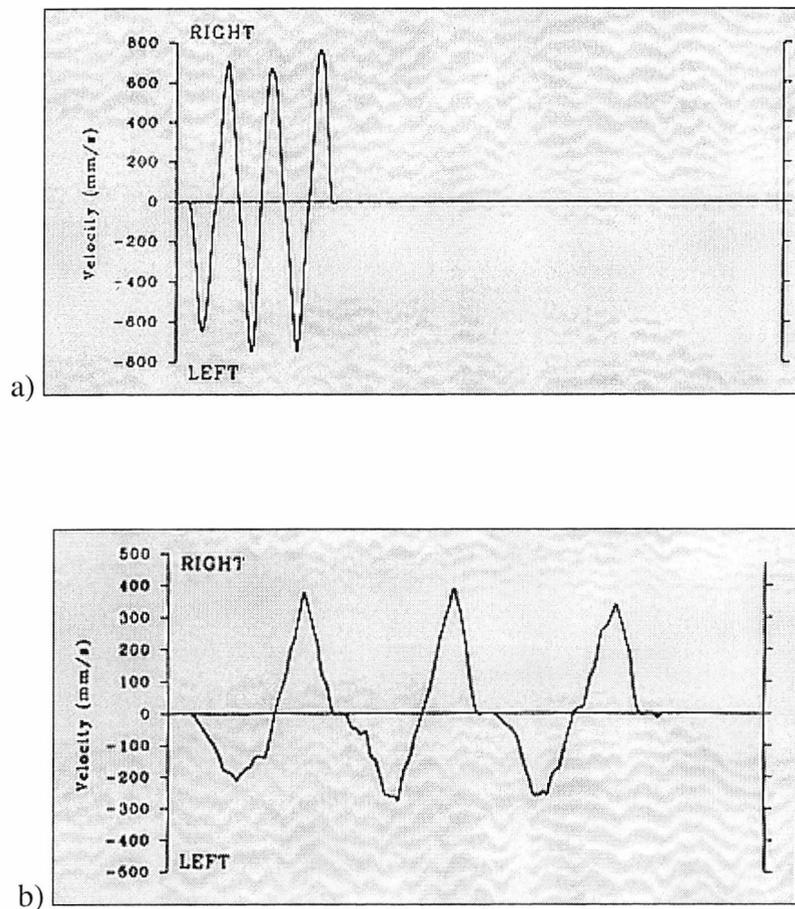
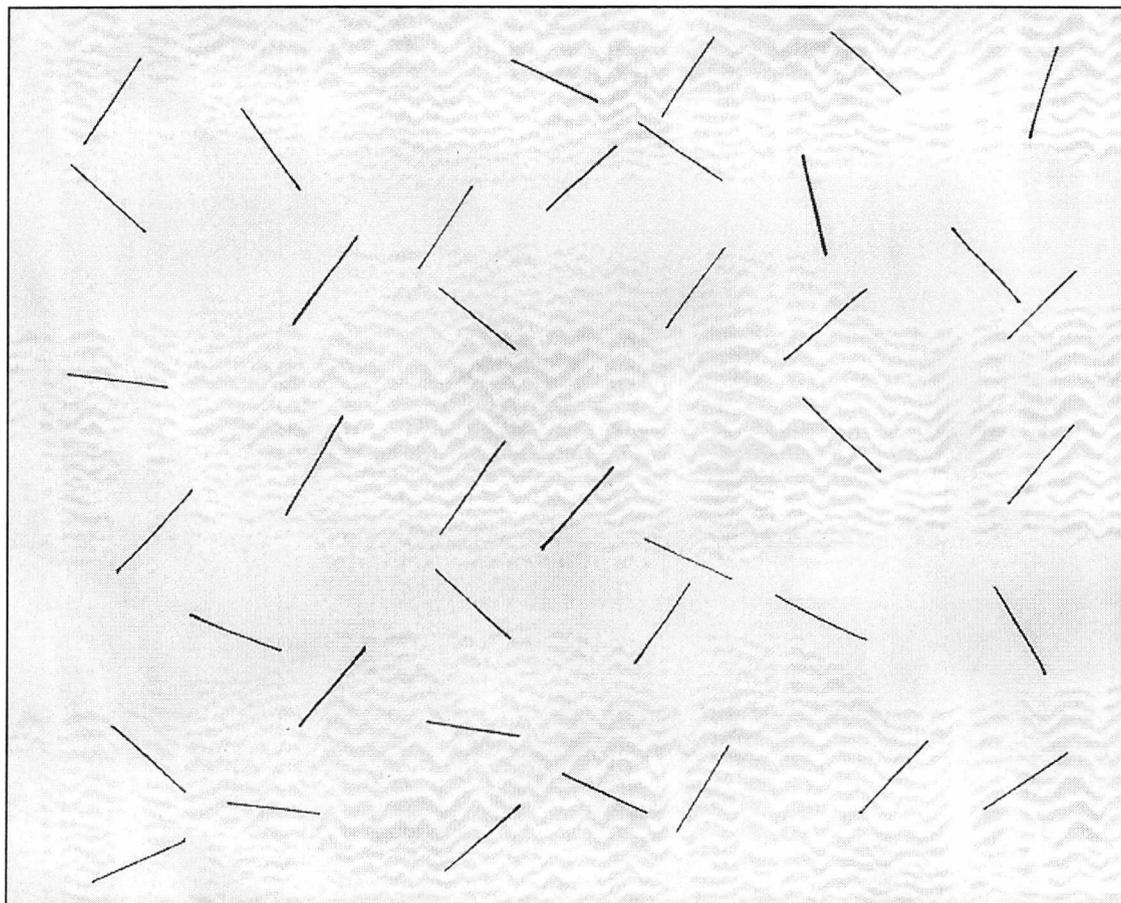


Figure 2.4 : Velocity profiles of an RCVA neglect subject (Mattingley et al. [2])

With all these cancellation tasks, typical right CVA neglect performance results in failure to cancel the targets on the left hand side of the overlay. The severity of the neglect can be assessed by counting the number of targets not cancelled on the complete overlay [51]. Chatterjee et. al [52] propose a power function to express neglect severity in relation to cancellation performance and number of targets on the cancellation sheet. Thus:

$$\text{Targets cancelled} = K \times (\text{Targets presented})^B \quad (2.1)$$



*Figure 2.5 : Albert's cancellation task [49]*

In this study, improvements in performance over time were noted by the increase in the constant,  $K$ . Chatterjee et al. reasoned that, as the exponent,  $B$ , did not change across the same multiple test attempts, an aspect of the neglect dysfunction remained the same. Chatterjee also concluded that an increase in the number of targets contained within the cancellation task also increased the sensitivity to detecting neglect. Studies have also shown that there is a timing increase for completion of the test overlay in proportion to the number of cancellation targets [53].

The power function performance relationship was derived from a series of cancellation tasks without distractor targets (without selective attention). Kaplan et al. [54] found that an increase in the number of distractors on the cancellation overlay caused neglect patients to omit more of the targets and hence increased the sensitivity of the task. Further experiments indicated that if the objects used for cancellation targets and distractors are similar, this also increased the sensitivity in detecting neglect and slowed completion time. Henderson [55]

demonstrated this effect using two cancellation tasks, the first where the shape of the targets were similar ('C' as a target, 'c' as a distractor) and the second where the shape differed in each case ('A' as a target, 'a' as a distractor). Accuracy was increased for the second test where the shape differed significantly (using the 'A'/'a' characters). Geldmacher [56] also showed this effect by conducting a series of experiments using different target and distractor characters. Using the letters 'I' and 'O' as targets and the letter 'L' as a distractor, cancellation accuracy for 'I' was lower than for the letter 'O'. In an earlier study, Geldmacher [57] investigated the ratio between the number of targets and distractors contained on an overlay. The findings of the study showed that all test subject groups were slower and less accurate when the ratio of distractors to targets was higher. Cancellation tasks using a large number of distractors such as the Bells test [58] have been shown to be more sensitive to neglect than the Albert's test.

Several studies have investigated the cancellation performance of neglect patients dividing the cancellation overlay into quadrants rather than on left and right visual fields. Using a standard Albert's Task with Right CVA neglect patients, both Halligan and Marshall [59] and Mark [60] found that as in previous studies, more omissions were made on the left hand side of the task. Quadrant analysis showed that the greatest number of omissions occurred in the bottom left hand quadrant of the overlay. Figure 2.6 represents the findings of Mark. The squares represent the locations of the cancellation targets within the task and the number of omissions made at each point. The diagram shows how the number of omissions can be represented by a series of diagonal contours running across the overlay from the top left to the bottom right.

The timing and constructional properties of cancellation task completion have been investigated by observed and videoed analysis. Search patterns and cancellation strategies have been analysed by Chatterjee et al. [61] by forcing test subjects to cancel in horizontal and vertical movements. While typical target omissions on the left hand side of the overlay were reported, regular patterns of cancellation with movement predominantly in the vertical plane were made by the single neglect patient used as a case study for the trial. Introducing more targets in the horizontal plane made no difference to the regularity of the cancellations. As with other measures, the cancellation sequence may be sensitive to the severity of neglect patient exhibits. Further investigation of this feature across a larger population of test subjects is therefore necessary.

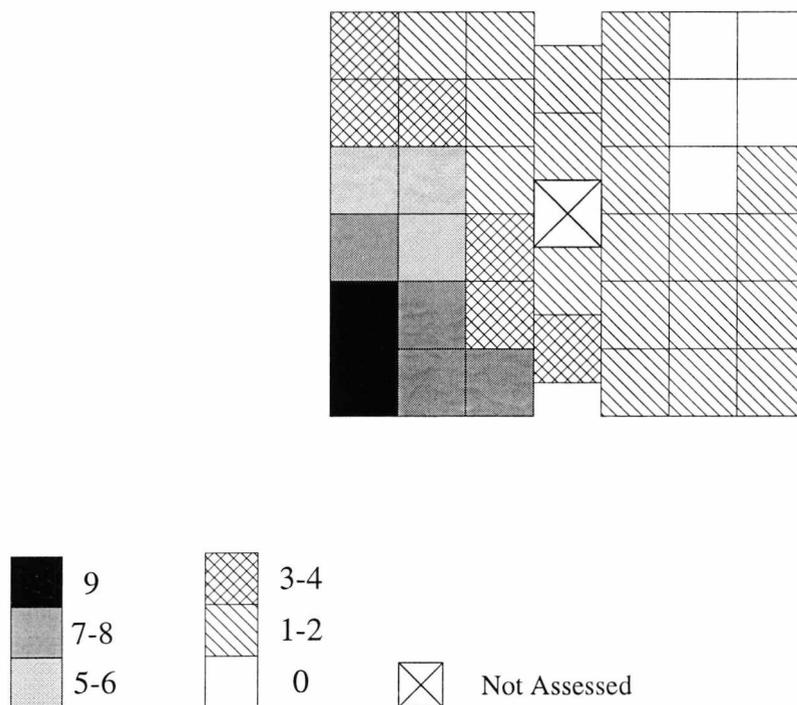


Figure 2.6 : Number of omissions in Albert's cancellation task

Age-matched control performance for the cancellation task was established by Geldmacher et al. [62] using a letter cancellation task containing 10 targets and 45 distractors. 26 % of all control subjects failed to completely cancel all of the targets. However, 74 % of these subjects only omitted a single target, leading to a figure of only 3.9 % of age matched subjects failing to cancel more than 2 targets. The study found that more cancellation omissions were made by older subjects and that omissions were generally made on the right hand side of the test overlay. Normal non-geriatric population performance on the cancellation task shows overall greater accuracy, confirming the effect of age on task performance [63].

### 2.3.3 Line Bisection Tasks

The bisection task is widely used as a simple clinical diagnostic test for neglect and results and observations are well documented within the medical and neuropsychological literature [64][65]. Right CVA neglect patients characteristically tend to bisect the line to the right of the centre point [66][67] which can be explained by an attentional bias causing the patient to

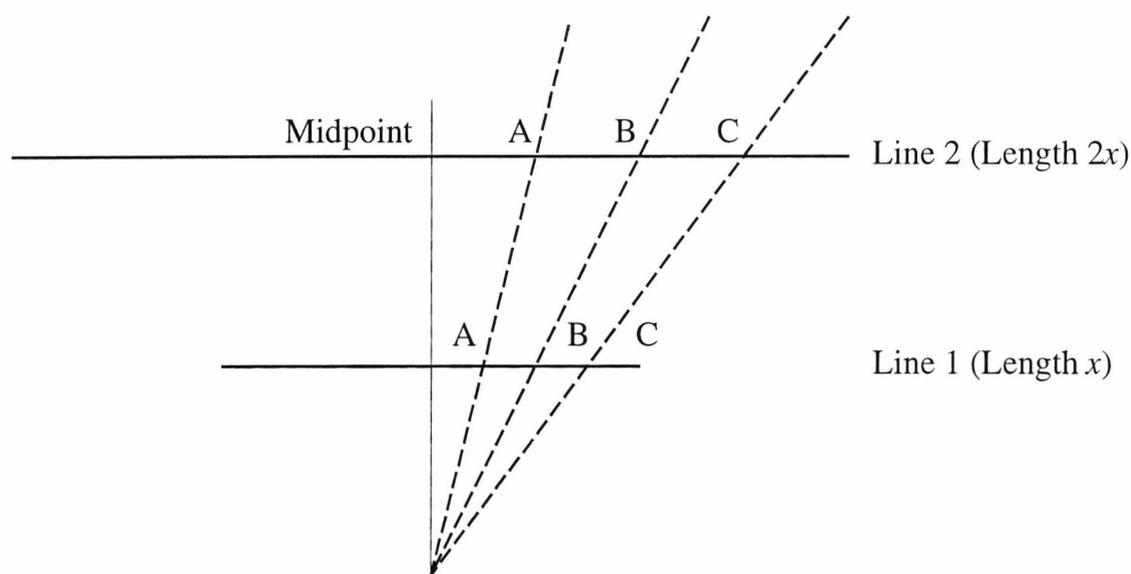
overestimate the left hand segment of the line and consequently underestimate the right segment. Several studies have focused on obtaining normal performance for the task [68][69][70][71] finding that accuracy amongst healthy adults at locating the midpoint is high with slight deviations usually to the left of centre. Chokron and De-Agostini [72] examine this slight left deviation in relation to normal reading and scanning direction. By analysing the normal performance of 30 French (Western, left to right scanning) and Israeli (Arabic, right to left scanning) subjects, the slight left deviation was again found in the French subjects but for the Israeli subjects a deviation to the right was observed. Scarisbrick et al. [73] assessed the effect of normal writing hand on normal performance within a left to right scanning normal population. Again a slight deviation to the left of the midpoint was observed regardless of dominant hand.

These studies also assess the effect of line length on accuracy, which can be exploited in testing neglect patients. Bisection error is linearly related to the overall line length in that the longer the target line, the greater the bisection error. Two studies have looked at this effect within a population of neglect patients [74][75] and have found that the bisection-error-to-line-length ratio is greater for neglect patients than it is for a normal population. Attempts have also been made to model the error mathematically as a power series [52] whereas recent studies [76] have developed a computer-based connectionist model to produce a quantitative analysis of the bisection results.

Halligan and Marshall [77] found that the severity of neglect can be directly obtained from the bisection error distance (again scaled by the line length effect), linked by the formula:

$$\frac{\text{Bisection error distance}}{\text{Line length}} = \text{Neglect severity} \quad (2.2)$$

This effect is shown in Figure 2.7, using hypothetical performances by mild, moderate and severe neglect subjects over two line lengths  $x$  and  $2x$ .



*Figure 2.7 : Hypothetical performance of (A) mild, (B) moderate and (C) severe neglect patients over two line lengths*

Other studies of normal bisection performance have concentrated on the effects of age, gender and scanning direction found in oriental languages. Fujii et al. [78] assessed patients across the entire adult age range (21 to 82) and have found that performance deteriorates in the oldest age group (61 to 82). As this grouping encapsulates the population of the test subjects to be included in the current neglect trial this deterioration can be ignored. The study showed that even though there was an overall performance deficit compared to two younger age groupings, the deviation of the results within the old group was small, indicating that age will not affect the validity of a control population of geriatric subjects. Roig [71] examined the effect of gender on bisection performance within a young population (16 to 42 years) and found no significant variation between performance.

Several studies have examined different assessment techniques for the administration of the line bisection task. Halligan and Marshall [67] implemented a computer-based system using a mouse to control an on-screen cursor indicating the midpoint. Hjaltason [79] used a head-mounted pointing device to locate the centre of the line. These studies show that although similar rightward deviation are noted for neglect subjects, the effects are not so pronounced as with traditional pencil and paper administration.

### 2.3.4 Test Sensitivity and Neglect Test Batteries

The tests of neglect defined in the previous three sections are usually administered to subjects as a battery of subtasks. Batteries such as the Rivermead Behavioural Inattention Test [80] are widely used as a diagnostic tool (Chapter 3 provides a full exploration of this test battery). Other test batteries that have been recently devised specifically for the detection of neglect include a modified Milner Landmark task (physical and verbal location of objects) [81], a battery of identification and location based tasks designed specifically for geriatric patients requiring coarser motor control [82] and a series of reading based tests [83].

CVAs in particular areas of the brain are known to cause varying performance characteristics across the range of tasks and therefore neglect severity assessed by a single task may not provide an accurate analysis of the extent of the neglect in a subject. Marshall et al. [84] documented the results from a range of tasks for three neglect subjects and identified task specific performance in relation to the location of the subject's lesion (Table 2.1)

Patient Number	Lesion Location	Deficit
1	Posterior Parietal	Poor line bisection and drawing
2	Temporoparietal/ Occipital	Poor line bisection response
3	Anterior / Subcortical	Good bisection, poor drawings

*Table 2.1 : Lesion location and task specific deficit*

Other studies [85] have examined the drawings produced from a series of figure copying tasks and have concluded that a CVA in the parietal region of the brain causes the most drawing dysfunction.

In an attempt to assess the diagnostic ability of the subtasks, Marshall and Halligan [86] presented the performance results of a single case study tested with a variety of tests of neglect. They found that the subject performed very badly on a line bisection and a cancellation task, whereas the drawings made for a geometric copying task contained only a few errors; the test subject could accurately copy most of the left hand side components of the shapes. From these results they concluded that the bisection and cancellation tasks were able

to identify less severe cases of neglect and were therefore more sensitive to detection of the condition. It should be noted that these findings applied to the generalised testing of neglect are hypothetical. As only a single case study patient was used in the trial, other performance variables such as lesion location and post-stroke testing time could affect performance on any of the sub-tasks.

Within the Rivermead BIT, Halligan and Robertson [14] established the ability of each of the tasks to detect neglect within 30 patients who had been diagnosed with the condition by clinical assessment and functional observations. Table 2.2 details their findings in order of task effectiveness. From this study, the cancellation tasks clearly provide the best test for the detection of neglect.

An important finding for the administration of tests of neglect (and other clinical based tests) within a geriatric population was presented by Casagrande et al. [87], who found that the performance on a task was sensitive to the time of day, test subject energy levels and the amount of sleep obtained by the test subject prior to testing.

<b>Task</b>	<b>Number of Neglect Patients Detected</b>	<b>% of Neglect Patients Detected</b>
<b>Star Cancellation</b>	30	100
<b>Letter Cancellation</b>	24	80
<b>Figure Copying</b>	22	73
<b>Line Crossing (Albert's)</b>	17	57
<b>Line Bisection</b>	16	53
<b>Representational Drawing</b>	11	37

*Table 2.2 : Test sensitivities of Rivermead BIT subtasks*

## 2.4 Computer Based Neuropsychological Testing

The use of computer based systems within the field of neuropsychological testing has enabled the measurement of features with greater resolution and accuracy. Elithorn et al. [88] presented the general principles and practices of *automated testing* and outlined the advantages of using such systems. Computer implementation allows greater accuracy in timing measurement and enables the recording and storage of response data while test

parameters such as stimulus size, position and display times can be modified with ease. More importantly for use in large scale trials, procedural consistency between tests can be maintained and an objective outcome can be produced algorithmically which is not susceptible to fatigue or assessor experience.

The majority of existing computer-based tests use on-screen prompting and reaction time assessment [89] by some form of interface, usually keyboard or external button or trigger. A number of studies have implemented on-screen bisection tasks [90] [67] with the test subject moving a cursor left and right to mark the centre of the bisection line. The results from these tests have been found to correlate with the standard pencil and paper implementation, but with reduced sensitivity to neglect. Rehabilitation therapy strategies have also been implemented using on-screen prompting and scanning exercises where a patient has to discriminate between different stimuli placed at varying positions within the visual field [91]. To date, the effectiveness of these rehabilitation strategies has been very limited. Lincoln [92] concluded that there was currently no effective computer-based rehabilitation treatment to compensate for neglect, indicating the need for innovation in methods of stimulus and assessment [93]. The study by Bergego et al. [94] supported these findings, noting no reduction in neglect severity following a scheme of computer-based recreation and rehabilitative on-screen scanning tests. Studies that do claim to improve performance are almost universally assessed on a single case study, highlighting the need for increased population trials [95].

While a computer based implementation has many advantages in accuracy and consistency, the use of technology can modify test subject performance. This is particularly prevalent within a geriatric community where apprehension towards using unfamiliar technology is increased [96]. Tseng et al. [97] concluded that a quarter to a third of the population are anxious about using computers and that test performance is affected proportionally to anxiety levels. The main cause of unfamiliarity is the communication interface between the human and the computer, in particular the use of a standard keyboard and inadequate instructions for use. Test subjects who are not familiar with a keyboard and general computer use worry about 'breaking' the computer or causing an unexpected response. This effect is compounded by the situation when on-screen or audible feedback of user action is delayed or not evident, causing multiple responses by the test subject [98].

Roberts et al. [99] and Collinson et al. [100] both surveyed input devices for handwritten responses to neuropsychological tests. Both studies found that a traditional pencil and paper infrastructure provided the best method for testing, preventing the introduction of other test

variables such as equipment unfamiliarity. However, a pen based graphics tablet maintained the test construct equivalence to pencil and paper tasks while enabling computer based analysis. As direct contact with the technology (such as keyboards, mice, cursor keys and touch screen) is abstracted, the test subject did not feel that the computer was imposing on his test performance or that he had to attain a level of computer literacy to perform well in the test. Other studies [101] [102] have supported these findings by examining pen-based technology for clinical use focusing on the speed of data entry and response times for a range of input peripherals. For the inexperienced user, the pen-based interface provided the fastest and most efficient method of computer communication.

The types of data that can be extracted from the use of pen-based capture systems have been widely explored within the fields of biometrics, signature verification and handwriting recognition [103][104]. Data capture and analysis methods can be classified into two groupings:

- **On-line analysis** examines the position and other pen status data in real-time or as a stream of coordinates stored in a file. This enables the extraction of both static data pertaining to the measurement of drawn images [105] and also dynamic time-based, movement and constructional data [106]. On line analysis is the obvious assessment method for data capture using a graphics tablet as additional dynamic data can be explored.
- **Off-line analysis** uses image processing techniques to analyse the drawn (static) image. Attempts have been made within the fields of forensic analysis of documents to obtain dynamic data from these static images [107], extracting features such as direction of stroke, pen movement velocity and pressure. Some of the techniques developed in this field will increase the measurement accuracy of drawn images from neuropsychological tests such as spatial and angular assessments of drawings.

With reference to neuropsychological testing, the types of features that can be extracted and the additional diagnostic information than can be obtained from pen-based drawing studies are now examined.

### 2.4.1 Static Features

The static features extracted from drawings or other handwritten responses are based on the accurate measurement of *outcome*. Static features such as length of drawing stroke, drawing location from an origin, angle between two components in a geometric shape, distance from target, number of formed corners, error distance in formed corners, drawing area and level of perseveration (number of times a single side has been drawn) have been used effectively in diagnostic studies [108][45]. Task specific features (such as the number of cancellations/drawing components) defined as the standard marking criteria for the detection of neglect [33][34] (Section 2.3) can be extracted using static analysis. Any measurement that can be extracted from the final drawing response - the completed test overlay that would traditionally be assessed by a therapist – is classified as a static feature.

The major advantages of computer based static assessment are accuracy and consistency. An algorithmic approach can be applied to the assessment of drawings which enables the application of identical marking criteria across the entire test population. Applying such a marking scheme is one of the inherent problems in the assessment of drawings (Section 2.3.1), so defining a set of rules based, for example, on the number of sides or components drawn or accuracy in corner formation provides a clear and consistent marking criterion and system of assessment.

Use of an algorithmic approach to task assessment removes any ambiguity in the interpretation of marking criteria. For example, failure by an assessor to mark a particular cancellation target will affect the overall score for the task. A computer based approach will sequentially visit each target enabling accurate inspection of the response.

As shown in Chapter 3, the assessment of drawing tasks varies because of a subjective interpretation of the marking guidelines described in the BIT manual. Using a standardised static analysis of the drawn image removes any assessment variation between test subjects.

### 2.4.2 Dynamic Features

The dynamic features extracted from the test responses are derived from the sequencing and time-stamping of pen coordinates (and other data such as pressure and tilt) returned by the

capture device. With reference to the detection of neglect, the research documented in this thesis attempts to establish the diagnostic properties of these dynamic features. In particular, it is interesting to ask, for example, whether spatial differences that are evident within the static parameters, such as drawing or cancelling to one side of the page, are replicated by dynamic features. Marquardt and Mai [109] provide a background to the extraction of dynamic features from pen-based applications including guidelines to the necessary sampling frequencies and resolutions required to avoid loss of data and techniques for eradicating errors caused by sampling noise.

The velocity (amount of movement per second) of the pen has been investigated over a series of tasks and neuropsychological conditions and has proved to reveal interesting diagnostic differences between test groupings. Using a series of straight line drawings, Plamondon [47] defined a model describing rapid-aimed movements made by subjects which can be linked to physiological impairments. Deviations from the normal bell-shaped velocity profile indicate motor and cognition problems within a test subject. A velocity skew measure calculated from the ratio between the time to reach peak velocity (acceleration time) and time after peak velocity (deceleration time) enables profile differences to be quantified. MacKenzie et al. [110] explored the target specific nature of movement-based tasks with reference to Fitt's Law of movement and the normal velocity profile. Across all test groups, the time to peak velocity within the profile increased proportionally to both the amplitude of movement (distance between targets) and to the target size, leading to the definition of a power function linking time after peak velocity, distance between movement targets and target size.

Teulings et al. [111] used the bell-shaped profile to note differences within Parkinson's patients in overall stroke size, drawing duration, peak velocity and time to peak velocity. Other studies [112] have examined the velocity profile at angles within geometric shape drawings constructed by a normal population which identified a higher pause time (no movement) at obtuse angles.

The work of Mattingley et al. [2] indicates that velocity profiling can be used within a population of neglect patients to identify velocity changes within spatial areas of the visual field (Section 2.3.1). Examination of normal performance for velocity profiling within a healthy geriatric population [113][114][115] reveals a deterioration in movement efficiency (a skew from the normal bell-shaped curve), with increased hesitation and sub-movements within the drawing. This deterioration increases the difficulty in differentiating between a healthy geriatric subject and a dysfunctional patient especially if the dysfunction is slight.

These results explain why Mattingley et. al. only found significant differences in a severe neglect group. Other studies have examined the velocity profiles in a range of dysfunctions: Tourette's Syndrome [116], Schizophrenia [117], Alzheimer's Disease [118] and Huntington's Disease [119].

Apart from the examination of velocity profiles, movement disorders such as tremor [120] have been investigated within dysfunctions with motor-based symptoms (for example Parkinson's Disease). Whilst tremor may be of interest to a neglect-based study, the main effects for examination are the spatial differences throughout the overlay. Investigation of dynamic features such as cancellation sequencing, starting position, timing regression and quadrant analysis of movements will indicate if dynamic features replicate the visual static differences within the drawing. Differences in movement times towards targets located to the left of the visual field [121] have indicated that a detailed timing analysis on a side or quadrant basis provides significant performance differences between neglect and other subject groups. Extracting time, movement and sequence based dynamic features for task specific analysis may enable detection of differences which cannot be identified by static features and hence increase the sensitivity of the test battery.

## 2.5 Classification Techniques

In the previous section, the possible types of data that can be extracted from a series of pen-based visuo-spatial tests have been investigated. While, ideally, every performance-based feature can be used to classify a patient, in reality each feature has a different classification ability. Analysing the interaction between two or more features may result in patterns of classification (or clusters) forming which can separate test subject groupings. This section presents an overview of a series of techniques that can be used to assess the ability of individual features for data classification and how combinations of features can be used to automatically classify the responses to indicate a patient grouping.

Classification architectures can be separated into two categories: *supervised* and *unsupervised*. The distinct property of a supervised classifier is that it requires a *training* phase involving the attribution of classifier behaviour to a class of input vector. Training the classifier involves providing the required system output value (or class) with an input feature vector and establishing the classifier performance characteristics common to all input vectors of a particular class. In some architectures (such as neural networks) this involves the

modification of internal weights to *reinforce* correct classifier performance. In other systems such as cluster analysis, training involves the identification of cluster centres which describe the mean position in  $n$ -dimensional space occupied by the features from a particular class. Upon completion of this first phase using a series of training data, the system can be used to classify using an input feature set.

An unsupervised classifier is not provided with a required output, but forms classifications based on similarities between input feature vectors. As such, the training phase of an unsupervised architecture is the process of the classifier forming an internal structure based on the provided input vectors. With a vector presented to the system, similarities with other vectors are analysed. The area of the network, or cluster, best representing the input feature vector can be identified and reinforced. These areas can then be labelled to indicate the particular classifications. Following training and classifier initialisation, the network area which becomes the most active or the cluster nearest to the position in  $n$ -dimensional space formed by an input feature vector indicates a classification.

There are several advantages to using unsupervised architectures, most importantly the ‘automatic’ nature of the result generation enables the abstraction of the data from any biased or miscalculated grouping. The classification patterns and self-organisation of the system can also be studied, which may indicate interactions and groupings that are not immediately apparent through direct inspection of data.

The performance of any classifier can be established by the error rate (Equation 2.3). This signifies the number of misclassifications by the trained system. An ideal classifier will have an error rate of 0%.

$$\text{Error Rate} = \left( \frac{\text{Number of incorrect classifications}}{\text{Number of classifications}} \right) \times 100\% \quad (2.3)$$

### 2.5.1 Principal Component Analysis

Principal Component Analysis (PCA) does not produce a classification of data, but it can be used to pre-process input feature vectors. PCA examines the correlation between independent

input features to establish any clustering or groupings within the provided data. PCA can be used to reduce the dimensionality of a feature vector by representing highly correlated variables as a single feature. Pre-processing the input feature reduces the amount of data presented to the classifier, generally improving speed performance. Dominant clusterings formed by highly correlated features within the input vector may saturate a classifier and prevent lesser correlated features contributing to the final classification calculation. Algorithms for performing PCA calculations are described widely within the statistical literature [122][123].

Recent data classification studies which have used PCA to assess data with a large dimensionality include handwritten digit recognition [124] and vision-based target classification for military purposes [125].

### 2.5.2 Bayesian Statistical Classification

The Bayesian classifier is a statistical approach to pattern classification. The classifier uses frequency distributions in calculating the probability that an input vector belongs to a particular class; the highest probability indicating class membership [8]. Where the frequency distributions of a particular class are unknown then the classifier is trained (supervised mode of classification) by obtaining a model of each class membership. The training set should be statistically representative of the entire range of class members to ensure classifier accuracy.

The Bayesian decision rule is shown as Equation 2.4. The probability of a vector  $D$  being assigned to a class  $G_i$  (of  $g$  classes) is defined as:

$$P(G_i|D) = \frac{P(D|G_i)P(G_i)}{\sum_{i=1}^g P(D|G_i)P(G_i)} \quad (2.4)$$

where :

$P(G_i|D)$  is the probability that  $D$  belongs to class  $G_i$ .

$P(G_i)$  is the probability of a case belonging to group  $G_i$  when no information about the case is available. This probability can be estimated from the observed proportions of cases in each group from the training data set.

$P(D|G_i)$  is the probability of obtaining vector  $D$  given class  $G_i$ . This determines the probability distribution that a class  $G_i$  yields vector  $D$ . In practical terms,  $P(D|G_i)$  can be calculated assuming a normal distribution from the training set. A method for calculating the probability from a set of training data can be found in Fairhurst [126].

Membership of vector  $D$  to group  $G_i$  is can be defined as:

$$D \in G_i \text{ iff } P(D|G_i)P(G_i) > P(D|G_j)P(G_j) \quad \forall i \neq j \quad (2.5)$$

### 2.5.3 Cluster Analysis

Cluster analysis is an unsupervised classification method which groups objects according to the similarity between feature vectors. The technique is widely used in the fields of biological and medical sciences where many data are collected from a particular patient and an attempt is made to classify a condition by grouping *similar observations* [127], [128] [129]. Figure 2.8 shows objects represented by two features in a two dimensional feature space. The object feature vectors have formed three clusters each with separate densities (represented by the circles by each grouping). Objects are classified by finding the nearest cluster centre to the input vector position within the feature space. The nearest cluster contains objects which are of greatest similarity as defined by the features under investigation. The object represented by the hexagon is closest to Cluster B using (in this case) a Euclidean distance measure and therefore can be classified as belonging to the group of objects represented by this cluster.

Objects can be classified in  $n$  dimensions where  $n$  is the size of the feature vector describing objects, although subsets of the vectors can also be analysed possibly following pre-processing by techniques such as PCA (Section 2.5.1)

The most popular method for classification using cluster analysis is the MacQueen K-means algorithm [130]. In this algorithm the number of clusters within the feature space is predefined, the inflexibility of which has led to much development and modification to the

basic algorithm. Having defined the number of clusters, *exemplar objects* from each of the classification groups are mapped into the feature space. These exemplars initially form the centre of each of the clusters from which distances will be measured. The distances between initial selection of exemplars therefore affects the performance of the classifier.

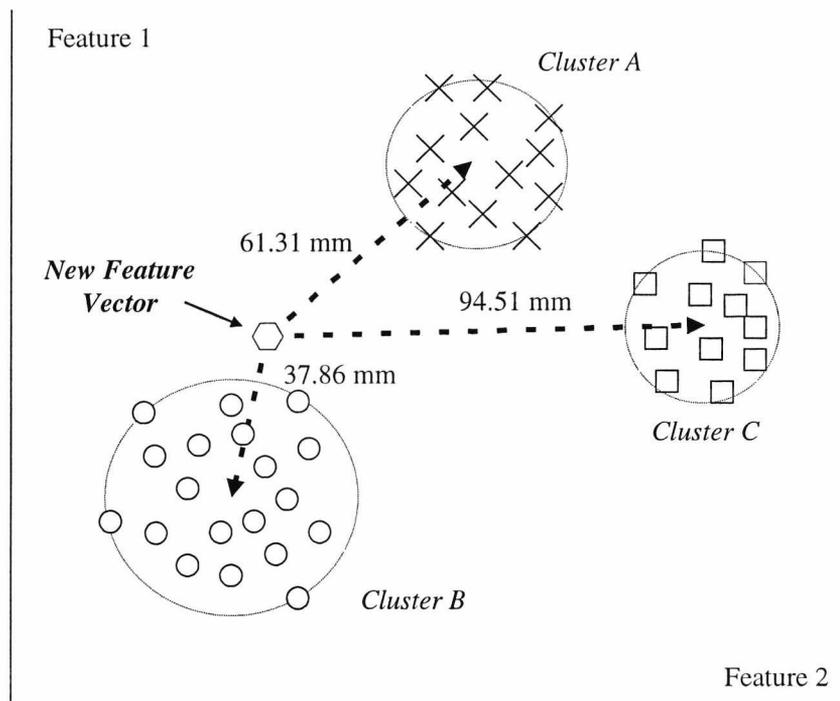


Figure 2.8 : Two dimensional Euclidean distance cluster analysis

Classification of objects is performed in two stages.

- Stage 1 – Determination of Cluster Centres : Objects (apart from those used as exemplars) are mapped individually into the feature space and the nearest cluster to the object is identified. After adding the object to the cluster, the new centre is computed. The resultant cluster centres after mapping all objects are then static for Stage 2 of the classification.
- Stage 2 – Classification : Using the cluster centres defined in Stage 1, classification is performed individually on objects by finding the nearest cluster centre using a selected distance metric.

Many improvements to this basic algorithm have been proposed [131][132]. Adaptive cluster creation introduces a new cluster centre if the distances between an object  $X$  introduced to the feature space and *all existing* cluster centres are greater than a predefined limit. The centre of the new cluster uses the mapped position of object  $X$  which is utilised in further distance calculations. Other strategies have included repeating the first stage of the K-means algorithm until a defined convergence threshold has been reached and using a learning rate factor within the cluster centre update calculations in Stage 1. This factor is decreased as the number of objects presented to the classifier is increased. This results in the cluster centres being able to adapt more to the initial objects when the classifier structure remains undefined. As the stability of the cluster centres increases so the amount of modification is restricted.

#### 2.5.4 Kohonen Self-Organising Map

A Kohonen Self-Organising Map (SOM) [133] is an unsupervised feed-forward learning neural classifier used widely for classification investigation [134][135][136]. Using the perceptron processing element [137] classification is similar to cluster analysis in that objects

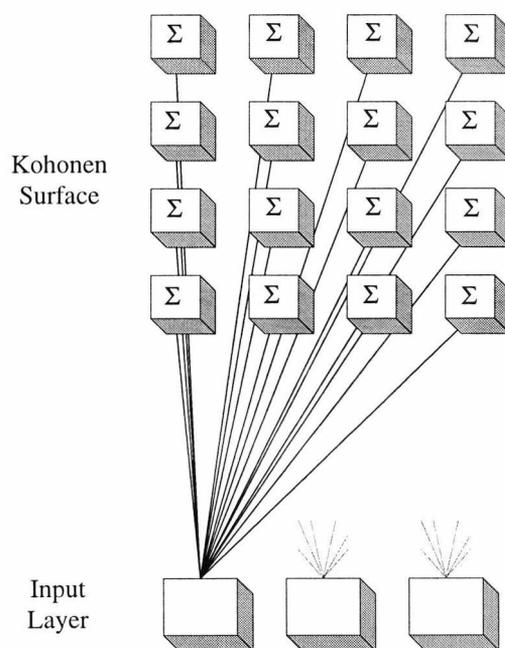


Figure 2.9 : 4 x 4 Kohonen map with 3 input and output nodes

with similar feature vectors are grouped together within the Kohonen surface. Figure 2.9 shows the configuration of a SOM. Feature vectors are introduced to the network at the input nodes, the number of which match the number of input values. These nodes are fully connected to the processing elements in the Kohonen surface (i.e. a connection exists between every input node and surface node). Similar objects are stored in a topologically correct position within the structure of the surface and thus nearby points within a surface also have similar feature attributes. Such topological mappings occur in many physiological processes, such as the mapping between the auditory cortex and the ear. As such, the self-organisation of the network given a set of objects is of particular interest to neuropsychologists, since these mechanisms may hold some clues about how neural systems in the human are organised and function.

Training of the SOM involves presenting the image to the network input nodes which is then propagated to all nodes of the Kohonen surface. The surface processing element with the highest output value is selected to represent the input object and is reinforced, along with surrounding neighbourhood nodes. Many models exist for neighbourhood reinforcement but to extend the similarities with a human neural system, a Mexican hat or Gaussian function models localised brain cell activation. To classify input objects with the trained SOM, the node with the highest activation on the surface indicates the classification.

Several problems exist with the basic SOM model. Uncertainties arise when defining the network topology in deciding how many processing elements are required to accurately classify data. Optimum performance occurs when the number of processing elements equals the known number of categories, which relies on having a priori knowledge of data segregation. Methods, such as K-means cluster pre-processing and thresholding of the SOM processing surface for estimating the number of output classes within the data when this is unknown prior to classification have been discussed in a number of papers [138][139].

### 2.5.5 Adaptive Resonance Theory

Since the Adaptive Resonance Theory (ART) classification architecture was first described by Grossberg in 1976 [140] there have been many modifications and variations to the basic system structure. The ART was developed to model the brain's ability to store and generalise the classification of objects, the main objective of which was to enable the formation of self-organising stable clustering of data. The main advantage of the architecture is its ability to

modify the internal weightings (train) and classify objects in real time, removing the distinct training and recognition phases of traditional neural architectures. The original ART architecture was devised to categorise binary input patterns. ART2 [141] provided an extension to the basic system to allow analogue values to be used as input features. Other extensions have allowed faster response (ART-2A) [142], implementation of fuzzy logic set theory for assessment of analogue patterns (Fuzzy ART) [143] and supervised learning of object data (ARTMAP) [144].

A detailed description of the functioning of the ART classifier can be found in many neural network references [145][146]. The basic functioning of the system (Figure 2.10) involves the use of two layers of perceptron elements: Input and Comparison (F1) and Output and Recognition (F2). An input vector is presented to the F1 layer which causes the selection of a single representative node within the F2 layer. By propagating a prototype vector, indicating the pattern represented by the F2 ‘winning’ node back to F1, the difference between the input vector and the prototype vector can be assessed. Output classification occurs if the pattern is above the threshold set by the vigilance parameter.

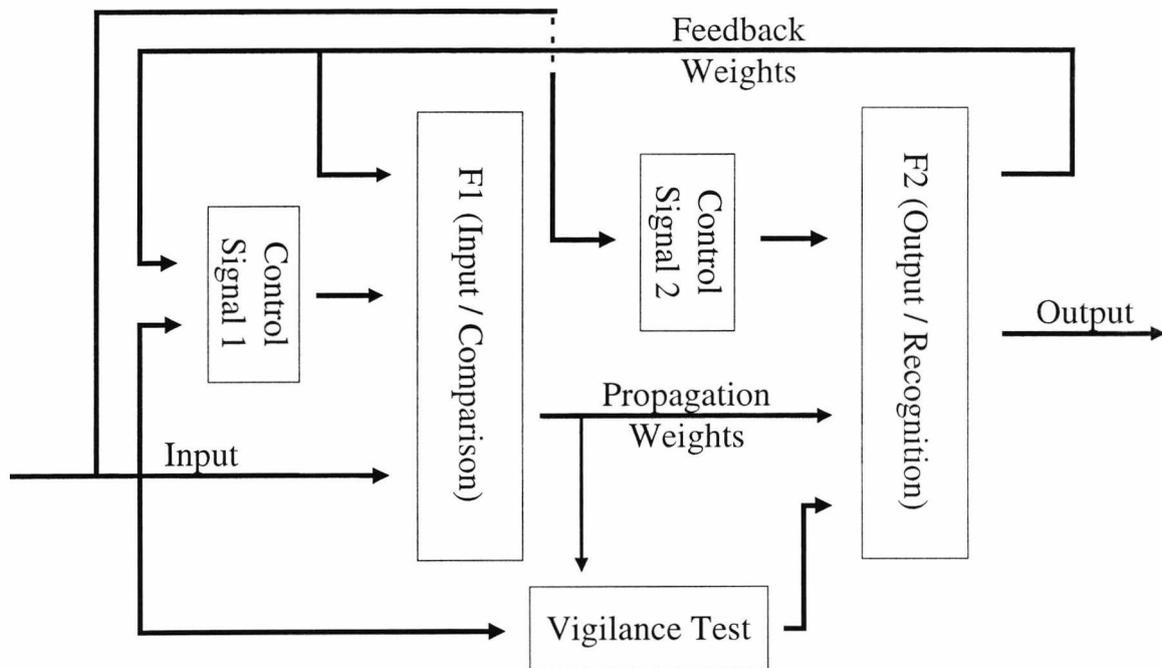


Figure 2.10 : ART data flow schematic

Recent examples of classification problems investigated using an ART architecture include the diagnosis of chronic inflammatory bowel disease [147] and the classification of base oils using their infrared spectrum [148].

## 2.6 Design Objectives

Having reviewed the background to the study and investigated the possible areas for development and improvement over conventional testing systems utilising a computer-based response capture system, a set of more specific objectives for the study can be defined. These objectives draw on the neuropsychological and clinical basis of pencil and paper tests for neglect combined with the advantages of accuracy, consistency and the range of dynamic constructional properties that a computer-based implementation delivers.

Hence, the principal design objectives are as follows:

1. To establish the reliability of the existing testing batteries and time overheads to administer and assess an individual patient.
2. To implement a battery of pencil and paper tasks to accurately assess visuo-spatial neglect. Pen movements and timings are captured using a computer-based system.
3. To collect a series of data with the computer-based system from patients with visuo-spatial neglect, stroke patients without neglect and age matched control subjects.
4. By applying an algorithmic approach to feature extraction, improve the reliability, resolution and consistency of static-based assessment.
5. To extract a series of dynamic time and constructional-based features and establish performance levels within test groups.
6. To assess whether the dynamic features replicate the static spatial differences between patient groups.
7. To enable the use of the system within a clinical environment both in terms of software interface and hardware design.

8. To establish which features extracted from the test battery provide the best discrimination between test subject groups.
9. To evaluate the ability of a series of pattern recognition and classification architectures to diagnose a test battery response.

## 2.7 Summary

This chapter has introduced and investigated the clinical condition of visuo-spatial neglect. The manifestation of neglect within stroke patients reveals areas which can be exploited through conventional pencil and paper tasks. In particular the need for accurate definition and standardised assessment of neglect is highlighted by the wide variation in reported incidence. The effects on everyday living show that thorough diagnosis of the condition is important for choice of rehabilitation programme and support within the hospital and once the patient has been discharged.

Several classes of task to detect neglect have been described and their relative diagnostic ability discussed. While the cancellation and bisection based tasks are more sensitive to detection of neglect, the drawing tasks contain dynamic or constructional data which can be used to identify neglect severity. Dynamic features of other tasks suitable for computer implementation have also been explored. Currently the drawing-based tasks suffer from non-standardisation of assessment criteria and prove too difficult for patients with severe cases of neglect. The variation in performance across a range of tasks, indicating the need for assessment using a test battery, has been highlighted.

The use of computers for neuropsychological assessment has been investigated with respect to interface peripherals and modification of test performance. Assessment using a graphics tablet has been identified as introducing the least disadvantage in use. Clear user instructions and an abstraction of the patient from direct contact with the computer is desirable as it reduces anxiety in use, which itself affects test performance. The possible types of static and dynamic features that can be extracted using a graphics tablet have been identified by examining the current literature in neuropsychological testing and assessing how the symptoms of neglect can be exploited.

Four types of classification methodology have been described, all of which can be used to analyse the feature vector of a patient's performance across a battery of tests. By using an unsupervised classifier, observations can be made about how clusterings are automatically formed. Relative performance in terms of classifier size, learning rates and feature vector size need to be investigated with trial data extracted from the responses from a set of patients.

Finally, in this chapter, the design objectives for the study have been defined on the basis of the subject review.

## Chapter 3

# Reliability of the Rivermead Behavioural Inattention Test

### 3.1 Introduction

To support the rationale of implementing a computer-based neglect test, the advantages and disadvantages of the existing conventional assessment method need to be established. One of the aims of automating the assessment process is to remove the subjectivity in the assessment of patient test responses. In this chapter the objectivity of current testing methods will be investigated by studying the extent of the agreement correlation between trained assessors.

A review of the development and use of the current neglect testing standard, the Rivermead Behavioural Inattention Test (BIT) is presented along with individual assessment techniques for the subtasks which comprise the conventional test battery. An interrater trial methodology is discussed along with appropriate assessment statistics. The agreement results of the BIT conventional assessment trial are presented followed by a discussion of the level of agreement between assessors on particular subtasks. It will be shown that agreement is acceptable, in terms of the Kappa statistic agreement criteria, for the subtasks with basic marking criteria (such as the number of cancellations on an overlay). However, in subtasks where objective judgement is required (such as the drawing tasks) the rater agreement is low, particularly where interpretation of the marking scheme supplied with the BIT is ambiguous. The implications for accurate assessment of neglect using the existing methods are also discussed.

### 3.2 The Behavioural Inattention Test

The Behavioural Inattention Test [50] was developed in 1987 specifically as a test of unilateral visual neglect. Intended for use within the fields of clinical evaluation and Occupational Therapy, a clearer understanding and common interpretation of the level of neglect within a patient is enabled through test standardisation across a range of tasks and

scoring methodologies. Combined with a series of behavioural tasks, the BIT enables a therapist to diagnose and monitor the effects of neglect on everyday activities which in turn can be used to select a course of rehabilitation relevant to the patient's condition. The use of the BIT for assessment of neglect is widespread [149] and is an accepted standard within the medical profession.

The test consists of two strands of subtasks:

- A *conventional battery* comprising a series of traditional pencil and paper based neuropsychological tests such as the Albert's cancellation task [49] and line bisection. Assessment is made on items such as correct number of cancellations and quality of drawings (Section 3.2.1)
- A *behavioural battery* comprising assessment of everyday activities such as reading, telling the time and telephone dialling. Assessment is made on items such as the dialling sequence in the telephone task, the number of items read on a menu and the correct reading and setting of the time.

In analysing the agreement between assessors scoring a common set of responses from the BIT, only the conventional battery will be used. Whereas the responses from the conventional battery can be distributed to assessors on sheets of completed overlays, assessment of the behavioural subtasks requires the use of techniques such as video recordings or arranging for all assessors to observe the same testing session. The ability of a remote assessor to observe specific marking items from the video is uncertain. The computer-based testing system is designed only to implement pencil and paper tasks with the same testing methodology as the conventional battery. The interrater study will therefore only concentrate on these subtasks.

### 3.2.1 Conventional Subtasks of BIT

The conventional battery of the BIT consists of six subtasks: three cancellation tasks, a line bisection task and two drawing tasks. The following is a brief summary of individual tasks and scoring methods. All methods and drawing examples are contained within the BIT Reference Manual [80]. This manual is the only documentation provided to assist with marking of the BIT.

### 3.2.1.1 Line Cancellation

Using a standard Albert's cancellation task [49] test subjects are presented with an overlay containing 40 lines positioned in a pseudo-random arrangement. Subjects are required to locate and 'cancel' (place a single pen stroke through) all of the lines. The overlay used is shown as Figure 2.5 in Chapter 2. Assessment of this task involves counting the correct number of line cancellations made on the overlay ignoring the central vertical column of 4 lines which are used to demonstrate the cancellation process to the test subject. This results in a maximum score of 36 for the overlay. To aid the assessor in marking the task, a 'mask' overlay highlighting the targets using transparent areas is placed over the test subject's response.

### 3.2.1.2 Letter Cancellation

Figure 3.1 shows the overlay used for the letter cancellation task. The test subject is required to locate all of the 'E' and 'R' characters amongst distractor characters in the 34 by 5 grid. 40 correct targets are printed on the overlay, 20 to the left of the vertical centre and 20 to the right. Again, the assessor uses a marking mask to aid the scoring.

### 3.2.1.3 Star Cancellation

The final cancellation task is shown in Figure 3.2. The test subject has to locate 56 small stars randomly positioned amongst larger stars and distractor characters. In assessing the overlay, the two stars directly above the central arrow are not counted as they are used to demonstrate to the test subject the cancellation technique required. This gives a maximum score of 54. A further marking mask is used for this overlay. This is particularly useful for this subtask as the targets are not arranged linearly in columns or rows.

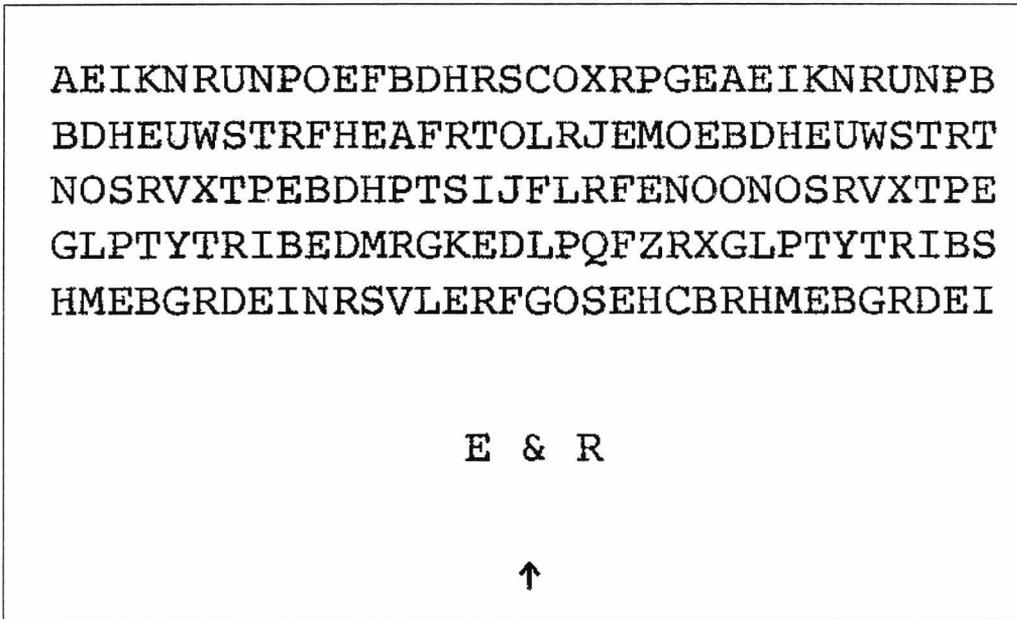


Figure 3.1 : Letter cancellation task overlay

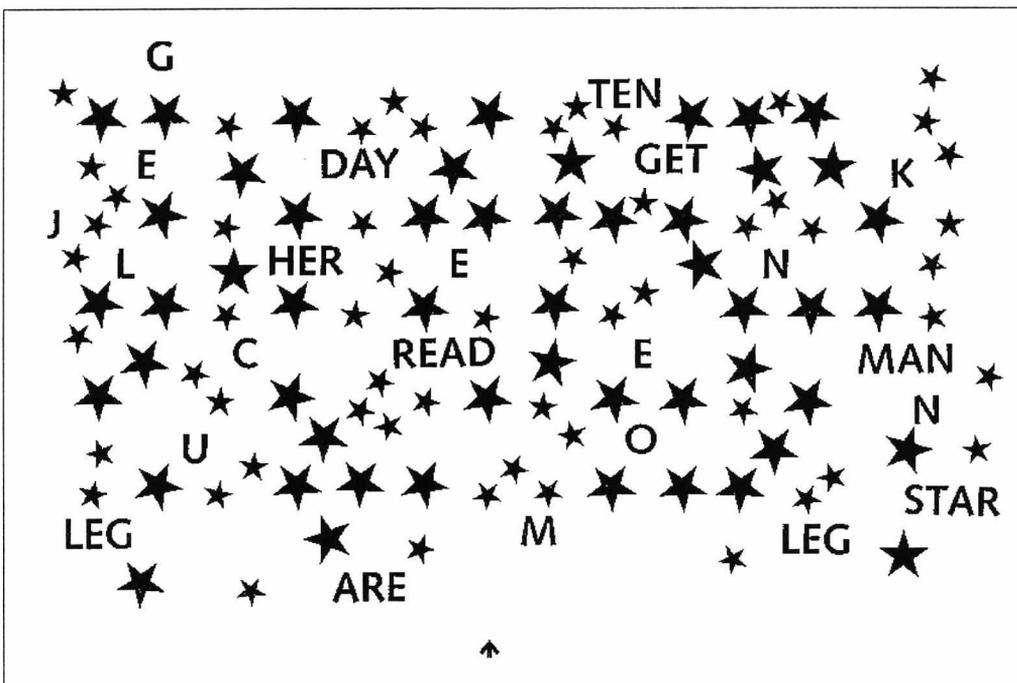
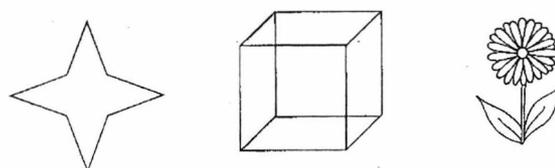


Figure 3.2 : Star cancellation task overlay

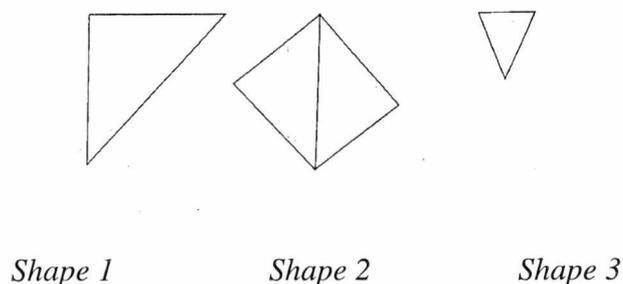
### 3.2.1.4 Figure and Shape Copying

Two overlays are used in the figure and shape copying task. The test subject is first presented with an overlay containing three simple drawings: a star, a cube and a flower (Figure 3.3). The model shapes are located on the left of the overlay and the test subject is required to copy the shapes directly to the right. This arrangement is also used on the second overlay which requires three simple geometric shapes to be copied (Figure 3.4).

A single assessment score in the range of 0 to 4 is awarded across all six drawings. This score reflects the ‘completeness’ or presence of major components within a drawing [80]. An assessor is provided with a single or pair of reference examples for each of the drawing or shapes. However, as example scores are not given with the drawings, assessors have to use individual judgement and experience to mark each response on a component level. This results in high subjectivity in assessment.



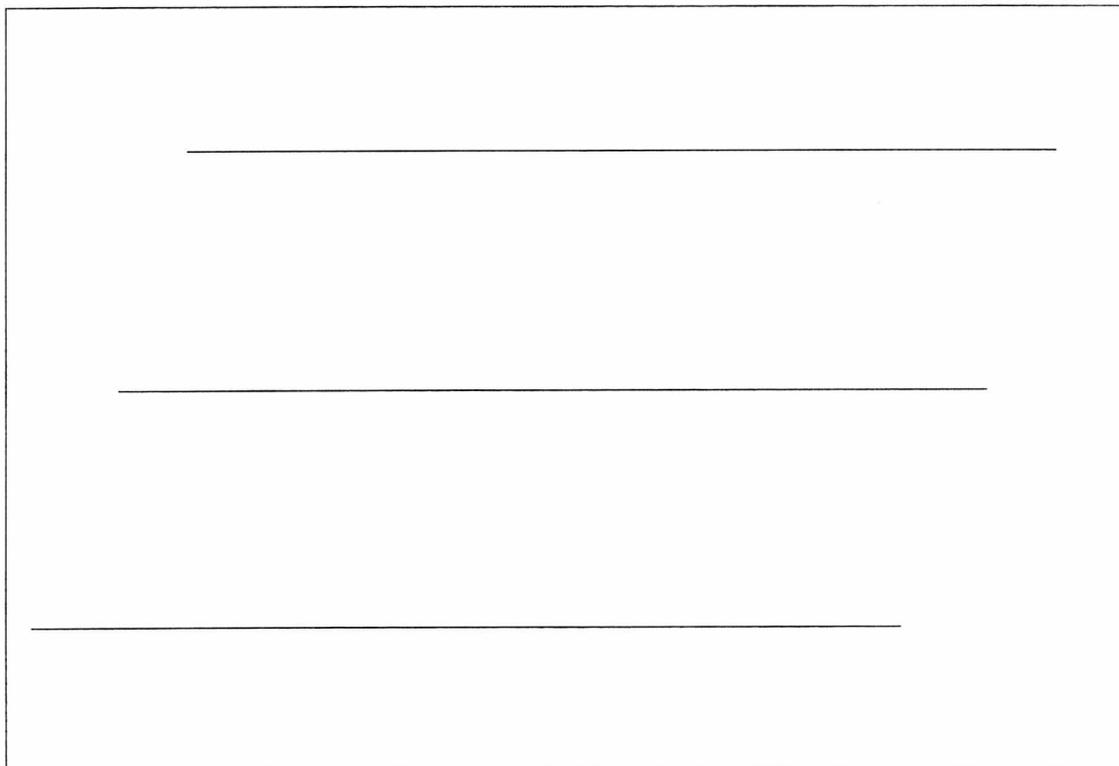
*Figure 3.3 : Figure copying models*



*Figure 3.4 : Shape copying models*

### 3.2.1.5 Line Bisection

This task consists of a single overlay on which are printed three horizontal lines of length 8 inches (176 mm). The test subject has to locate and bisect at the midpoint of each individual line (Figure 3.5). Each line is scored on the bisection deviation from the true midpoint. A marking mask contains the scoring distance limits from the centre of each line. A score of 3 is awarded if the bisection is within 0.5 inch (12 mm) of the true mid-point, 2 marks if within 0.75 inch (18 mm) and 1 mark if within 1.0 inch (22 mm). Separate marks are awarded for the three lines on the overlay, resulting in a maximum score of 9.



*Figure 3.5 : Line bisection overlay*

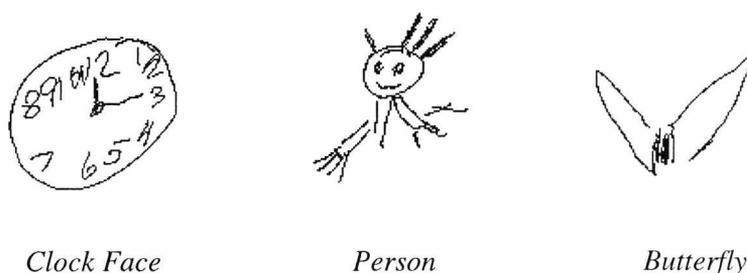
### 3.2.1.6 Representational Drawing

For the final task of the conventional battery, the test subject is required to draw a clock-face, a person and a butterfly, all without reference models, on a blank sheet of paper. Figure 3.6 shows some example responses from this task. As with the figure and shape copying task,

marks are awarded globally across all three drawings based on the presence of major components. Again this is very subjective as, although example drawings are provided, the associated scores are omitted. To demonstrate the subjectivity in drawing assessments, the following marking criteria (the *only* marking instructions provided for the drawing tasks) is taken directly from the BIT manual [80]:

*“The scoring of this subtest is based on the completeness of the respective drawing (0 to 3). Failure to complete is defined as the omission of any major component of the drawing”*

Without reference drawings for each score awarded, the assessment across the three drawings requires the individual interpretation from each assessor.

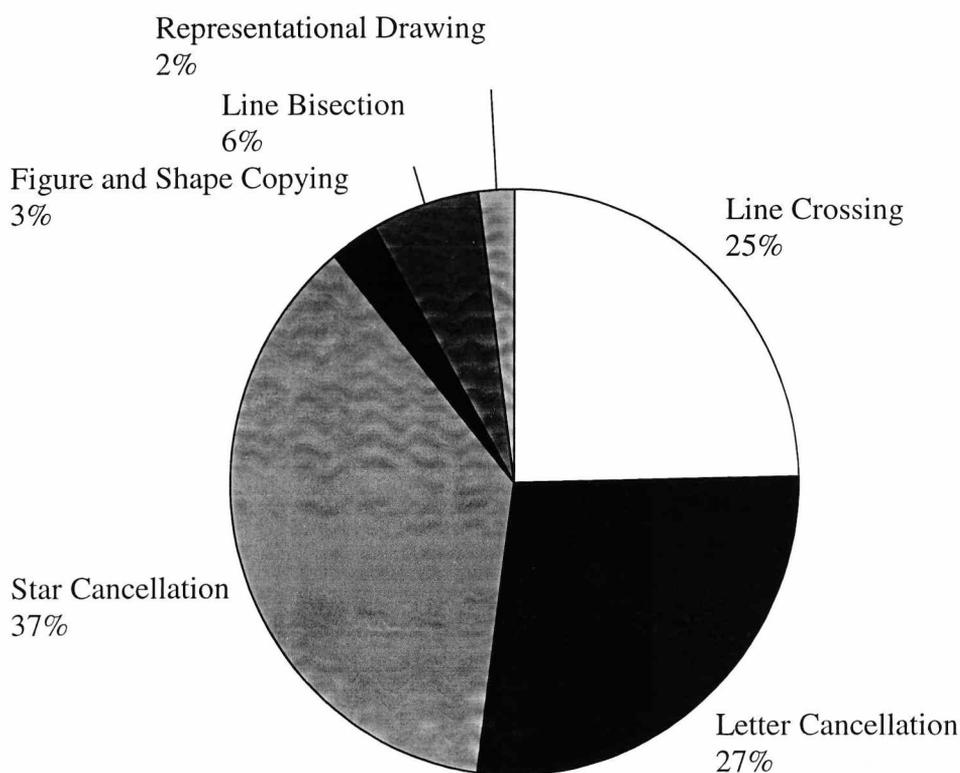


*Figure 3.6 : Figure and shape copying models and representational drawing responses*

### 3.2.2 Battery Score Interpretation

To obtain a total score for the conventional test battery, the marks awarded for each sub-task are summed producing a maximum score of 146. If a total score of 129 or below is awarded, then the test subject can be diagnosed as exhibiting a visuo-spatial deficit such as neglect. This neglect threshold was derived experimentally by Wilson, Cockburn and Halligan [80] from the results of 50 asymptomatic control test subjects; a score of 130 represented the lowest score obtained by this group and hence defined the threshold between neglect and control subject detection. The precise implications of a test subject’s deficit can be investigated further by the behavioural battery or other clinical investigations.

On examination of the marking scheme for the conventional battery it is apparent that there is a heavy bias towards the cancellation tasks (Figure 3.7). Indeed, as these three tasks account for 89% or 130 marks, it is possible to 'pass' the test without scoring on the drawing and bisection tasks. The BIT scoring scheme does recommend however that any sub-task score below a defined cut-off point (detailed in Table 3.1) should be investigated further, even if a total score of above 129 is obtained. This indicates the importance of each sub-task in the assessment of attention.



*Figure 3.7 : Distribution of marks with the BIT conventional battery*

Wilson, Cockburn and Halligan validated the BIT using two independent assessors to score 13 test subject responses [80]. The result of their trial revealed total agreement in scoring between the two assessors. It is unclear, however, if the agreement was based on the total battery score or individual sub-task scores. If the former was the basis for agreement assessment, then a variation in marks awarded may still have occurred within the subtask scores.

Hartman-Maeir and Katz [150] validated the behavioural sub-tasks by comparing the results from 40 test subjects against their activities of daily living (ADL) scores [151]. They found that seven of the nine subtasks significantly differentiated between neglect and non-neglect subjects and that six subtasks correlated with the finding of the ADL. Conventional subtasks were not included in the trial. Further studies [152] have provided additional validation of the behavioural subtests and the ADL. Cermak and Hausser [153] suggested areas on which the functional assessments of the BIT could be validated more thoroughly including effects of age, gender and education level. The reliability analysis of the behavioural subtests has not been extended.

<b>Subtask</b>	<b>Max Score</b>	<b>Cut-off Score</b>	<b>% of Marks</b>
Line Crossing	36	34	24.6
Letter Cancellation	40	32	27.4
Star Cancellation	54	51	36.9
Figure and Shape Copying	4	3	2.7
Line Bisection	9	7	6.2
Representational Drawing	3	2	2.0

*Table 3.1 : Conventional subtask marks and cut-off scores*

### 3.3 Interrater Methodology

This section describes the experimental trial to explore levels of agreement between BIT assessments. The objectives of the study are principally to:

- a) investigate agreement in the conventional battery total score awarded over a common set of test responses and establish implications for the diagnosis of individual test subjects.
- b) identify the conventional subtasks that produce the most disagreement between assessors and examine why these subtasks produce marking variance.
- c) identify the shapes within the drawing tasks produce the most disagreement between assessors and why particular responses result in varied interpretation of the component-based assessment criteria.

Eleven Occupational Therapist assessors from four separate test centres in East Kent participated in the study. Ten sets of completed overlays from the BIT conventional battery were presented independently to individual assessors who were asked to mark each set separately to the guidelines defined in the assessment manual. Each subtask score was recorded for all ten sets of responses. A battery score was obtained from a sum of these subtask scores. All BIT sets were from patients admitted to Nunnery Fields Hospital, Canterbury with a right sided cerebro-vascular accident (CVA). Of these ten test sets, five were completed by patients who had been identified as exhibiting visuo-spatial neglect from clinical examination (other than the BIT) by doctors and/or therapists.

Following the battery assessment using the conventional marking scheme defined by the BIT manual, assessors were required to rate each of the *individual drawings* made for the Figure and Shape Copying and Representation Drawing tasks using a scale of 0 (very poor) to 4 (excellent) as opposed to awarding a global subtask mark. This facilitated the investigation of individual shape assessment agreement between raters.

To obtain a clearer understanding of the assessment correlation in the drawing-based tasks, a further ten sets of completed Figure and Shape Copying and Representation Drawing overlays were presented to each assessor. These overlays were responses from a further ten RCVA test subjects, five of which exhibited neglect. Thus, a total of twenty subject's drawings were assessed. Again the drawings were assessed using the BIT defined guidelines and by individual shape. This enables an examination of marking variance, and hence assessment ambiguity, within each of the drawing shapes.

### 3.3.1 Assessment of Results

Several statistical methods exist to analyse the agreement between assessors' scores and hence the interrater reliability of the BIT. Many of the statistics (such as Cronbach's Alpha, Cohen's Kappa and Kendal's Coefficient of Correlation) compute similar or ranked equivalent results, but all involve finding the level of agreement between an assessment made by two or more raters over a single or range of subjects. Prior to presenting the results from this particular interrater study, measures for analysis are considered along with an interpretation of scores. For a general discussion on interrater measures see Bakeman [154] and Williamson [155].

The simplest measure of agreement is the *pairwise correlation* between raters (when more than two raters are used in a trial then a mean can be taken of individual pairwise correlations). This method, however, leads to errors when applied to the interrater agreement of continuous scoring data such as that awarded for the BIT assessment. Instead of measuring direct numerical or ranked agreement, correlation measures the relationship between assessors' results sets. For example, *Assessor A* may constantly mark a single subject 5 marks higher than *Assessor B* across all of the subtests resulting in a perfect correlation score of 1. The assessors' direct agreement (i.e. identical marks were awarded for an individual assessment), however, is 0.

### 3.3.1.1 Percentage of Agreement ( $P_0$ ) and Agreement by Chance ( $P_C$ )

The obvious solution for assessment of agreement is the direct comparison between data items (is *Assessor A's* mark for a single test subject the same as that awarded by *Assessor B's* ?), leading to a 'percentage of agreement' measure. For example, if 65 out of 80 assessments were identical then an percentage of agreement of 81.25 % is obtained. This method, however, suffers from judging bias if a large population of the test subjects belong to a particular category or categories. Consider, for example, the data contained in Table 3.2. This shows the agreement between two raters assessing 20 responses from the figure copying task of the BIT.

		<i>Assessor A marks awarded</i>					
		<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>	<b>0</b>	<b>Total</b>
<i>Assessor B – marks awarded</i>	<b>4</b>	8	-	-	-	-	8
	<b>3</b>	3	-	-	-	-	3
	<b>2</b>	1	1	1	-	-	3
	<b>1</b>	-	-	-	-	-	0
	<b>0</b>	-	-	-	-	6	6
	<b>Total</b>	12	1	1	0	6	20

Table 3.2 : Two assessor agreement in figure and shape copying task

As can be seen from the diagonal entry, the percentage of agreement is:

$$P_o = \frac{\text{Number of agreements}}{\text{Number of assessments}} = \frac{(8 + 1 + 6)}{20} = 75\% \quad (3.1)$$

This result is biased towards the two extremes of the marking scheme (4, a 'perfect' drawing and 0, a very poor drawing). To quantify this bias, the probability of the assessors awarding identical marks due to chance ( $P_c$ ) is calculated by the following formula :

$$P_c = \sum_{i=1}^n \left( \left( \frac{\text{number of } i \text{ awarded by assessor A}}{\text{number of assessments}} \right) \times \left( \frac{\text{number of } i \text{ awarded by assessor B}}{\text{number of assessments}} \right) \right) \quad (3.2)$$

where :

$n$  = maximum in range of marks awarded

$i$  = score under observation

Examining the single case of an award of 4 marks, the 'perfect' drawing mark, the probability of assessor A awarding a mark of 4 is estimated at  $(12 / 20) = 0.6$  and a probability of  $(8 / 20) = 0.4$  for assessor B. Operating independently this means that there is an overall probability of  $(0.4 \times 0.6) = 0.24$  for both assessors classifying with a mark of 4.

This table also hints at one of the problems associated with the existing marking scheme of the BIT. Given a range of marks that a set of drawings are to be assessed between (in the case of the figure copying task between 0-poor and 4-excellent), assessors tend to award marks at the extremes of the scale. One hypothesis of why this occurs is that without a reference drawing for any of the marks, assessors use the extremes of the scale to denote a simple pass/fail assessment. Indeed, examination of the marks awarded for the 220 drawing tasks assessed in this trial using the BIT guidelines showed that 80% of drawings were awarded either a minimum or maximum score.

### 3.3.1.2 Cohen's Kappa ( $\kappa$ )

Cohen's Kappa statistic [156] [157] was devised to overcome the chance related assessment problem described above, thus removing any errors in agreement due to the sample distribution. The Kappa calculation produces a result between 0.0 (no agreement) and 1.0 (total agreement).

Kappa is defined by :

$$\kappa = \frac{P_o - P_c}{1 - P_c} \quad (3.3)$$

where :

$P_o$  = the proportion of observed agreement.

$P_c$  = the proportion of agreement due to chance.

$P_c$  is calculated as described in Section 3.3.1.1. by summing the chance probabilities for each marks awarded by a pair of assessors. Applied to the data in Table 3.2:

$$\begin{aligned} P_c &= (12 / 20) \times (8 / 20) + (1 / 20) \times (3 / 20) + (1 / 20) \times (3 / 20) \\ &\quad + (6 / 20) \times (6 / 20) \\ &= 0.24 + 0.0075 + 0.0075 + 0.09 \\ &= 0.345 \end{aligned}$$

Again, applying the data in Table 3.2, the overall Kappa agreement, correcting for chance is:

$$\kappa = \frac{0.75 - 0.345}{1 - 0.345} = \frac{0.405}{0.655} = 0.618$$

Comparing the chance corrected agreement ( $\kappa$ ) of 61.8% against the direct percentage ( $P_o$ ) of 75 % we can observe that the agreement is lower.

Interpretation of the Kappa statistic is not strictly defined. Fleiss [158] characterises the agreement of the Kappa calculation in broad terms as follows:

<b>Kappa Result</b>	<b>Rating</b>
<0.40	Poor
0.40 to 0.60	Fair
0.60 to 0.75	Good
>0.75	Excellent

*Table 3.3 : Fleiss Kappa statistic interpretation*

This reliability criterion is also derived by Landis and Koch [159]. Others such as Krippendorff [160] are more subjective in their rating, concluding that a value of  $\kappa > 0.8$  indicates good reliability and  $0.67 < \kappa < 0.8$  “allows tentative conclusions to be drawn”. A more statistical assessment of whether the Kappa result indicates significant agreement between assessors is to examine the  $z$  statistic or the *standard score*. This is obtained by the following formula:

$$z = \frac{\kappa}{\sigma_{\kappa}} \quad (3.4)$$

where :

$\kappa$  = Kappa statistic

$\sigma_{\kappa}$  = standard deviation of data used in calculating Kappa

$z$  produces an assessment which indicates how the value of Kappa deviates from the zero position in a normal distribution. A value above 1.65 (or  $-1.65$  due to the symmetry of the distribution) indicates that Kappa differed significantly from zero at the 95% significance level or better.

Applied to the 11 assessor BIT interrater trial, an overall statistic for a particular subtask is obtained by calculating a Kappa score (and standard error) for all pairs of assessors and a mean taken of these values.

### 3.3.1.3. Kendall's Coefficient of Concordance ( $W$ )

Kendall's  $W$  statistic is also used widely to assess the agreement between multiple assessors. Instead of using the raw data values, results are *ranked* and then analysed to find rater agreement. Because of its method of computation, the  $W$  statistic is most suited to trials where assessors have to rank (and thus find the agreement in ordering) a range of objects on a particular feature, rather than applied directly to quantitative data such as that obtained from the BIT trial. Kendall's  $W$  has the advantage of dealing with multiple raters in a single calculation, therefore not requiring pairwise-means to obtain an overall result.  $W$  is obtained by finding the variance in ranks for each assessment variable which is then divided by the maximum variance in column totals to obtain a value between 0 and 1. If there is no variance in ranks then there is total agreement between assessors for a particular variable.

$W$  cannot be easily utilised in assessing the BIT interrater results. As the marking range for each of the sub-tests differ then separate assessment and ranking must be computed separately for each task. As in many cases the quantitative data from the sub-tests are at the extremes of the marking scheme (see Section 3.3.1.1), many of the ranks would be identical and not contain a continuous range of values that are most suited to ranking variables.

### 3.3.1.4. Cronbach's Alpha ( $\alpha$ )

Cronbach's Alpha [161] measures the average covariance between items within a series of test results therefore not requiring the standardisation of a marking range across all test data. This is useful in the assessment of BIT data which has variability in the number of marks awarded for each subtask. Bland and Altman [162] state a more direct interpretation of Alpha. If two random samples of  $k$  items were taken from a data set and summed then this would result in two separate scores from these selected items. Alpha represents the expected covariance between these scores.

Alpha is calculated using the following formula :

$$\alpha = \frac{\left( \frac{k \times cov}{var} \right)}{1 + \frac{((k-1) \times cov)}{var}} \quad (3.5)$$

where :

$k$  = number of items in test.

cov = mean covariance between items.

var = mean variance between items.

From an examination of Equation 3.5 it can be noted that the value of Alpha is dependent on the number of items within the test as well as the covariance between items. Consequently, if an identical covariance is obtained for two tests, a higher Alpha result is obtained for the test with more items reflecting a wider agreement between assessors.

Interpretation of the Alpha score is again not standardised mainly because of the relationship between sample size and Alpha result. Bland and Altman [162] also highlight several other medically based studies using this statistic, suggesting a satisfactory agreement between assessors produces a value of  $\alpha > 0.7$ .

Cronbach's Alpha is implemented within the SPSS statistics package [163] facilitating simple calculation of the statistic. In particular, by calculating an individual Alpha score following the removal of a specific individual data item (for example the number of cancellations from the star task), it can be observed how that particular item affects the overall reliability scale, providing evidence about which results cause the most agreement and disagreement between assessors.

### 3.3.1.5 Intraclass Correlation Coefficient (ICC)

The ICC [164] [165] measures agreement by assessing both the variance between assessors and within individual rater assessments. Specifically, it analyses the interaction between an assessor and test subject's responses (how does an assessor modify his assessment when presented with the responses of a particular test subject).

ICC is calculated using the following formula :

$$\text{Intraclass correlation} = \frac{MS_{tsub} - MS_{rate \times tsub}}{MS_{tsub} + (r - 1)MS_{rate \times tsub} + \frac{(r(MS_{rate} - MS_{rate \times tsub}))}{n}} \quad (3.6)$$

where :

$MS_{tsub}$  = Between test subjects mean-square.

$MS_{rate}$  = Within test subject mean-square.

$MS_{rate \times tsub}$  = Interaction mean-square between rater and subject.

$r$  = number of raters

$n$  = number of subjects

The mean squared results can be taken from a standard two way analysis of variance (ANOVA) calculation [166] .

As with Kappa and Alpha, the result of ICC is in the range of 0.0 (no agreement) to 1.0 (perfect agreement). ICC will return a value nearer to 1.0 if the agreement between raters is high with small differences caused by interaction effects despite a potential large variability between test subject performance. However, if there is a global disagreement between assessors or an interaction effect caused by a particular assessor not consistently applying a marking scheme across all test subjects, then the ICC result will be lower. The results of a low ICC indicate that there is an inability for the assessors to apply a consistent and uniform marking scheme either *individually* (variability of a single rater applying a scheme) or *globally* (raters cannot agree on how to apply the scheme).

### 3.4 BIT Agreement Results

Table 3.4 details the total conventional battery marks awarded by each of the assessors for the 10 complete sets of BIT responses.

Of particular interest is the mark variation for each test response set. Here, a larger value indicates that the patient's test responses produce a greater error in interpreting the marking scheme by the assessors. This may be due to ambiguity in a particular drawing or response

Test Response Set	1	2	3	4	5	6	7	8	9	10
<b>Assessor</b>										
<b>1</b>	146	138	146	106	42	146	122	129	81	146
<b>2</b>	145	138	145	105	42	146	122	128	87	146
<b>3</b>	144	136	146	105	42	147	124	130	86	146
<b>4</b>	144	136	146	100	42	146	122	127	84	146
<b>5</b>	143	137	146	104	42	146	112	128	84	146
<b>6</b>	144	138	146	103	42	146	125	143	87	146
<b>7</b>	144	138	146	104	42	146	122	144	83	146
<b>8</b>	146	137	146	104	42	145	122	129	84	146
<b>9</b>	146	137	146	104	42	145	126	131	89	145
<b>10</b>	146	138	146	105	42	145	123	128	84	146
<b>11</b>	146	137	146	108	42	145	122	130	85	146
<b>Max</b>	146	138	146	108	42	147	126	144	89	146
<b>Min</b>	143	136	145	100	42	145	112	127	81	145
<b>Mean</b>	144.91	137.27	145.91	104.36	42	145.73	122	131.55	84.91	145.91
<b>Variation</b>	3	2	1	8	0	2	14	17	8	1

*Table 3.4 : Mean BIT results and marking variation*

(for example, uncertainty to whether a particular target has been cancelled or to whether a component has been drawn or not) or the assessor miscalculating the result (for example, miscounting the number of cancellations made or assessing incorrectly the distance from the midpoint to the bisection line by misreading the marking scale) on one or more of the subtasks. Various possible reasons for the ambiguity are examined further in Section 3.4.3.

By examining the variation in marks, it can be seen that test subjects 4, 7, 8 and 9 produce the most variation. Test set 8 is of particular interest as six of the assessors score the battery at or below the 129 neglect threshold, whereas the other five are above this mark. This has implications for the clinical classification of the particular patient based on the interpretation of assessment. Only one out of the ten battery sets (test set 5) produced total agreement between assessors. This test subject exhibited severe neglect and failed to produce any responses for the drawing tasks. All assessors uniformly scored these tasks with 0 marks. Due to the severity of the neglect, the cancellation tasks contained many omissions which simplified counting the number of correct cancellations made. For example, only 12 of the 40 lines were cancelled on the Line cancellation (Albert's) task, all of which were positioned in two columns to the extreme right of the page which aided the assessment of the cancellations. Again, all assessors uniformly scored the cancellation tasks due to the low number of responses.

Applying the Kappa statistic across all assessors for the total battery marks, a mean Kappa score of 0.795 is obtained. Using Fleiss' Kappa statistic interpretation, this indicates a satisfactory agreement between assessors which is confirmed by the high significance of the  $z$  value (21.62). The direct number of agreements ( $P_o$ ) is 0.81 or 81%. Comparing the Kappa score (0.795) with  $P_o$  (0.81) shows that the difference between the agreement probability due chance and actual observed direct agreement is minimal (0.015). The levels of agreement are supported by a high ICC score of 0.994.

Analysis of the agreement between assessors using the overall conventional marking scheme presents two opposing outcomes on reliability. Whilst the calculated statistics indicate very satisfactory agreement between assessors, inspection of the individual battery scores shows that misclassification does occur (in test subject 8). Given that a diagnosis of neglect is a contributing factor to the selection of an individual rehabilitation scheme, accurate assessment is critical.

### 3.4.1 Subtask Agreement

One of the design objectives behind a computer-based assessment scheme is to remove any subjectivity within a marking scheme. To obtain a more detailed understanding of the component parts of the BIT, the assessors' agreement in individual subtasks of the BIT was analysed. By assessing the agreement, an indication to the ambiguity of each subtask's marking criteria can be established. A disagreement indicates a varied interpretation of the defined assessment scheme. Table 3.5 presents subtask Kappa agreement statistics for the 11 assessors over 10 sets of complete BIT responses.

Task	$\kappa$	$P_o$
Line Cancellation	0.879	0.938
Letter Cancellation	0.765	0.816
Star Cancellation	0.590	0.655
Figure and Shape Copying	0.677	0.798
Line Bisection	0.813	0.895
Representational Drawing	0.578	0.765

Table 3.5 : Subtask agreement for 10 sets of BIT responses

Unsatisfactory agreement (using Fleiss' Kappa interpretation) occurs for the star cancellation task and the two drawing tasks. Whilst the marking scheme for all three of the cancellation tasks involves the assessment of the correct number of targets marked, the arrangement of the targets on the star overlay is in a random configuration, as opposed to a pseudo-random grid formation of the line and letter cancellation tasks. Beckwith and Restle [167] found that counting objects in a random array was more difficult than when objects were arranged in linear rows (horizontal or vertical) as the former required the implementation of a scanning strategy. Without the imposition of linear structure, an assessor does not have a forced scanning strategy for the assessment of individual cancellation targets and is not immediately aware if a target has been 'counted', resulting in scoring errors. The application of an automated computer-based assessment of the cancellation tasks would eliminate these scanning and counting errors.

Reasons for disagreement in assessing the drawing tasks are investigated in Section 3.4.3.

### 3.4.2 Drawing Subtask Agreement

As shown in the previous section, the drawing tasks cause an unsatisfactory level of agreement between assessors who have to establish their own marking rules and criteria which is subject to variation during the marking process. Two hypotheses to the self-devised criteria used by assessors when marking drawings are:

- **Component Level** : assessment of number of sides or objects (such as wings or body for the butterfly task) within a drawing or range of drawings and applying a scheme based on the presence of components. While awarding marks may be a simple task for the geometric shapes, for example a mark for each side or vertex of a triangle or square, definition of components is more subjective for the representational drawing tasks. Failure to complete a component, accuracy in forming vertices and spatial arrangement may also influence whether the assessor judges a particular component as being present.
- **Marking Range** : assessment of shapes is performed by assigning grading criteria across the range of marks available. For example, full marks would be awarded to a 'perfect' drawing down to zero marks for 'no response' or 'unrecognisable drawing' with the intermediate classifications assigned subjectively.

As well as applying the standard BIT marking scheme to the drawing tasks, each assessor was asked to mark individual shapes and drawings from the 20 sets of BIT data. To further highlight the subjectivity within the drawing tasks, Table 3.6 shows the levels of agreement over these 20 test responses using the standard BIT assessment for the two drawings tasks. This table shows increased disagreement over a larger number of assessments particularly within the Representational Drawing task.

<b>Task</b>	$\kappa$	$P_o$
Figure and Shape Copying	0.578	0.732
Representational Drawing	0.178	0.427

*Table 3.6 : Drawing subtask agreement between assessors for 20 responses*

From the assessment of *individual* shapes and drawings (rather than an single assessment over the 3 or 6 shapes drawing) we can establish which produced the most disagreement, indicating a level of difficulty in applying a standardised marking scheme. The range of marks used was 0 to 4 for the figure and shape copying and 0 to 3 for the representational drawings. Table 3.7 details the agreement analysis.

<b>Drawing</b>	$\kappa$	$P_o$
Star	0.460	0.691
Cube	0.226	0.389
Daisy	0.300	0.452
Shape 1	0.323	0.460
Shape 2	0.413	0.565
Shape 3	0.186	0.391
Clock Face	0.352	0.515
Person	0.254	0.429
Butterfly	0.419	0.565

*Table 3.7 : Drawing object agreement between assessors for 20 responses*

While all the agreement values for Kappa and  $P_o$  are unsatisfactory, the data shows that the cube, daisy and shape 3 (as shown in Figures 3.3 and 3.4) produce the most disagreement for the figure and shape copying task and the drawing of a person for the representational drawing task. The results, however, show similar disagreement levels across the entire range of drawings. The next section describes an investigation of the reasons why particular drawings cause such ambiguity in assessment.

### 3.4.3 Patient Response Agreement

The results documented in Sections 3.4.1 and 3.4.2 show the agreement between assessors over a set of 11 conventional test battery responses. In this analysis it has been identified which subtasks produce the most disagreement. In this section, the 10 individual test responses (20 for the drawing tasks) are analysed to establish whether a particular set of responses causes ambiguity in assessment and why such ambiguity should arise.

Cronbach's Alpha was used to establish which of the ten sets of patient responses caused the most disagreement between assessors. By removing from the Alpha calculation the marks awarded to a particular test subject by all of the assessors and then noting whether the overall agreement improves or deteriorates, it is possible to establish the level of disagreement caused by the excluded subject's responses. The response sets were then ranked on levels of disagreement for each subtask.

Table 3.8 shows the 3 response sets for each subtask that cause the most disagreement. The values alongside the response set number are the standard deviation of the marks awarded for the particular subtask response drawn by the identified test subject. These results indicate the severity of the disagreement caused by the specified test response. Line cancellation and line bisection tasks only have two entries as the other 8 batteries result in perfect agreement between assessors.

	<i>Rank</i>		
	<b>1</b>	<b>2</b>	<b>3</b>
<b>Line Cancellation</b>	8 (0.467)	4 (0.302)	-
<b>Letter Cancellation</b>	8 (1.689)	4 (1.213)	9 (0.603)
<b>Star Cancellation</b>	8 (3.668)	7 (3.015)	1 (1.136)
<b>Figure and Shape Copying</b>	14 (1.293)	8 (1.264)	16 (1.213)
<b>Line Bisection</b>	8 (0.522)	6 (0.504)	-
<b>Representational Drawing</b>	9 (1.167)	7 (1.035)	8 (0.934)

*Table 3.8 : Test subjects ranked by disagreement on standard conventional task marking scheme*

The same analysis technique was used on the individual drawing images across 20 battery sets. Table 3.9 shows the three test response sets causing the most disagreement between

assessors. Again, the numbers in parenthesis alongside the battery number detail the standard deviation in marks awarded to a particular response and hence indicate the severity of the disagreement.

<i>Rank</i>			
	<b>1</b>	<b>2</b>	<b>3</b>
<b>Star</b>	1 (1.213)	20 (1.136)	9 (1.078)
<b>Cube</b>	8 (1.264)	3 (0.981)	12 (0.924)
<b>Daisy</b>	8 (1.103)	11 (1.035)	12 (1.026)
<b>Shape 1</b>	12 (1.221)	9 (1.206)	14 (1.128)
<b>Shape 2</b>	9 (1.264)	17 (1.190)	7 (1.120)
<b>Shape 3</b>	4 (1.537)	17 (1.420)	13 (1.414)
<b>Clock Face</b>	19 (1.361)	20 (1.341)	14 (1.167)
<b>Person</b>	8 (1.439)	15 (1.136)	4 (1.128)
<b>Butterfly</b>	19 (1.272)	8 (1.264)	17 (1.136)

*Table 3.9 : Test subjects ranked by disagreement on drawing tasks*

Examining the results contained in Tables 3.8 and 3.9, test response batteries 8, 9 and 12 cause the most disagreement between raters. Visual analysis of these patient's responses shows that very light pen markings were made on all response overlays, causing some of the drawings and cancellations to appear to be absent to the assessor unless closely analysed. Because of this, cancellations and drawing components were calculated incorrectly.

As an investigation to why certain drawings resulted in disagreement between assessors, four responses were analysed from each of the representational drawing models (Section 3.2.1.6), two of which caused the most disagreement and the other two resulting in the maximum agreement between assessors. The responses and results are detailed in Tables 3.10 (clockface), 3.11 (man) and 3.12 (butterfly) showing the two responses assessed with the most agreement (Rank 1 and 2) followed by the responses assessed with the least agreement. The range, average and standard deviation of marks is presented for each response.

Assessments made on test set number 5 were not included in this analysis. No drawing task responses were made by this test subject and all assessors accordingly awarded a mark of 0 for all drawings. While this shows perfect agreement between assessors, the study's aim is to assess reasons for agreement (and disagreement) with drawn images. Two other test subjects failed to produce responses for the man image, resulting in the lowest rank (causing the most disagreement between assessors) of 17.

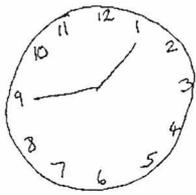
Test Set	Agreement Ranking	Response	Comments
10	1st		Mean score = 4, SD = 0 Max score = 4, Min score = 4
1	2nd		Mean score = 3.909, SD = 0.301 Max score = 4, Min score = 3
20	18th		Mean score = 3, SD = 1.341 Max score = 4, Min score = 0
19	19th		Mean score = 2.636, SD = 1.361 Max score = 4, Min score = 0

Table 3.10 : Clockface drawing responses

Examining the results from the clockface drawings (Table 3.10), the assessors are most in agreement when an image is drawn with a circular edge, 12 clearly drawn hour indicators at the correct positions and two hands. Assessors have a higher level of agreement on shapes that appear 'well drawn' (the subjective assessment of the rater on the overall quality of the drawing - correct number and spatial arrangement of the drawing's components) and hence are given high marks. Shapes that are 'badly drawn' (again, a subjective assessment), where a lower average mark is awarded for the drawing, results in lower agreement (the highest ranked 'badly drawn' image was 7th), although these drawings do not cause the largest disagreement between assessors.

The two images causing the most disagreement both have the correct number of hour indicators, however the numerals are unidentifiable in certain positions. The clockface edge is not at all circular and is drawn with tremor. The response causing the most disagreement is drawn smaller than the other images. Disagreement seems to occur between assessors

marking on a component basis (awarding full marks as all are present) and those assessing on the 'clarity' of the drawing both on a component level (they are all visible *but not* identifiable) or as an entire drawing.

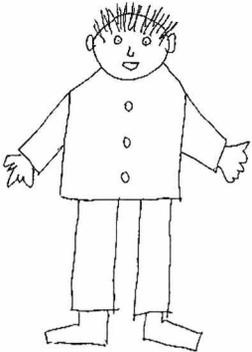
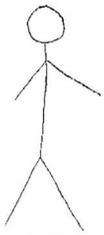
Test Set	Agreement Ranking	Response	Comments
10	1st		Mean score = 4, SD = 0 Max score = 4, Min score = 4
2	2nd		Mean score = 3.909, SD = 0.301 Max score = 4, Min score = 3
15	16th		Mean score = 3.090, SD = 1.136 Max score = 4, Min score = 1
8	17th		Mean score = 2.454, SD = 1.439 Max score = 4, Min score = 0

Table 3.11 : 'Man' drawing responses

Examination of the responses from the representational drawing of a man (Table 3.11) shows an agreement when components (limbs, body, head) are all drawn in the correct position. As with the clockface results, there is also agreement (to a lesser extent) in assessing 'badly drawn' responses. Most importantly, the drawings which cause the most disagreement comprise of simple 'stick' components. As with the clockface, these drawings are drawn smaller than the other responses. While, to some assessors, these stick representations (with the limbs positioned correctly) form a perfect representation of a man, to others, the image is too simple and thus is scored lower.

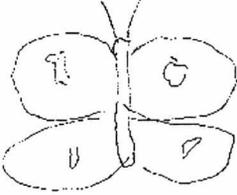
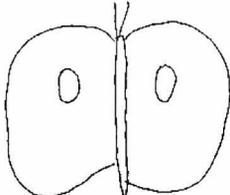
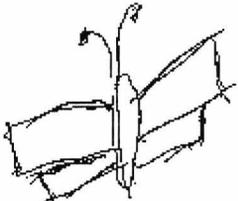
Test Set	Agreement Ranking	Response	Comments
18	1st		Mean score = 4, SD = 0 Max score = 4, Min score = 4
10	2nd		Mean score = 3.909, SD = 0.301 Max score = 4, Min score = 3
8	18th		Mean score = 3, SD = 1.264 Max score = 4, Min score = 0
19	19th		Mean score = 1.727, SD = 1.272 Max score = 4, Min score = 0

Table 3.12 : Butterfly drawing responses

The butterfly drawings (Table 3.12) again show that assessors are able to agree on drawings which are clearly constructed (symmetrical, semi-circular wings) and, as with the other two tasks, are also able to agree on poorly drawn images. The shapes producing the least

agreement once more show that individual interpretation of what constitutes a 'good' image can be based on the number of components present or the overall quality (correctly formed symmetrical components, no perseveration) of the drawing.

This examination of the drawings indicates that it is not the very poor drawings that cause scoring variability, rather it is those drawings which contain the correct number of, and correctly positioned components, but are poorly constructed, non-symmetrical, contain tremor or are not instantly identifiable. Raters assessing on a 'component-present' basis award higher marks than those assessing the quality and clarity of the overall drawing for these images. Very poor and very good drawings are universally recognised by all assessors.

### 3.5 Assessment Timings

A further trial concerning the current administration of the BIT was undertaken to establish the amount of time required to assess a patient. Performance from a patient can be affected by fatigue caused by the test administration, therefore a quicker testing time reduces any result modification due to tiredness. The testing time can also have an effect on the assessor, particularly in marking accuracy.

<b>Patient ID</b>	<b>Conventional Administration</b>	<b>Conventional Marking</b>	<b>Behavioural Administration and Marking</b>
<b>1</b>	26	22	62
<b>2</b>	37	20	58
<b>3</b>	31	25	45
<b>4</b>	50	25	49
<b>5</b>	25	20	42
<b>Mean Values</b>	33.8	22.4	51.2

*Table 3.13 : Assessment and marking times (in minutes) for the conventional and behavioural BIT batteries*

The times taken for a therapist to *administer* and then *mark* five individual conventional BIT test responses from stroke patients at the Nunnery Fields Hospital, Canterbury, were recorded separately during normal therapy sessions. These times were also recorded for the behavioural

subtasks, indicating the resources required to assess an individual patient with both BIT batteries (The behavioural subtasks were marked as they were administered). As trained therapist resources are at a premium within a hospital environment, this indication of resources consumed in testing can be seen as a notional ‘norm’ in current practice, on which an automated system should seek to improve. Table 3.13 details these timings.

From these data, we can calculate an average assessment and marking time of 56 minutes for the conventional BIT test battery. If supported by the behavioural test battery, the average time is increased by an additional 51 minutes. This compares with the computer based test which on average takes 25 minutes to both administer and assess the responses.

### 3.6 Summary

In this chapter, the subtasks of the current standard for visuo-spatial neglect testing, the Rivermead Behavioural Inattention Test, have been introduced and methods for assessment examined. The proportion of marks each subtask contributes to the total battery score shows a bias towards the cancellation based tasks.

A methodology to assess the levels of disagreement between assessors of the BIT was defined and suitable statistical measures for examining the agreement levels were presented. The study validates the interrater reliability of the BIT using the total battery score for the conventional tasks. This indicates that the BIT performs satisfactorily in identifying neglect patients across multiple assessors. Combined with the original validation of the test battery against both clinical assessment and ADL and the acceptance of the BIT within the medical profession, the BIT can be used with confidence as a standard on which to base patient grouping, and with which to validate a computer-based system.

Examination of the subtask scores reveals significant levels of disagreement in the drawing tasks, particularly where the assessment scheme is subjective (the drawing tasks) or confusing for the assessor (the star cancellation task). This identifies an area in which a computer-based test system can improve accuracy and repeatability in marking a test subject’s response. Agreement across all of the drawing attempts was unsatisfactory. Subtasks with simple and unambiguous marking schemes such as the line cancellation and bisection tasks produce satisfactory agreement.

---

Levels of disagreement on drawings are caused by individual interpretation of the basic assessment criteria. Drawings which contain the requisite number of components and correctly positioned but are badly formed cause the most disagreement between assessors. Marking assessment rules can be imposed algorithmically by a computer-based assessment system which can also be used to detect fine and light-pressured pen movement by normalising drawing pressure or lowering the threshold at which a drawing mark is detected on an overlay. This consistent application of a marking scheme leads to an increase in accuracy and standardisation. The reduction in test administration and assessment time over the BIT using the computer-based system lessens patient fatigue and frees therapist resources.

## Chapter 4

# Experimental Infrastructure for Pen Based Data Capture

### 4.1 Introduction

Before implementing software to capture data and extract features from a computer-based test of neglect, the design options, constraints and requirements for the test infrastructure along with the range of data required to maximise extractable features need to be considered. This chapter explores the practical and theoretical issues concerning pen-based data capture, in particular with reference to the design of a system for use within a hospital and clinical environment.

Following an assessment of the system requirements, the options for an input peripheral device are investigated and the selected Wacom graphics tablet and communications protocol are presented. Data handling and storage requirements for captured test responses and practical issues of data pre-processing prior to feature extraction are addressed, specifically the filtering and interpolation of raw pen coordinate data. Finally, examples of feature extraction methodologies are presented.

### 4.2 Handwritten Data Capture Requirements

The highest level design requirement for the system is the capture of drawing data in real time and the extraction of a series of diagnostic features based on the pen movements. Research into handwriting dynamics has defined many physical properties of normal handwritten performance [168][169][109], the bounds of which must be within the specification of the chosen input device to prevent any restriction in selectable features.

As defined in Chapter 2, the principal aims of a computer based implementation are to firstly to validate the computer based system by obtaining comparable *static* (positional/image)

feature results between the computer and the traditional pencil and paper-based neuropsychological tasks. Secondly, to investigate when novel *dynamic* (constructional /timing) features extracted from response data can be used to classify a test subject. To attain test environment consistency between the computer test and the traditional test battery the choice of the hardware must not impair or modify standard pencil and paper drawing conditions, thereby preventing the introduction of another patient performance variable. Geriatric patient performance is often affected by apprehension about using technology [98] meaning that direct access, for example through the use of keyboards and mice which require a degree of computer literacy (albeit modest) and competency in use are not suitable for this type of patient testing.

#### 4.2.1 Data Capture Peripheral

Teulings and Maarse [168], Maarmari and Plamondon [169] and Marquardt and Mai [109] all present detailed theoretical background to the area of handwriting data capture, all three studies using a graphics tablet as an input device. Whilst the graphics tablet is the standard data capture device for handwriting and drawing analysis, other pen based devices exist such as the digital ink pen and screen based tablets. These products all enable the capture of drawing data, but performance is modified in that the feedback from the pen device is not identical to a pencil and paper task.

The digital ink device consists of a conventional pen with a small roller-ball mounted at the pen tip. As the pen is moved, its position relative to the previous location is reported. Pen movement can therefore be detected, but only within individual drawing components when the pen is on a drawing surface. Relative spatial positioning of components within drawings are not obtainable as the pen does not report locations when the pen is removed from the table surface. The screen based tablet, while using a pen for input does not use paper markings for visual feedback, instead presenting the drawn data graphically on a screen. The current cost of these devices is also prohibitive from routine use.

A graphics digitisation tablet with a marking ink pen is the obvious peripheral for data capture. Paper can be overlaid on the tablet surface and marked using the pen, maintaining the test environment of the traditional test system. Many test subjects have been impaired by the adoption of older style graphics tablets [170] where the pen is attached to the tablet via a

cable. Whilst not directly intrusive to the writing and drawing style, the standard pencil and paper test configuration is not maintained. The introduction of tablets with cordless pens of similar dimensions to a normal pencil overcomes this problem. Standard graphics tablets report the position of the pen relative to an origin and also a range of additional items such as pen pressure and tilt. Analysed, these data allow the extraction of static features. Dynamic features require the use of a time-stamp which specifies the time offset at which a pen data packet was captured. Order of construction, pen velocity and other rate of change measures can be extracted by using this time-stamp which is added by the computer to the incoming status packet as the data is stored.

Marquardt and Mai [109] define three type of errors inherent in collecting data from a graphics tablet. These need to be considered to ensure the data provides an accurate representation of the drawn response:

- *Spatial* errors are caused by a limited resolution or sample rate, noise, non-linearity of reporting surface and missing coordinates. Interpolation of positional data can restore lost coordinates and smooth non-linearity. Filtering of data signals can remove noise and hence ‘smooth’ the response.
- *Temporal* errors are caused by irregularities in data sampling times. Again, time based interpolation can restore regularity.
- *Intrinsic* errors are introduced by the method of sampling such as pen tilt affecting the  $x$  and  $y$  position coordinates. Smoothing of data can remove the signal noise introduced by this error.

By designing and configuring the system to limit the impact of these inherent errors on the quality of captured data, for example correct choice of peripheral and selection of adequate sampling rates, the reliance on pre-processing operations implemented to overcome data loss can be reduced.

### 4.2.2 Sampling Rates and Spatial Resolution

The sample rate of the tablet determines how often the pen position and other pen status data (such as pen pressure and tilt) is transmitted to the connected computer. Selecting too low a sample rate introduces spatial errors and movement data is lost, resulting in an inaccurate representation of the drawing. Too high a sample rate and the computer is unable to fully buffer the incoming data and temporal timing error are introduced. Maarmari and Plamondon [169] analysed normal handwriting performance by extracting a range of features such as pen velocity and acceleration as well as displacement calculations. They found that a sample rate of 100Hz did not allow the accurate extraction of acceleration components [171] and fine transient responses contained within several biomechanical models of handwriting performance (such as that described by Plamondon [172]). This problem can be overcome by interpolating data as described in section 4.6.2., effectively increasing the sample rate. For general movement and displacement calculations, however, experimental evidence has shown that the highest frequency observed in drawing displacement data is in the range of 13.6 to 20Hz [173][169]. Taking this upper limit, the Nyquist frequency required to sample pen movement without loss of data is 40 Hz which is well within the range of the graphics tablets currently available. Any high frequency components of the input signal can be considered to be insignificant as part of the hand-drawn image (for example noise from the tablet) and can be removed using low-pass filtration techniques.

Spatial resolution defines the level of detail that it is possible to capture using the digitising device. A coarse resolution leads to omission of fine details within drawings (spatial error). Typical values of resolution for current graphics tablets are around 100 lines per mm. However, the overall resolution is affected by the accuracy of the chosen pointing device. Using the Wacom Inkpen with the UltraPad series of tablets reduces the overall resolution to 6.25 lines per mm. This resolution produces handwriting data which, although it suffers from quantisational spatial errors, can be pre-processed (interpolated) prior to feature extraction to produce a representation of handwriting data of acceptable quality for most current applications.

### 4.2.3 Additional Pen Based Features

Alongside the positional data, the graphics tablet can report other information such as pen pressure and  $x$  and  $y$  coordinate tilt. Schomaker and Plamondon [174] examined the relationship between pen pressure and tilt by constructing a simple biomechanical model. By resolving the forces in each axis, the findings showed a small or negative correlation in normal and cursive handwriting indicating the high complexity of the motor control involved. The pressure data, however, resulted in a cyclical profile when executing a cursive script. While this profile occurs in responses captured from a normal asymptomatic population, comparison of the cyclical pressure profile obtained from stroke patient responses may reveal differences and changes throughout the duration of the task. Other studies [175] have examined the relationship between pen force and velocity. These additional data items provide information on the motor-based constructional aspects of drawing. Although not a primary aim of the neglect assessment, the relationship between the motor and positional components of a hand-drawn attempt could provide additional diagnostic indicators.

### 4.2.4 Data Capture in the Clinical Environment

The ergonomic specifications of the data capture device are of importance especially within the field of neuropsychological and clinically based studies. A number of recent studies have used graphics tablets to capture hand-drawn data from stroke patients [2] and an elderly population [113] proving the suitability of using the device with the target patient group. As many of the test subjects will be confined to a bed, the ability for the test equipment to be set up on a bedside hemi-table or other confined environment whilst not restricting the standard pencil and paper test environment is important. The equipment therefore needs to be portable, lightweight, robust and able to be used in environments with limited space and access. For use within a clinical environment, the developed software must be intuitive, simple to use and robust. Emphasis within neuropsychological testing is placed on standardisation in test procedure [176] with patient instructions, environment [177] and seating position [178] specified for all test attempts. The tablet must therefore be used in an environment in which the tablet can be uniformly positioned directly in front of the test subject and does not restrict particular sub-sets of test subjects (for example test subjects confined to wheelchairs). The physical size of the tablet surface must be able to accommodate the test overlays and the paper fixing system without impairing the normal handwriting movement of the test subject.

### 4.3 System Infrastructure

The schematic of the developed computer-based system infrastructure is shown in Figure 4.1. Data capture, feature extraction and processing occur in three distinct phases. The computer-based system has been designed so that these phases are non time-dependent, meaning that data can be captured and analysed (and re-analysed) at the convenience of the test administrator.

**Phase 1 - Data Capture** : Test overlays are placed and secured individually on to the surface of the graphics tablet. The test subject completes each task by drawing or marking directly onto the surface of the overlay. Each overlay constitutes a *test attempt*. As the test subject draws on the overlay with the pen, *raw drawing data* such as pen coordinate position and pressure are captured and transmitted via a serial link to an attached computer. As the *raw drawing data* is received by the computer, each packet is stored sequentially along with a timestamp (referenced to the start of data capture) in an ASCII *test response* file. A separate *test response* file is created for each test attempt.

**Phase 2 – Feature Extraction** : Features are classified as being either *static* - based on the drawn pen coordinates and include features such as drawing area, number of shape components - or *dynamic* - based on timing or constructional properties of the drawn data, such as construction order or time to complete a specific element of a shape drawing. Individual features are extracted from the completed test response files rather than in real-time from the tablet data-stream. The advantages of storing the raw data on the computer are twofold: Processing the input stream in real-time reduces the report rate of the tablet due to the increased processing required. This would be particularly pronounced if the raw data is concurrently stored in a file. Secondly, by re-assessing individual response files, new and developing features can be extracted without the necessity to retest subjects; particularly important within a research environment where feature parameters and calculations are constantly under review. Data can also be analysed repeatedly and replayed for on-screen visual analysis. The feature set of extracted results for each test response file is stored back on the computer in a separate *feature set* file.

**Phase 3 – Classification and Report** : The amount of the output data from the system depends on the requirements of the end-user and thus the system's output must be flexible to a wide range of needs. Whilst a researcher will often be interested in individual features from

all tasks as a performance metric, a clinician requires a concise performance indicator over the entire test battery to assess a patient's progress. To obtain a single indicator or series of such indicators, multidimensional pattern recognition techniques can be used to reduce and classify the feature sets from subtasks and the entire battery (Section 6.12). Features and classifications are written to a *results file* which may be investigated using a spreadsheet or other means, whilst the program must provide a concise report of the patient's performance for use within a clinical environment. For the researcher, the individual feature set files will allow quantitative feature data to be analysed.

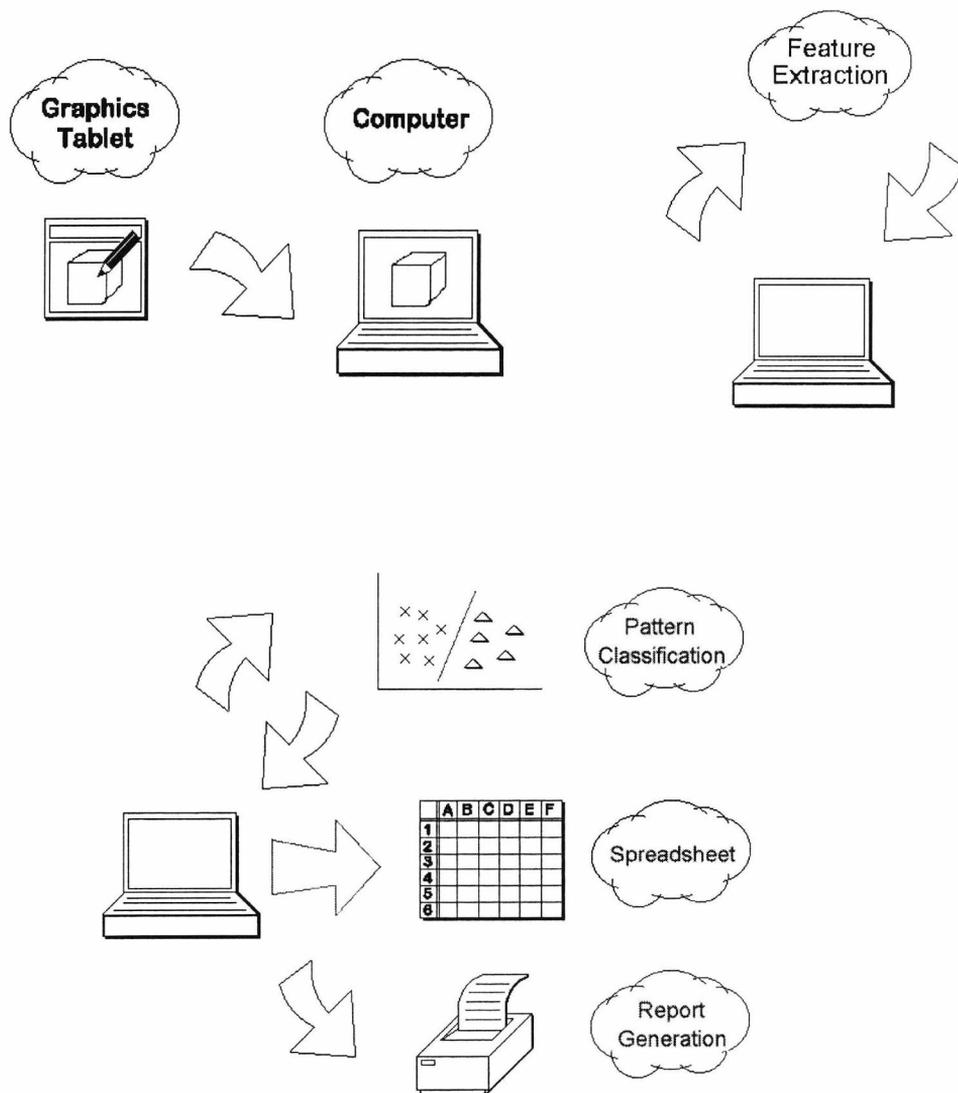


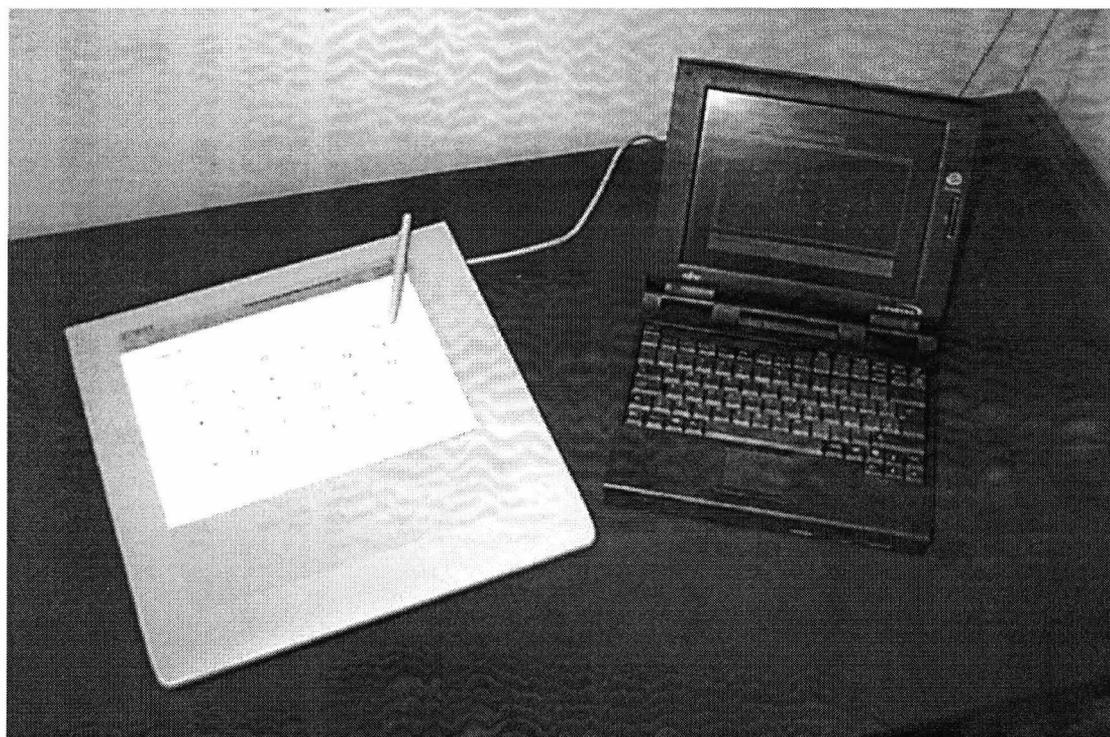
Figure 4.1 : 3 phase test system schematic

## 4.4 Wacom Tablet and Portable Computer

The implemented system uses a Wacom WD1212 UltraPad graphics tablet to capture the drawing data. This tablet has a drawing area of 304.8 mm<sup>2</sup> enabling an A4 sized overlay to be placed on the tablet in both orientations (portrait or landscape). The Wacom tablet uses a series of magnetic inductance based pens all of which are cordless and require no internal power supply. This means that the input device can be of comparable dimensions to a normal pen, thus maintaining the standard test configuration. An ink (biro) pen is used to draw onto the overlay providing normal visual feedback for the test subject. The tablet has a resolution of 100 lines per mm, resulting in a coordinate range in each axis of 0 to 30480. The pen has a lower resolution than the tablet, being accurate to 0.15mm, so in practice the range of coordinates reported are not continuous. This spatial resolution, however, is sufficient for the capture of hand-written data as demonstrated by the system's use in other pen based studies [2].

The test overlays are fixed on the tablet surface by a specially designed clamping system. Firmly securing the overlay to the surface is important for two main reasons: firstly to prevent the paper moving when drawn upon, thus distorting the recorded image. Secondly, as many of the tasks are location specific, relying on assessment of accuracy against a mask file containing model response coordinates (for example location of targets on an cancellation task), the positioning needs to be exact. The fixing method must not, however, restrict movement across the tablet. Because of this, large clamping mechanisms either side of the tablet cannot be used.

Figure 4.2 shows the selected test apparatus. The implemented system uses a straight location edge mounted at the top of a plastic base. The overlay is secured at each side by two low profile paper clamps adapted from a paper document folder. These clips allow the rapid interchange of overlays during the administration of the test battery, whilst keeping the paper taut and flat during drawing. The plastic base is attached to the tablet surface by strips of Velcro. This prevents the base from slipping from a calibrated position but allows removal if normal use of the tablet is required. Because pressure values are calculated at the pen tip and transmitted back to the tablet rather than by measuring pressure on the actual tablet surface, pressure values are still accurate despite the pen status being monitored through the layers of the overlay and the base plastic sheet.



*Figure 4.2 : Test apparatus*

The computer used for the trial is a standard portable PC. No special interfacing is required as the graphics tablet directly interfaces with the serial port. Consideration was given to the amount of file storage that would be required to hold a large number (~200) of test battery responses. The average storage requirement for a complete set of battery responses is 590Kbytes, so 200 test subject each performing a single attempt of the test battery requires disk space of around 100Mbytes. This capacity is easily obtainable using current storage technology and therefore file compression is not required.

#### 4.5 Data Transmission Protocol and Storage

Pen position data is transmitted in packet form to the computer via a serial link at a baud rate of 19200 bits per second using the Wacom IVE protocol. A total of nine data items reporting the status of the pen (Table 4.1) are sent in a packet of nine bytes with certain data items being distributed over two or more bytes [179]. This results in a maximum report rate of 205 points per second (205 Hz). In practice this rate is reduced by the processing time required to

obtain the information by the tablet, transmit the data, decode the data packet and write the data to a file. On average 100 data packets per second (100 Hz) are transmitted to the computer. Data is not transmitted from the tablet when the pen is out of range of the tablet (over 5 mm above tablet surface).

<b>Data Item</b>	<b>Bit Count</b>	<b>Value Range</b>	<b>Description</b>
Proximity	1	0 to 1	1 if pointing device detected, 0 otherwise
Pointer	1	0 to 1	1 if pointing device is a cursor, 0 otherwise
Button Flag	1	0 to 1	1 if pointing device pressed, 0 otherwise
X Position	16	0 to 30480	X coordinate
Y Position	16	0 to 30480	Y coordinate
Button Value	4	0 to 15	*Button data (pen tip, pen barrel button etc.)
Pressure	8	-127 to 127	Pressure of pen on tablet
X Tilt	6	-31 to 31	X tilt value (-ve to the left)
Y Tilt	6	-31 to 31	Y tilt value (-ve to top)

\* Button Value item describes the combined status of the pen barrel button (whether pressed or not) and, mirroring the Button Flag item, whether the pen tip is on the tablet surface. The data field is 4 bits wide to accommodate the three additional buttons found on the Wacom data entry puck not used in this research.

*Table 4.1 : Wacom IVE data components*

An example of a single data packet returned from the tablet is shown in Table 4.2. The 9 raw data bytes (a) are converted into binary representations (b) from which the individual data items are reconstructed (c) according to the defined Wacom IVE protocol. Once the data packet has been assembled into individual data items, they are written to the test response file, stored within a separate directory structure organised by test subject identifier and test battery attempt number.

<b>Byte</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
<b>Value</b>	200	15	92	24	8	45	50	1	20

*Table 4.2a : Initial decimal packet contents*

	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0
<b>Byte</b>								
<b>1</b>	1	Prox. 1	Point. 0	0	B.Flг. 1	0	X.15 0	X.14 0
<b>2</b>	0	X.13 0	X.12 0	X.11 0	X.10 1	X.9 1	X.8 1	X.7 1
<b>3</b>	0	X.6 1	X.5 0	X.4 1	X.3 1	X.2 1	X.1 0	X.0 0
<b>4</b>	0	B.3 0	B.2 0	B.1 1	B.0 1	P.0 0	Y.15 0	Y.14 0
<b>5</b>	0	Y.13 0	Y.12 0	Y.11 0	Y.10 1	Y.9 0	Y.8 0	Y.7 0
<b>6</b>	0	Y.6 0	Y.5 1	Y.4 0	Y.3 1	Y.2 1	Y.1 0	Y.0 1
<b>7</b>	0	SP 0	P.6 1	P.5 1	P.4 0	P.3 0	P.2 1	P.1 0
<b>8</b>	0	SXT 0	XT.5 0	XT.4 0	XT.3 0	XT.2 0	XT.1 0	XT.0 1
<b>9</b>	0	SYT 0	YT.5 0	YT.4 1	YT.3 0	YT.2 1	YT.1 0	YT.0 0

Bit identification:

**Y.xx** – Y coordinate bit (0 to 15)

**X.xx** – X coordinate bit (0 to 15)

**P.xx** – Pressure bit (0 to 6),

**SP** – 0 if pressure is positive, 1 if negative (two's complement)

**XT.xx** – X axis pen tilt (0 to 5),

**SXT** – 0 if x axis tilt is positive, 1 if negative (two's complement)

**YT.xx** – Y axis pen tilt (0 to 5),

**SXY** – 0 if y axis tilt is positive, 1 if negative (two's complement)

**Prox.** – Proximity Flag, **Point.** - Pointer Flag, **B.Flг.** – Button Flag.

*Table 4.2b : Binary representation of packet and bit identification*

<b>Data Item</b>	<b>Value</b>
Proximity	1
Pointer	0
Button Flag	1
X Position	2012
Y Position	1069
Button Value	3
Pressure	100
X Tilt	1
Y Tilt	20

*Table 4.2c : Reconstructed data items*

Figure 4.3 is an example of a section of a test response file containing the reconstructed raw drawing data. When the data is written to a file, a time stamp is added to the data. This represents the time in milliseconds since the start of the capture file and enables the use of time-based dynamic parameters such as pen velocity and constructional analysis. The first column of this data is the time stamp in milliseconds followed by the pen  $x$  and  $y$  coordinates. The other data items in the file represent pen proximity, button value, pen pressure and  $x$  and  $y$  axis tilt. This data can be pre-processed prior to extraction of features, for example by increasing the sampling rate through interpolation between data items (Section 4.6.2) and filtering the data to remove noise and other high frequency components (Section 4.6.3).

```

510 2471 2107 1 1 25 14 35
520 2470 2118 1 1 28 14 35
530 2469 2128 1 1 30 14 34
540 2468 2138 1 1 32 14 34
550 2468 2149 1 1 36 14 34
560 2466 2163 1 1 39 14 33
570 2461 2177 1 1 46 14 33

```

*Figure 4.3 : Example test response file*

Individual data response files are identified by two numbers combined to form a single filename. The test battery comprises seven subtasks, each subtask being assigned a number detailed in Table 4.3. Modifications to the test battery have occurred during the course of the study and subtask identifiers for redundant tasks have been reused. The test response filename is composed of the subtask identifier followed by the overlay number within the subtask. A file extension of ‘.tst’ is used for all test responses files.

<b>Subtask Identifier*</b>	<b>Subtask</b>
1	Point Location Task
2	Line Bisection Task
3	OX Cancellation Task
4	Albert's Cancellation Task
5	Figure Completion Task (Ver.1) Drawing Tasks (Ver.2)
6	Figure Copying Task (Ver. 1)
8	Drawing from Memory Task (Ver. 1)

\* ID 7 was assigned to a Trail Making Task which was included during task development but removed from both versions of the test battery.

*Table 4.3 : Task identification numbers*

A particular test subject may undertake the test battery several times. To prevent overwriting, the test response files for each attempt at the battery is stored in a separate directory within the test subject identification structure (Figure 4.4). A *data set presence file* stores the time and date of each test attempt and is used to calculate the next set number when additional attempts are made by a particular test subject at the test battery. This file is created at the start of each battery attempt thus preserving files if only part of the whole battery is completed.

#### 4.6 Pre-processing of Raw Test Response Data

The test response files contain raw data directly reported by the graphics tablet. In a number of cases this data needs to be pre-processed prior to extraction of features. This section details three operations performed on the raw data to increase accuracy in assessing time based features (interpolation), remove noise and other high frequency components which are not part of the handwritten response (Gaussian filtering) and location offset to remove errors introduced by the misalignment of overlays on the surface of the graphics tablet.

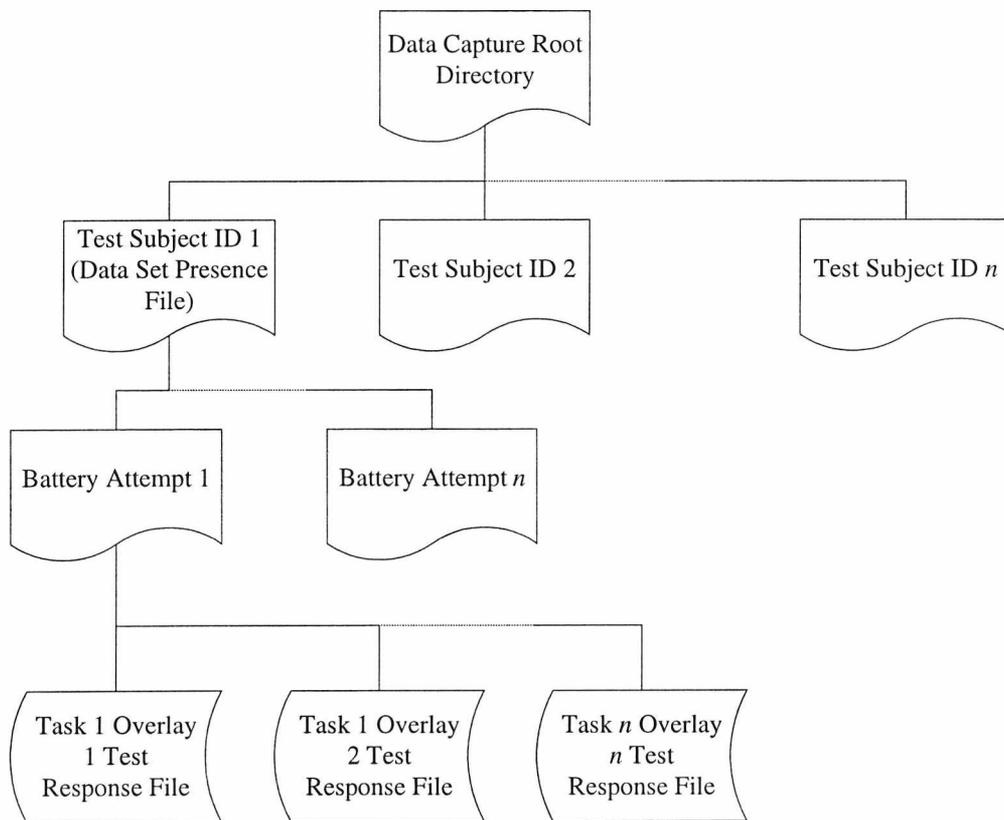


Figure 4.4 : Test response file directory structure

### 4.6.1 Offset Calculation

Upon completion of a test battery, offset calculations are computed to overcome any movement in the overlay positioning. This may have been caused by a misalignment of either the overlays within the paper clamps or the clamping system itself. Using an additional overlay containing two location crosses located in top left  $(x_1, y_1)$  and bottom right  $(x_2, y_2)$  corners (Figure 4.5), the test administrator positions the pen at these two locations. As the mask files have been calibrated with the overlay clamp in a default  $x$  and  $y$  position, any movement can be calculated from the *actual* (default) and *drawn* sets of coordinates. The offset for each individual test battery attempt is stored in the *offset data file* which list test ID, battery attempt number and  $x$  and  $y$  coordinate offset. The offset can be subtracted from the  $x$  and  $y$  coordinate data within the test battery attempt to align each test response data to the mask files.

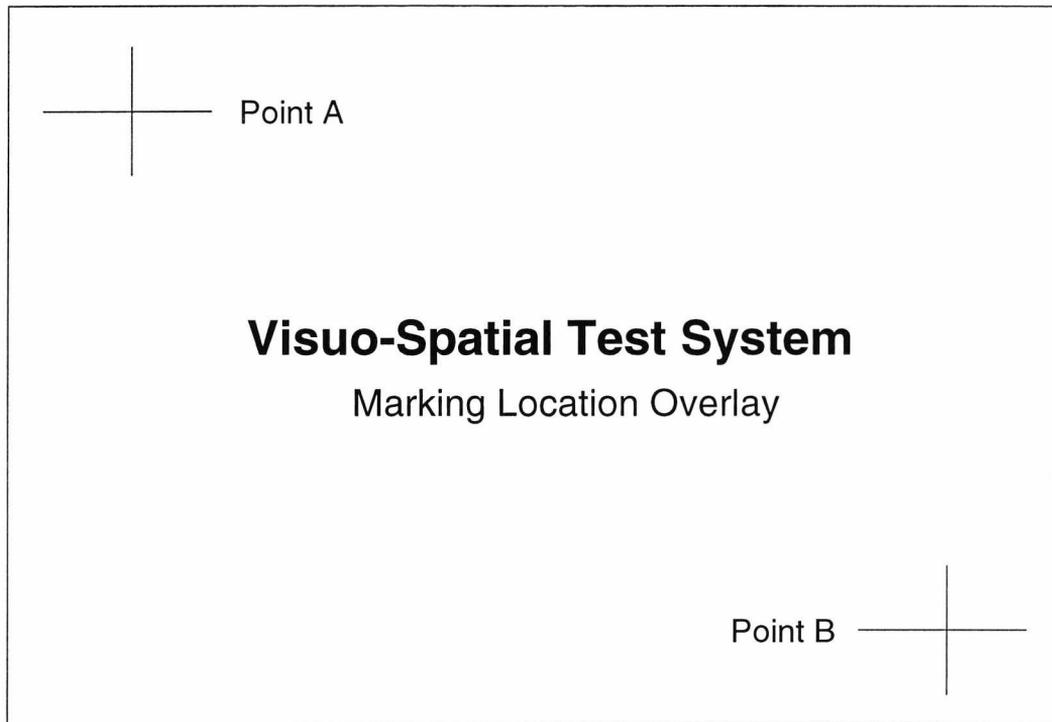


Figure 4.5 : Offset calibration overlay

The offset is calculated by the following formulae:

$$x_{offset} = \frac{x_1 error + x_2 error}{2} \quad y_{offset} = \frac{y_1 error + y_2 error}{2} \quad (4.1)$$

where:

$$\begin{aligned} x_1 error &= x_1 drawn - x_1 actual \\ x_2 error &= x_2 drawn - x_2 actual \\ y_1 error &= y_1 drawn - y_1 actual \\ y_2 error &= y_2 drawn - y_2 actual \end{aligned}$$

An examination of the offset data file over 155 battery attempts (5425 individual overlays) in a time period of 24 months without re-calibration shows that there is very little error variation within each axis. The mean  $x$  axis shift was 3.17 mm (standard deviation 7.24 mm) and mean  $y$  shift was 2.54 mm (standard deviation 6.88mm). This confirms the reliability of the paper fixing system. These errors can be incorporated as tolerances in location based assessments.

### 4.6.2 Quadratic Interpolation of Data

The number of data points extracted from the tablet (sampling rate) can be increased and quantised by interpolating between samples. This is particularly useful when using velocity and movement profile routines which perform optimally with a constant time-base at an increased frequency. Several methods exist for interpolating between data points including linear fitting, Fourier interpolation and the digital processing technique of oversampling. One of the widely used methods is to fit a quadratic to a series of data points  $x_1..x_n$  and calculate data points between 1 and n using the derived line equation [180]. Figure 4.6 shows a simple 2<sup>nd</sup> order polynomial fitted to three data points.

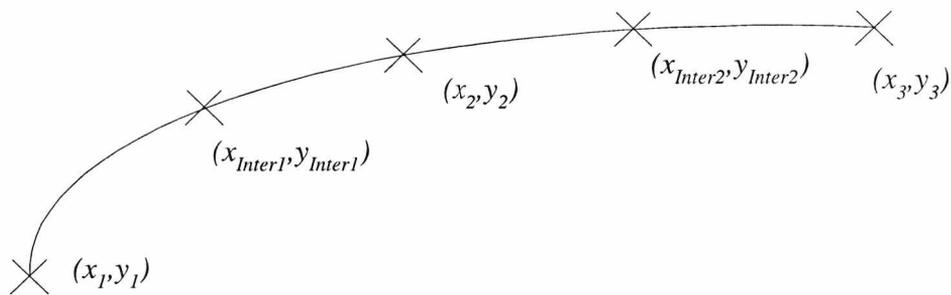


Figure 4.6 : Quadratic interpolation of line points

The points can be fitted to the line by the equations :

$$y_1 = p_2 x_1^2 + p_1 x_1 + p_0 \quad (4.2)$$

$$y_2 = p_2 x_2^2 + p_1 x_2 + p_0 \quad (4.3)$$

$$y_3 = p_2 x_3^2 + p_1 x_3 + p_0 \quad (4.4)$$

These three simultaneous equations can be solved by Gaussian Elimination resulting in the coefficients  $p_1$ ,  $p_2$  and  $p_3$  and hence the equation of the spline that fits all three data points.

The  $y$  coordinate of a 4<sup>th</sup> interpolated point with  $x$  coordinate  $x_{Inter}$  where  $x_1 < x_{Inter} < x_3$  is calculated by substituting the coefficients into the general 2<sup>nd</sup> order polynomial formula:

$$y_{Inter} = p_2 x_{Inter}^2 + p_1 x_{Inter} + p_0 \quad (4.5)$$

To double the sampling rate to 200Hz, the test response files are interpolated by taking three continuous timing points from the file and processing pen status data items ( $x$  location,  $y$  location, pressure etc.) individually. Visualising this graphically with relation to Figure 4.6, the individual data items ( $y$  axis) are plotted against the corresponding timestamp value ( $x$  axis). A quadratic polynomial is then computed between these point triplets. Any data item value can then be calculated using this polynomial at a given time value between the first and last selected timestamp forming the quadratic. This process is repeated for groups of three timestamps until the end of the data stream is reached resulting in a list of data values with a constant time-base and/or resampled frequency based on the existing feature data.

To double the existing sampling rate, three general solutions can be used to calculate the quadratic coefficients. If the middle of the three points is assigned as a zero time reference point ( $t=0$ ) and if a constant time-base is maintained between the points, then we can arbitrarily assign the first of the point to occur at  $t=-1$  and the last of the three points to occur at  $t=1$ . Substituting these times as  $x$  values into equations 4.2 to 4.4, the following general solutions are obtained :

$$p_0 = y_2 \quad (4.6)$$

$$p_1 = \frac{y_3 - y_1}{2} \quad (4.7)$$

$$p_2 = \frac{y_1 - 2y_2 + y_3}{2} \quad (4.8)$$

An interpolated doubled sample rate is obtained by finding the data item ( $y$ ) value at times  $t=-0.5$  and  $t=0.5$ . Figure 4.7 shows how the data stream sampled at 100Hz (a) is smoothed by

interpolating the  $x$  and  $y$  coordinate points separately to a new frequency of 200Hz (b). Data is taken from a square drawing task. The increased number of samples for the coordinate positions can clearly be seen to follow the existing data contour.

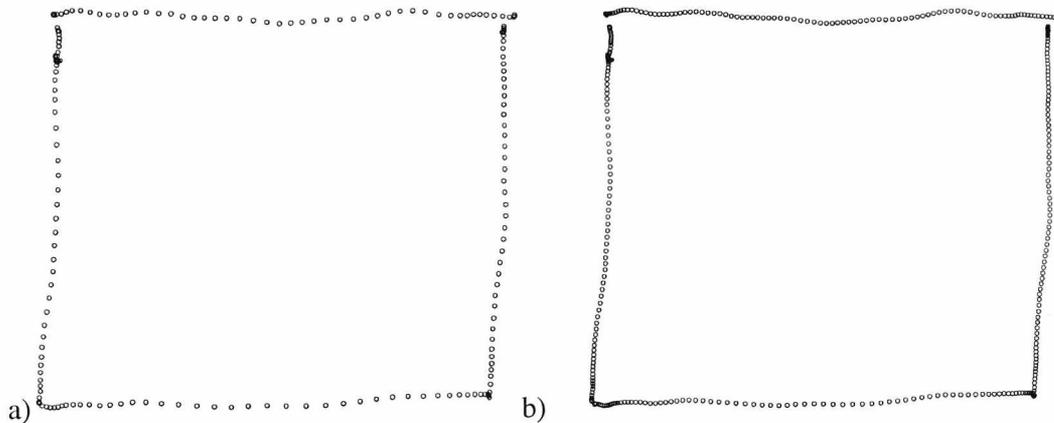


Figure 4.7 : Interpolation example a) original data (100Hz sampling rate), b) Data interpolated to 200 Hz

### 4.6.3 Gaussian Low-Pass Filtering

As stated in Section 4.2.2., normal handwriting typically has a maximum frequency in displacement movement of between 20 Hz and 13.6 Hz. Applying a low-pass filter with a cut-off defined at the maximum displacement frequency to individual data items (e.g. separate  $x$  and  $y$  coordinates) will eradicate any noise that is inherent to the output of the graphics tablet and other high frequency components of the data that are not constituent of a handwritten response (spatial errors). This will allow a reliable analysis of the handwritten data to be obtained. The Wacom tablet used in the research has an intrinsic fault that the tilt of the pen causes a modification to the  $x$  and  $y$  coordinate data including the condition when the pen is stationary (or consistency is maintained along a specific axis).

The simplest method for implementing a low-pass filter is to apply a window based *spatial convolution filter* to the data. Low-pass filtering of the data can be achieved using a *mask* loaded with a Gaussian profile. The profile shown in Figure 4.8 is obtained from Equation 4.9. Apart from the distance from the mean of the distribution ( $x=0$ ), the other variable is the

standard deviation of the distribution - increasing this value widens the spread of the distribution.

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-x^2}{2\sigma^2}} \quad (4.9)$$

where :

$\sigma$  = standard deviation of distribution

$x$  = distance from mean of the distribution (centre point)

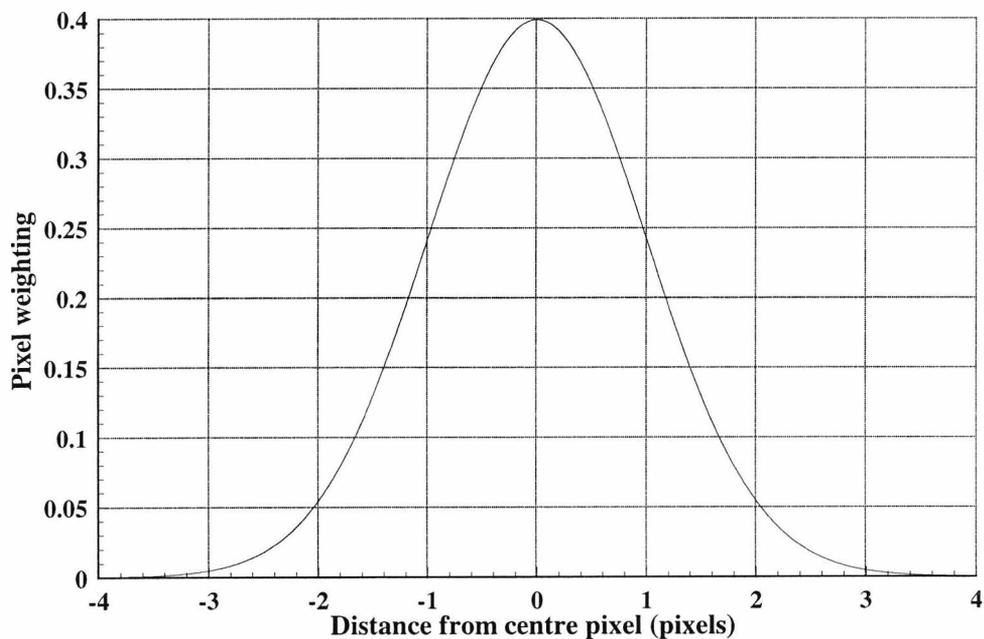


Figure 4.8 : Gaussian response profile ( $\sigma = 1$ )

Masks (in any dimension) can be generated from a calculated Gaussian. Figure 4.9 shows a one dimensional mask of width 2 (that is 2 data elements each side of the centre element). The mask is loaded with a profile with standard deviation=1. Using the mask defined in Figure 4.9, which operates on a single data stream (for example, on the pen  $x$  values), a

spatial convolution process is performed by multiplying the current value under investigation and two values each side within the stream by the values in the mask. This value is then divided by the sum of the mask (in this case 10) to produce the Gaussian weighted result. The mask is then shifted to the next item in the stream.

$$\frac{1}{10} \times \begin{array}{|c|c|c|c|c|} \hline 0.5 & 2.5 & 4 & 2.5 & 0.5 \\ \hline \end{array}$$

Figure 4.9 : One dimensional Gaussian convolution mask (width = 2,  $\sigma = 1$ )

The standard one dimensional convolution process is described by the function :

$$F(x) = \sum_{i=1}^j f(x + (i - (w + 1)))h(i) \quad (4.11)$$

where :

$w$  = width of mask

$j$  =  $(w \times 2) + 1$

$h [1..j]$  = mask elements

$f[x-j..x+j]$  = data stream elements contained within mask

This method, providing a smoothing effect on the data, is very simple to implement and computationally efficient. However, the response frequency of the filter cannot be controlled as accurately as operating in the frequency domain of the data. In this second implementation method, the data is transformed into the frequency domain by Fourier transform, filtered and then inversely transformed to recreate the original data stream with the filtered components removed. This method is computationally exhaustive involving the two Fourier transforms (normal and inverse) and a filtering process. As the data response file is typically in the range of 3000 to 6000 data items then each Fourier transform would be very slow.

The cut-off frequency of the convolution mask implementation and consequently the amount of smoothing to the data stream is controlled by the standard deviation of the Gaussian values contained within the convolution mask and the width of the mask itself. Empirical results show that as the width of the mask is increased and as smaller values of standard deviation of the Gaussian are implemented, the greater the cut-off frequency of the filter and hence the less smooth the data stream. An experiment was undertaken to numerically identify the relationship between these variables and thus increase the accuracy of the cut-off frequency of an implemented convolution mask. A random waveform sampled at 100Hz was filtered using a range of one dimensional convolution mask widths (between 1 to 10 data items wide) loaded with Gaussian profiles with standard deviations between 0.5 and 3. The resulting data stream was then transformed into the frequency domain using a discrete Fourier transform. The highest frequency component of the data contained within the frequency spectrum of the filtered stream was recorded as the corresponding cut-off frequency of the low-pass filter. Figure 4.10 shows the results of this study.

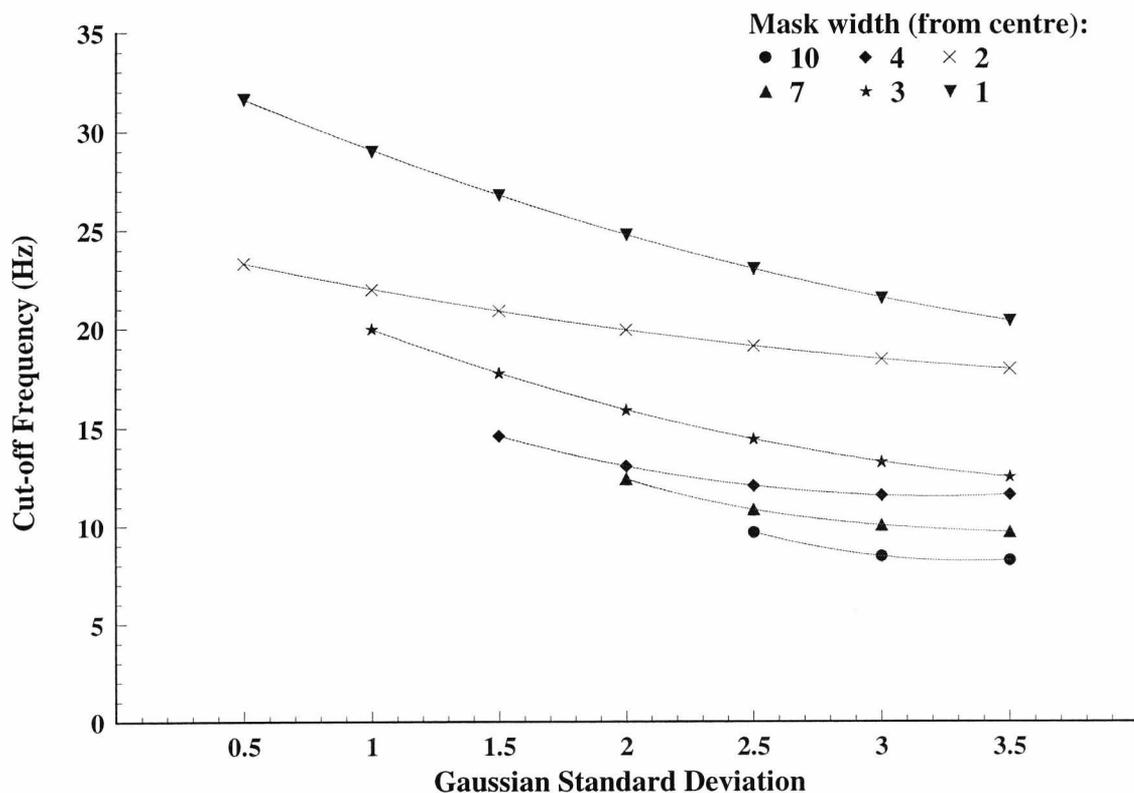
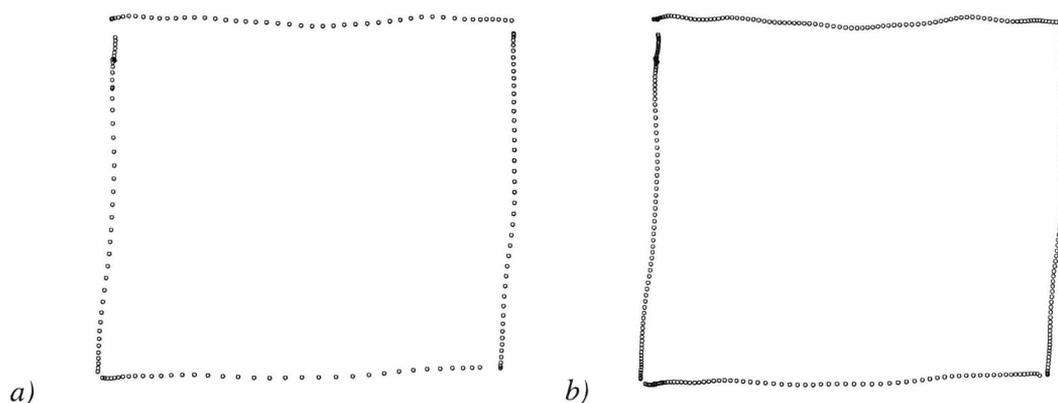


Figure 4.10 : Gaussian low-pass cut-off frequencies

With reference to the profile contained in Figure 4.8, it can be observed that values at a distance greater than  $3\sigma$  from the centre of the distribution are negligible to the output of the mask, therefore larger mask widths require higher standard deviations to enable values further from the centre of the mask to have a noticeable effect on the mask result. Because of this, the results from masks with a width greater than 3 times the standard deviation of the loaded Gaussian profile are ignored.

From these results a mask can be constructed to implement the required low-pass cut-off frequency of 20Hz (width = 3,  $\sigma=1$ ). Figure 4.11(a) shows the output of the designed Gaussian filter on the data presented in Figure 4.7. The  $x$  and  $y$  coordinate data streams were filtered separately. By comparing the two plots, the smoothing effect of the Gaussian filter can be seen. Figure 4.11(b) shows the combination of two pre-processing operations with the Gaussian smoothed data stream interpolated to a frequency of 200 Hz. This combination filters out noise and provides a cleaner data stream for feature extraction.



*Figure 4.11 : Gaussian filtration example a) Low-pass Gaussian filtering of drawing shown in Figure 4.7(a), b) Interpolated Gaussian filtering of drawing data*

## 4.7 Feature Extraction

Features are extracted from the test response files by processing the stored list of data items. Chapter 5 provides a full explanation of the range of features extracted from each task but this section contains two simple examples of the extraction process common for all features.

In particular, these examples represent two dynamic features which are not normally obtainable from the completed test overlays.

Figure 4.12 shows the flow chart for the calculation of overall drawing time of a test response. Coordinate data is read sequentially from the response file until pen contact is detected on the tablet. The tablet will return coordinates of the pen up to a distance of 5mm from the surface. When the button value indicates pen surface contact for the first time within the test response file, the drawing start time will be defined by the corresponding packet timestamp. The time at which the pen is finally removed from the tablet surface (button value = 0) will indicate at the drawing stop time. The overall drawing time is simply a matter of subtracting the starting from the stop timestamps.

Constructional order can be extracted from a test response file by segmenting and analysing rate of change in a pen data item. For the drawing tasks, side drawing order can be obtained from the change in  $x$  and  $y$  pen coordinates extracted from the segmented sections. Figure 4.13 show the  $x$  and  $y$  coordinate profiles of test response data items from a square drawing from memory task. Five segments can clearly be identified within the drawing and timings for each stroke can be extracted from these profiles. The constructional order is extracted in three stages:

1. The *axis* of movement is calculated (horizontal or vertical). Within each segment, does the  $x$  or the  $y$  coordinate have a greater rate of change. This feature is easier to detect within the drawing of the square where all the sides are parallel to axes. For a 45 degree diagonal line the rate of change of both coordinate data items are identical. In the given example, drawing movement in the first segment is in the  $y$  axis, followed by a movement in the  $x$  axis in the second segment.
2. Having identified the axis of movement, the *direction* of movement (left or right, up or down) is obtained from the start and end coordinates within the segment. With the zero reference point for both axes being in the top left hand corner, pen movement is down in the first segment and to the right in the second in the given example.
3. The final stage of extraction, *segment labelling*, concurrent analysis of both movement profiles indicates whether a vertical movement constitutes a left or a right side of the

drawn square. Again, with the example data, the first segment is to the left of the drawing and the second segment forms the lower side of the square.

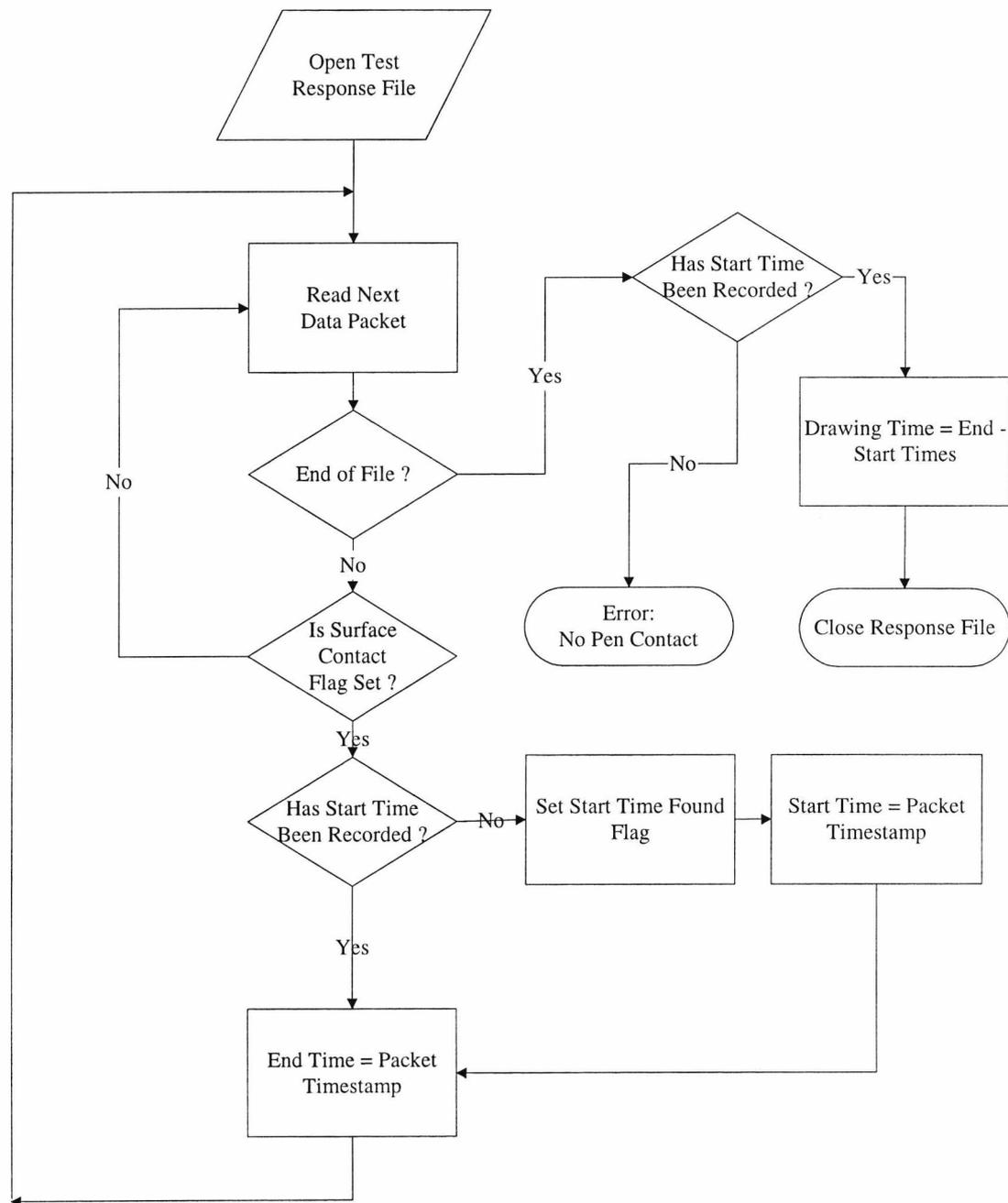


Figure 4.12 : Total drawing time feature extraction flow chart

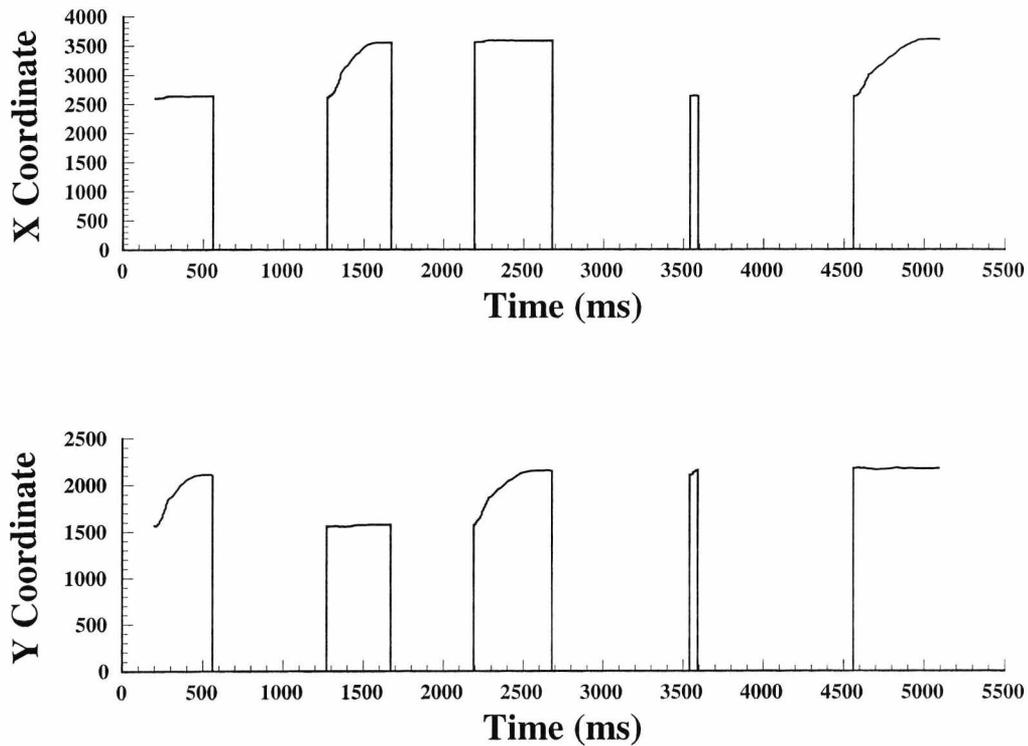
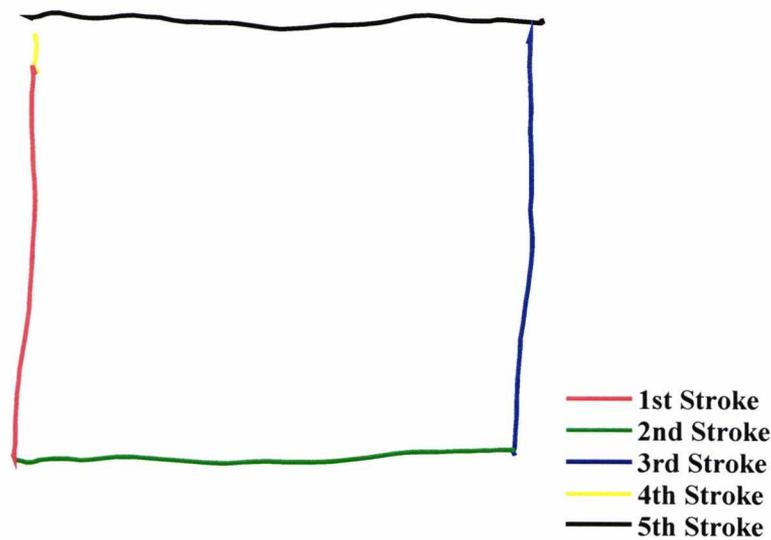


Figure 4.13 : X and Y coordinate profiles from a drawing from memory square task

An alternative graphical representation of the drawing is presented in Figure 4.14 using colour codes to represent individual segments. The drawing order and positioning is clearly visible from this representation.

## 4.8 Summary

This chapter has identified the central issues concerning the automatic capture of handwritten data. The movement characteristics of a handwritten response have been defined and a range of input devices explored with reference to both these characteristics and also the requirements of a hospital-based environment. The pen status data items reported by a graphics tablet were summarised, leading to a discussion about the types of features that can be extracted from this data.



*Figure 4.14 : Stroke order in a square drawing from memory task*

The Wacom WD 1212 tablet was examined in detail and the system infrastructure for the capture, feature extraction and result reporting explained. The accuracy of the system for attaching test overlays to the tablet surface was verified by analysing the  $x$  and  $y$  axis shift over 5425 overlays in a 24 month period without re-calibration. Performance was found to be satisfactory. Three pre-processing operations were investigated to overcome some of the defined inherent problems with data capture using a graphics tablet. These operations reduce noise and increase the accuracy of the signal prior to feature extraction. Methodologies for feature extraction were given and serve as a template for other extraction routines.

The defined system presents an infrastructure from which a range of existing and novel features can be extracted within bounds of handwriting movement specification and which can be used within the defined clinical environment and test subject grouping. Having defined the experimental infrastructure, the individual tasks and features can be defined and implemented.

## Chapter 5

# Task Definitions and Feature Extraction

### 5.1 Introduction

This chapter provides a detailed description of the tasks implemented as part of the computer-based assessment battery and defines the extractable performance features from each task. Features are described algorithmically and, where applicable, the assessment criteria are defined, resulting in a stable and objective set of marking rules for application to all test responses.

Designed to maximise potential differences between a neglect and a stroke control population, the testing procedures have also been devised on an experimental basis to test hypotheses concerning dynamic feature performance in relation to spatial awareness.

### 5.2 Test Battery

A total of 35 overlays are used in the test battery. The graphics tablet on to which the overlays are fixed is positioned on a table directly in front of the test subject, ensuring a normal writing position. All overlays are of size 296 by 209 mm (A4) in a landscape orientation. The overlays are fixed individually onto the surface of the tablet by the test administrator and are removed following completion; the next overlay in the battery is then placed on the tablet surface. A script containing verbal instructions given to each test subject is used by the test administrator to ensure that a uniform testing procedure is maintained. A copy of the test script, devised by a trained Occupational Therapist, is included as Appendix B. Capture of drawing response and general computer interaction is controlled by the administrator, ensuring that the patient has no direct contact with the computer and thus preventing any performance modification due to an unfamiliar testing environment. The test infrastructure appears to the patient to be identical to that used for conventional pencil and paper tests.

The task order within the test battery is: Point Location, Cancellation Tasks (OX1, OX2, Albert's, Star), Figure Completion, Figure Copying, Drawing from Memory and Line Bisection. In addition, a series of Movement Profile tasks were added to the battery to provide kinematic analysis of test performance.

No time limit is imposed for any of the tasks. However, aborting any or all of the overlays is an option left to the discretion of the administrator.

### 5.2.1 Point Location Task

The point location task is used as a simple screening test at the start of the test battery. A total of four overlays are used containing two black dots of diameter 15mm. One of the dots (*the starting dot*) is located in the same position, at the bottom centre, in all four overlays (33 mm from the bottom and 146 mm from the left of the overlay). The other *target* dot is positioned in one of the four quadrants of the overlay, thus over the four test overlays all areas of the immediate visual field are tested. Table 5.1 details the position of the target dot. Figure 5.1 shows overlay 1 with the target dot in the top left hand corner. The starting dot is positioned in the lower centre of the overlay.

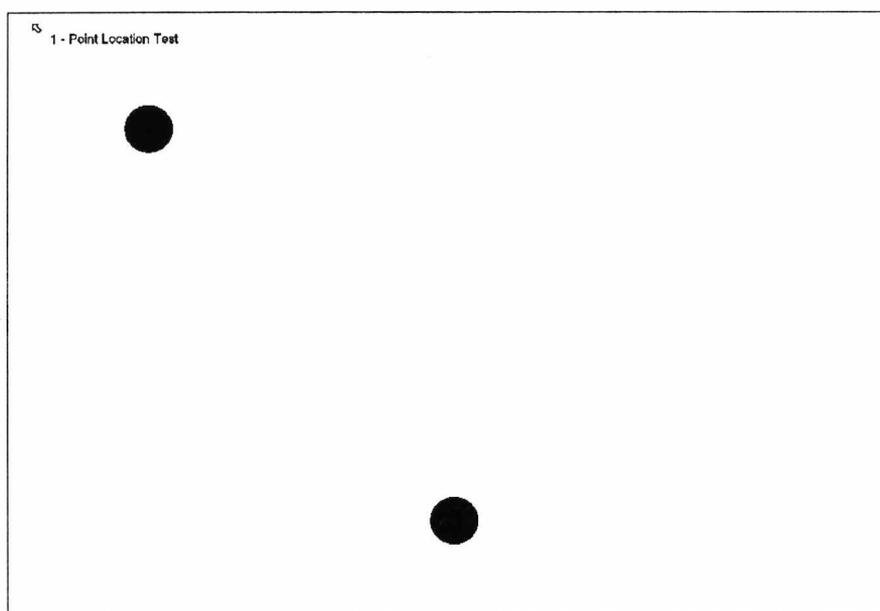


Figure 5.1 : Point location task overlay 1

Overlay Number	Quadrant	x Position from Left Hand Edge	y Position from Bottom Edge
1	Top Left	45 mm	165 mm
2	Top Right	220 mm	127 mm
3	Bottom Left	49 mm	44 mm
4	Bottom Right	239 mm	60 mm

Table 5.1 : Position of target dot in point location task

The test subject is presented with the individual overlays in sequence and is required to perform the following events to complete the task:

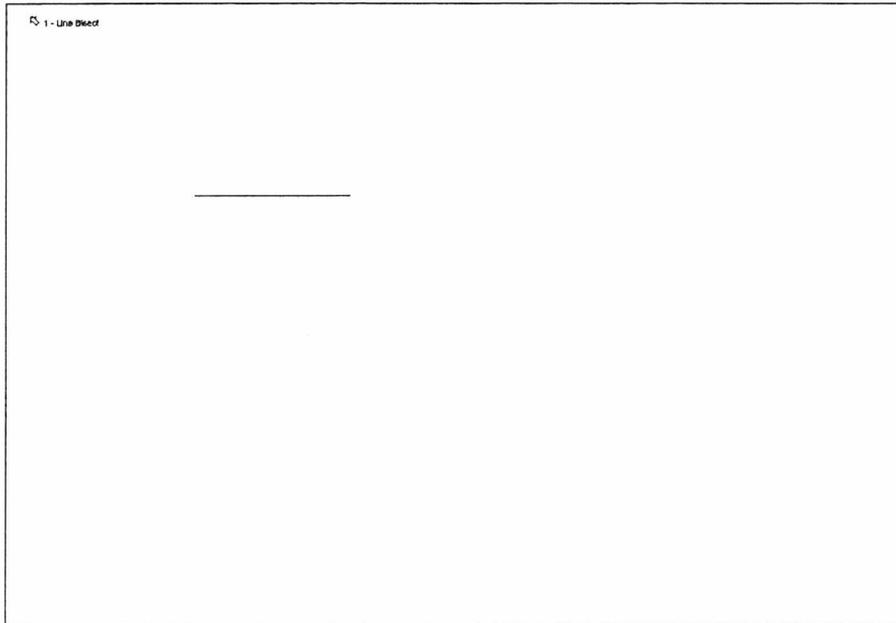
- Move the pen to the starting dot and place the pen on the dot.
- Remove the pen.
- Move to the target dot and place the pen on this dot.

### 5.2.2 Line Bisection Task

The line bisection task comprises eight overlays each containing a single horizontal *target line* which the test subject has to locate and bisect at the midpoint. The line length is 50mm for the first four overlays and 140mm for the second set of four. Figure 5.2 shows the first overlay from this task, with a short (50mm) line positioned in the top left hand quadrant. As with the Point Location Task, the line is positioned with its midpoint within each quadrant over a set of four overlays. Table 5.2 details the position and location of each line.

Overlay Number	Quadrant	Length	x Position from Left Hand Edge	y Position from Bottom Edge
1	Top Left	50 mm	85 mm	142 mm
2	Top Right	50 mm	254 mm	155 mm
3	Bottom Left	50 mm	85 mm	54 mm
4	Bottom Right	50 mm	196 mm	66mm
5	Top Left	140 mm	107 mm	141 mm
6	Top Right	140 mm	201 mm	152 mm
7	Bottom Left	140 mm	85 mm	44 mm
8	Bottom Right	140 mm	207 mm	77 mm

Table 5.2 : Position of target line in line bisection task



*Figure 5.2 : Line bisection task overlay 1*

### 5.2.3 Cancellation Tasks

Three cancellation tasks have been implemented in the computer-based test battery. The first task is a standard Albert's Cancellation Task [49] which has been used for many years as a neuropsychological test of neglect (Section 2.3.2). The test subject is presented with a single overlay containing 40 lines of length 25mm arranged at a random orientation and in a pseudo-random grid formation measuring 222 mm x 181 mm (Figure 2.5). The test subject is required to cancel all of the lines on the overlay. The test response is assessed by counting the number of cancellations, ignoring the centre vertical line of four targets which are used to demonstrate the cancellation process.

The other two tasks (OX1 and OX2) use a series of targets ('O' characters) and distractors ('X' characters) arranged in a pseudo-random grid formation (Figures 5.3 and 5.4). The test subject is required to cancel all the O characters on the overlay. In the first task, there are 12 of each type of character, of dimensions 7 x 7 mm, arranged in a pseudo-random grid of dimension 198 x 107 mm. The second overlay has 16 targets and distractors, also of dimension 7 x 7 mm, again arranged in a pseudo-random grid, with larger dimensions of 214 x 158 mm. Only 12 out of the 16 targets are assessed for the correct number of cancellations, resulting in 3 target and 3 distractors characters per quadrant. The boxed area in Figure 5.4

details the characters which are not assessed in calculating the correct number of cancellations. Studies [167] have shown that by fixing the positions of targets in linear horizontal and vertical rows, test subjects are able to form methodical axis-based cancellation strategies (for example, cancelling all targets in a particular row and then moving on to the next row) easier than if the targets were arranged in 'random' positions. Categorising the cancellation strategy is therefore simplified if a lesser number of strategies are used across the test population. Linear (or grid) positioning of the targets forces the test subject to scan and cancel along vertical and horizontal axes reducing the number of output sequences. As with the Albert's task, the basic analysis of this test is to count the number of cancellations on the overlay.

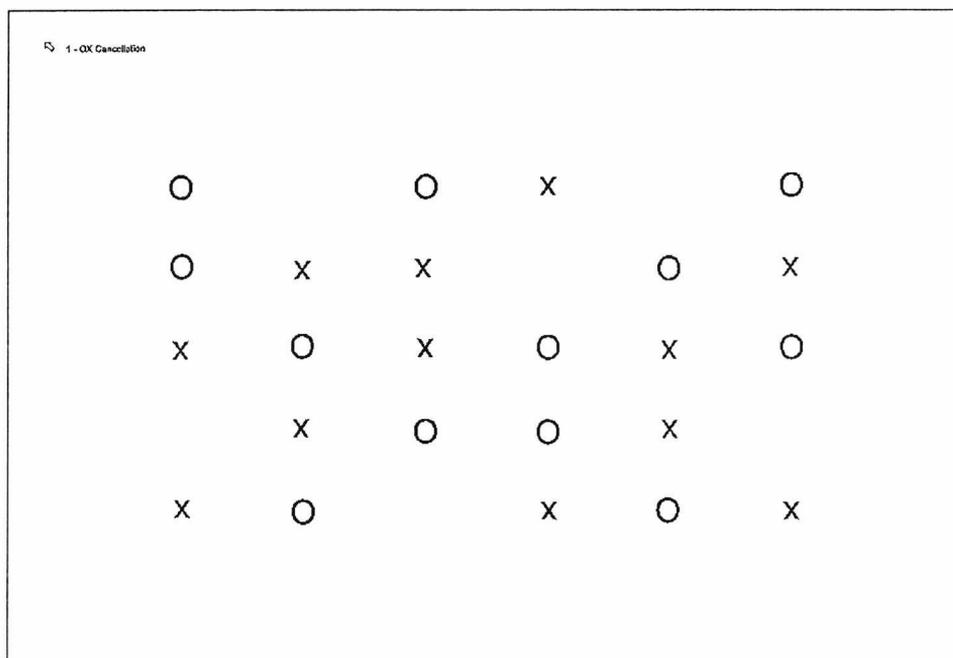


Figure 5.3 : OX cancellation task overlay 1

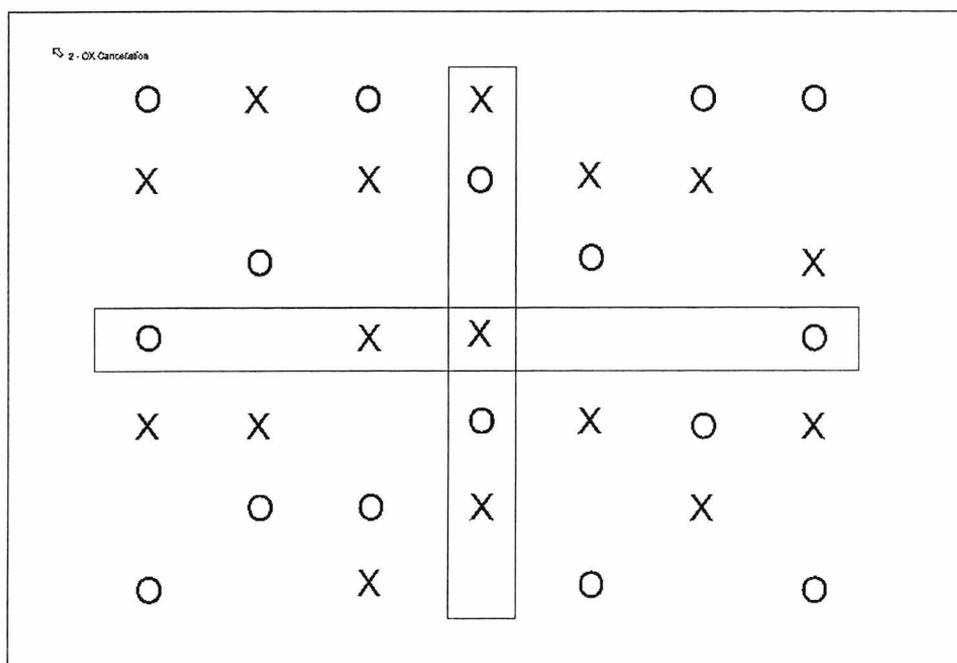


Figure 5.4 : OX cancellation task overlay 2 detailing targets which are not assessed in the number of cancellations

#### 5.2.4 Figure Completion Tasks

The figure completion task consists of six overlays each containing half of a simple representational shape (Figure 5.5) split vertically. The shapes are located so that the vertical split is positioned at the horizontal centre of each overlay. Three different representational shapes of increasing complexity are used: a diamond, a man and a house. The test subject is first presented with the left hand side of the image which requires completion (i.e. to draw the mirror image of the shape) to the right hand side of the overlay. The second in the pair of overlays, which requires copying of the same shape to the left of the overlay, is then presented. Table 5.3 records the maximum width and height of the half images. As a vertical mirror image is required to be drawn, this is also the size of drawing expected from each subject.

Analysis of neglect concentrates on the performance differences between copying *into* the affected visual field (the left side in right CVA patients) and when copying *from* this visual field.

Shape	Max. width (mm)	Max. height (mm)
Diamond	50	95
Man	36	83
House	62	83

Table 5.3 : Figure completion half images sizes

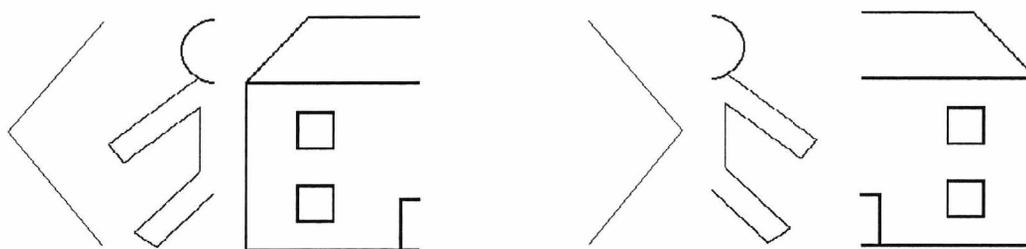


Figure 5.5 : Figure completion representational drawings

### 5.2.5 Figure Drawing Tasks

The figure drawing tasks involve the copying and drawing from memory of a series of simple geometric shapes. The four shapes used in these tasks are shown in Figure 5.6. Two of these shapes (the cross and the cube) were used in a copying trial by Warrington, James and Kinsbourne [85] and were found to give the best separation between neglect and non-neglect right CVA subjects based on assessed performance characteristics.

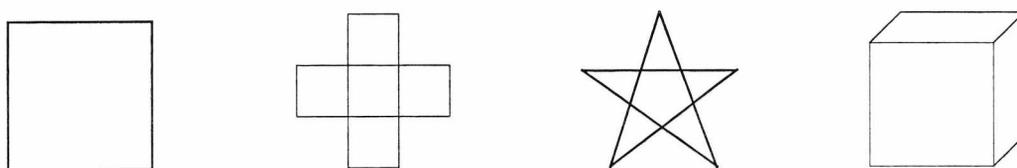


Figure 5.6 : Figure copying and drawing from memory models

The two other shapes were selected for their simplicity, signified by the low number of sides (the square) and their relative complexity to the other shapes used in the battery (the five pointed star). The star shape presents many dynamic feature possibilities as it is possible to construct without pen removal. As most subjects are familiar with the standard two-triangle six pointed star, the shape requires detailed copying analysis.

In the first of the drawing tasks, *figure copying*, the shapes are printed individually in the top horizontal centre of four separate overlays. The test subject is required to copy the shape directly below the printed image. The order in which the shapes are presented (square, cross, star and cube) remains constant for all attempts at this task. The sizes of the model/target image are given in Table 5.4.

Shape	Max. width (mm)	Max. height (mm)
Square	21	21
Cross	42	42
Star	22	22
Cube	36	36

Table 5.4 : Figure copying shape sizes

The second drawing task, *drawing from memory*, uses only two of these shapes, the square and the cube. The test subject is asked to draw these shapes on separate blank overlay sheets without any prompting or copying model. The choice of image represent the extremes of complexity of the shapes used in the figure copying task and enables the assessment of dimensionality within a drawing (the cube is particularly interesting as it embraces three dimensional perspective properties).

### 5.2.6 Drawing Profile Tasks

The final set of tasks is a series of line drawing assessments involving visual discrimination and sequence processing. Using 8 overlays, the drawing profile tasks enable a more detailed understanding of the movement dynamics involved in drawing images. The overlays can be divided into 4 distinct task groupings.

The first three overlays contain a series of dots at each side of the page. The dots are 10mm in diameter, located 57 mm from each vertical edge of the overlay and are separated vertically by 46 mm. Figure 5.7 shows the first of this group of three overlays. Starting at the bottom right hand side of the overlay (selected because the right hand side is least affected by the patient's neglect), the test subject must move the pen across the page to the lowest left hand side dot. Having located this target, a movement is made back across the page to the second lowest target. This 'zigzag' pattern is repeated until all dots have been visited, finishing in the top right hand corner. By asking the test subject to draw a line the length (or width) of the overlay, the pen is in contact with the tablet for a greater time, thus producing a clearer movement feature dynamic, such as velocity and acceleration, whilst also allowing the extraction of the various kinematic-based timing measurement, such as acceleration and deceleration phase timings. Repeated movement back and forth across the width of the page also means that differences can be observed as a neglect subject moves the pen in and out of the neglected field.

The two other overlays contain variations of this task (Figures 5.8 and 5.9), introducing distractors (10 x 10 mm squares) placed amongst the circles. The test subject must discriminate between the targets and only move the pen to the circles. Whilst the order of targets and distractors is uniform on each side of the second overlay, the order is randomised in the third overlay.

The second set of two overlays require the test subject to join numbered dots in sequence to form a square. The dots are located in the four quadrants of the overlay, 30 mm from each edge, so analysis can assess movement in and out of the affected vertical visual field as well as on a horizontal basis. Figure 5.10 shows the first of these square drawing overlays. In this overlay the dots are numbered in an anti-clockwise direction from the bottom right corner. The second overlay reverses this ordering whilst maintaining the positions of the dots.

Another square drawing task is implemented on overlays 6 and 7. The dots for this task are located 64 mm apart in a single side of the visual field. No drawing sequence is specified for this task. The first of the overlays has the dots printed in test subject's right visual field, swapping to the left in overlay 7 (Figure 5.11).

The final overlay in this task sequence contains twelve dots arranged at a 70 mm radius around a single central dot in a clock-face configuration (Figure 5.12). The test subject is required to draw from the central dot to each of the outlying dots. The pen is then picked up

and moved back to the central dot. Analysis of this task provides data on the sequence in which the dots were visited and also the timings and movement dynamics at the various positions within the overlay.

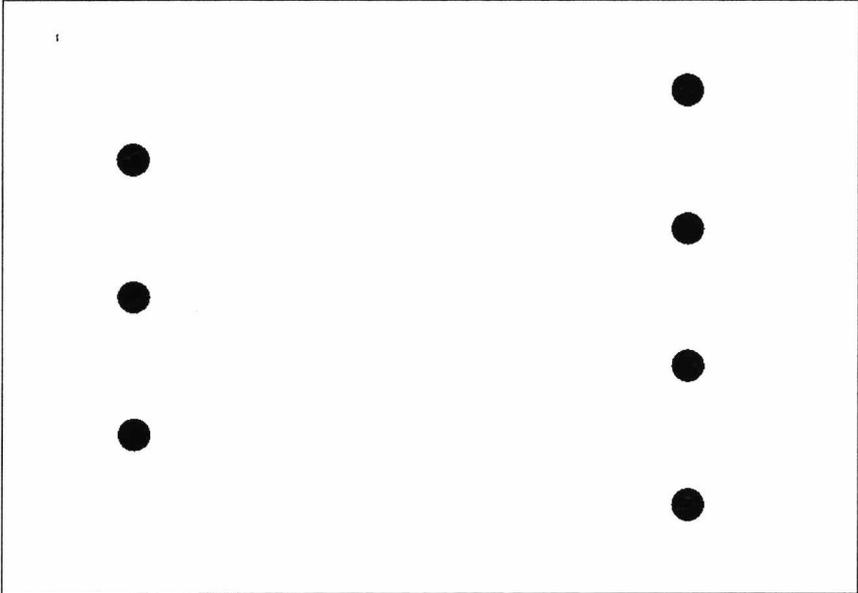


Figure 5.7 : Line drawing task

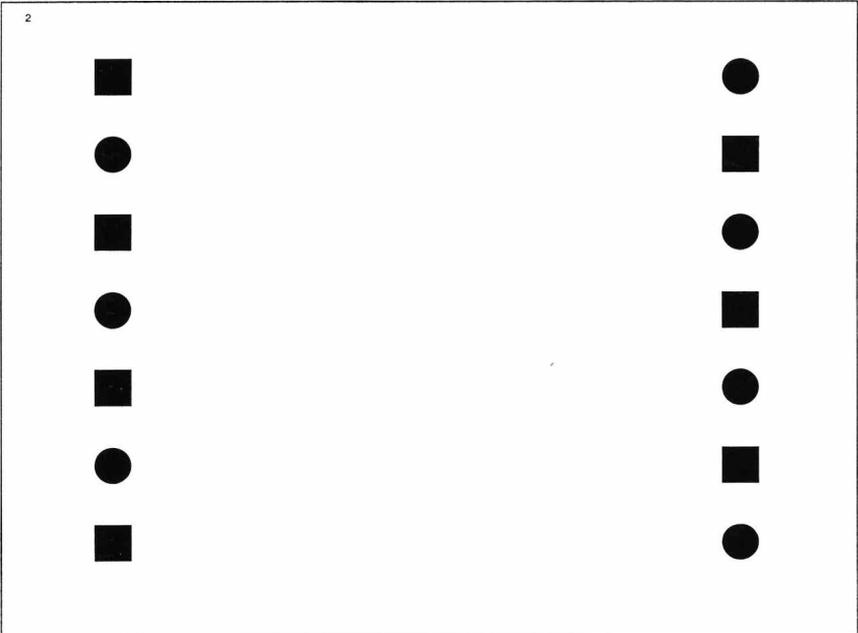


Figure 5.8 : Line drawing task with distractors

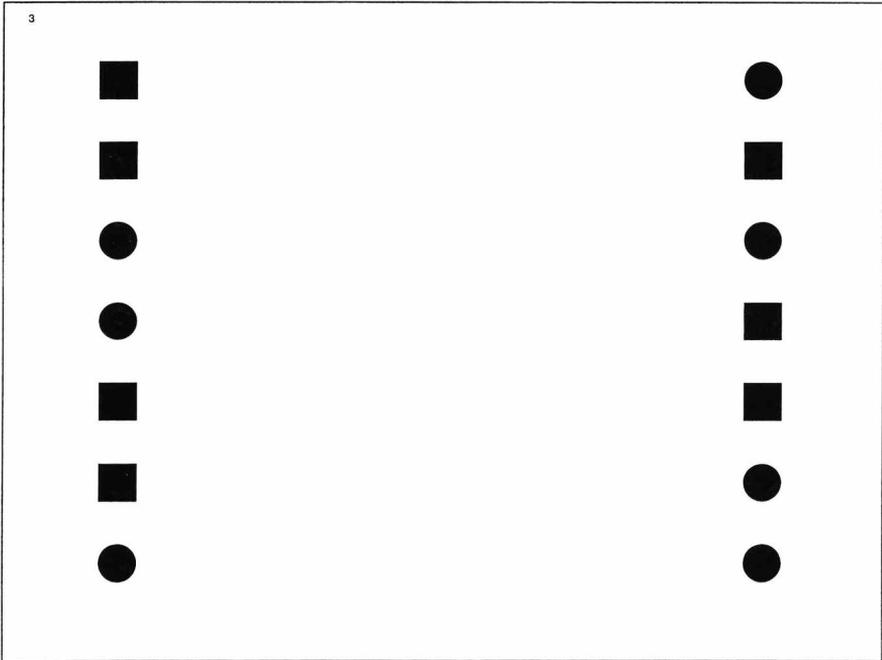


Figure 5.9 : Line drawing tasks with distractors and pseudo-random ordering

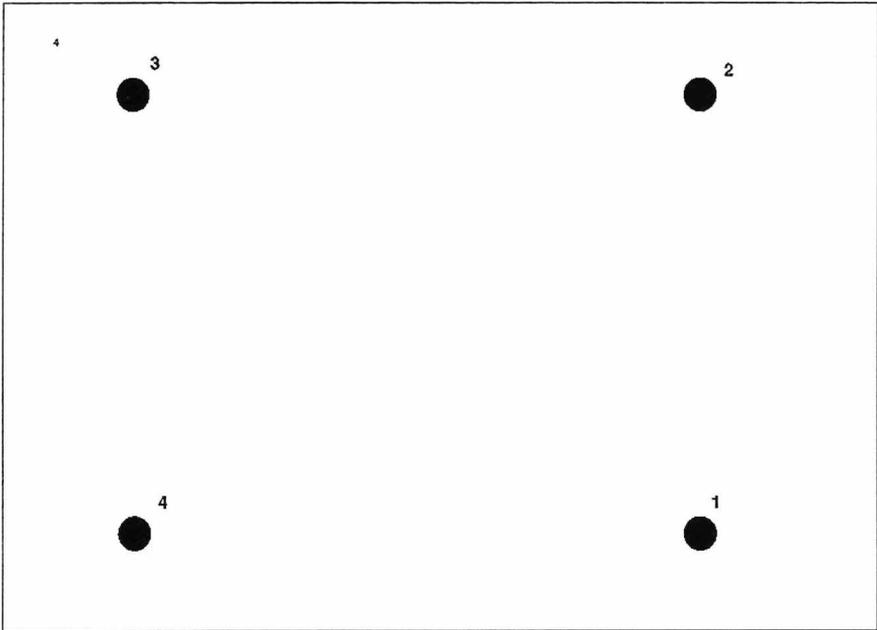
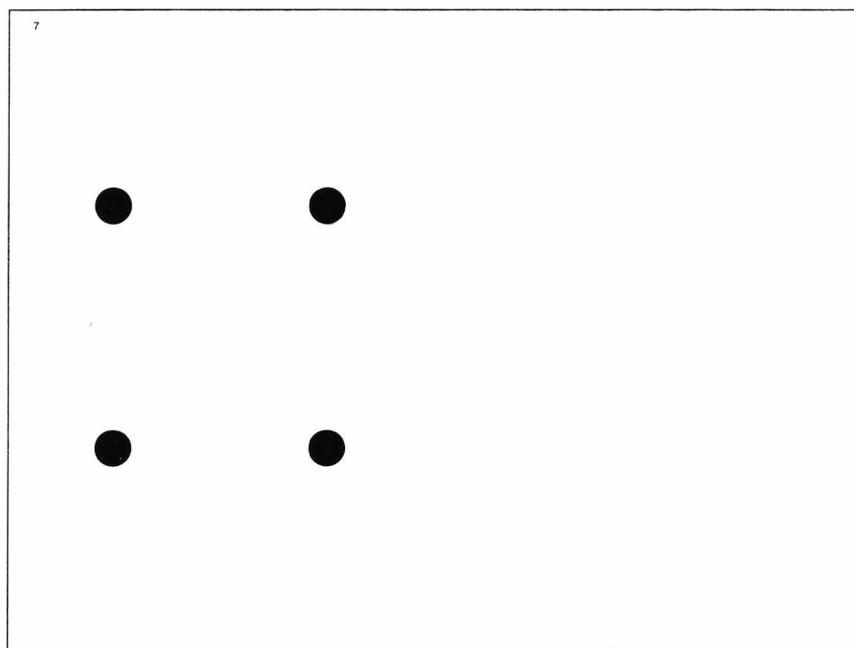
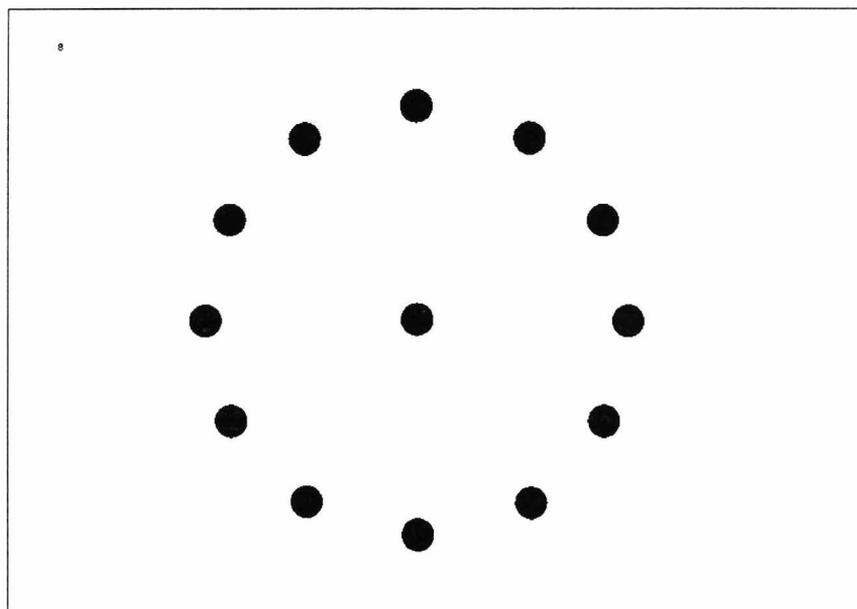


Figure 5.10 : Dot joining square task



*Figure 5.11 : Visual field square drawing task*



*Figure 5.12 : Clock-face line drawing task*

## 5.3 Performance Features

This section defines the features extracted from each drawing or task response undertaken by a test subject. Section 4.7 described the procedure for extracting features from a stream of raw coordinate and pen status data whereby calculations are made by reading sequentially through the data stream and extracting features from the test response file. All time-based features utilise the timestamping which is assigned (or extrapolated) to the pen status data during sampling. Each feature has been implemented as a separate function which can be applied to a standard test response file. Obviously, applying a task-specific feature function to a response file from the incorrect task causes an erroneous result.

The features that are documented in this section have been arranged according to the tasks to which they are applied. Several features are generic to all tasks (such as overall drawing time, number of pen lifts within drawing etc.) and many features are common across the analysis of all four of the drawing based tasks. These features are described in Sections 5.3.1 and 5.3.5 respectively. Features that are specific to an individual drawing task, such as number of cancellations on an overlay, are documented separately at the end of this section.

### 5.3.1 Generic Task Features

The features described in this section can be extracted from any test response file captured using the computer-based assessment system. This means that they are suitable for use outside the immediate scope of this project.

#### 5.3.1.1 Overall Task Execution Time

The overall task execution time indicates the real time taken to complete a task. As task complexity increases the overall time taken, this measure can only be used as a comparison between pairs or groups of test subjects *on a single task*. The time is calculated using the pen button flag contained in the test response file. This flag is set to >0 when the pen is placed on the tablet surface (i.e. a drawing is made), so the overall execution time can be calculated thus:

$$\text{Overall Execution Time} = t_{end} - t_{start} \quad (5.1)$$

where :

$t_{end}$  = last time when pen tip was on graphics tablet surface in test response file.

$t_{start}$  = time of initial pen tip contact.

### 5.3.1.2 Pen Contact/Pen Movement Ratio

This ratio is again obtained by monitoring the pen button status flag within the test response file. By individually accumulating the times within the response file when the pen is on and removed from the tablet surface, a ratio can be formed. A result over 1.0, indicates a greater time spent with the pen off the tablet which in turn indicates a larger planning phase of the task completion. Pen removal from the tablet surface usually occurs when the test subject is moving to the start of a new component.

The ratio can be defined:

$$\text{Pen Contact Ratio} = \frac{\text{Total time pen off tablet}}{\text{Total time pen on tablet}} \quad (5.2)$$

The time accumulation only occurs between the times  $t_{end}$  and  $t_{start}$  as defined in Section 5.3.1.1., so as not to include the variable time phases introduced by the test administrator arbitrarily starting and stopping the sampling procedure. This ensures that the ratio only accounts for timing within the *active period of drawing or response*.

### 5.3.1.3 Mean and Peak Pressure

The pressure on the tablet surface produced by the pen is returned in the range of 0 to 255. Accumulating the pen pressure value for every packet returned when the pen is on the tablet and then dividing the number of pen contact packets results in the mean pressure over the test response:

$$\text{Mean Pen Pressure} = \frac{\sum(p)}{n} \quad (5.3)$$

where:

$p$  = pressure values from a data packet when pen is on tablet (pen tip status  $\neq 0$ )

$n$  = number of packets when pen is on tablet

The maximum pen pressure recorded during the drawing process can be used as a performance feature. However, this suffers from a ceiling effect of the maximum pressure it is possible for the tablet to detect.

#### 5.3.1.4 Pen X/Y Tilt Standard Deviation

Pen tilt is recorded in both the horizontal ( $x$ ) and vertical ( $y$ ) axis. By computing the standard deviation of tilt values when the pen is on the tablet, the variation in posture throughout the drawing processes can be established.

The standard deviation of the pen  $x$ -axis tilt is calculated by :

$$x \text{ tilt } \sigma = \sqrt{\frac{\sum(xt)^2}{n}} \quad (5.4)$$

where:

$xt$  =  $x$  axis tilt value from a data packet when pen is on tablet (pen tip status  $\neq 0$ )

$n$  = number of packets when pen is on tablet

$y$  axis tilt is calculated using the same method.

### 5.3.1.5 Pen Lifts within Drawing

This is a measure of the number of times the pen was removed from the tablet *during* the drawing time (not including the final pen lift at the end of the drawing hence the deduction of a single occurrence within the calculation) and gives the number of movement segments within the drawing.

$$\text{Pen lifts within drawing} = \left( \sum_{t=0}^n [(Z(t-1) \rightarrow Z(t)) = (1 \rightarrow 0)] \right) - 1 \quad (5.5)$$

where:

$n$  = number of packets

$Z(t-1)$  = pen tip status at time  $t-1$

$Z(t)$  = pen tip status at time  $t$

$Z(t) = 0$  = pen tip not on tablet

$Z(t) = 1$  = pen tip on tablet

### 5.3.2 Point Location Features

The features extracted from the Point Location task are limited by the simplicity of the task. Time for pen movement between targets is calculated from the overall execution time defined in Section 5.3.1.1 indicating the processing time required for identification, location and movement to the target dot. The other two features obtainable from this task concern the accuracy of target location and a comparison between quadrant performance.

#### 5.3.2.1 Target Distance Error

Assessing if the test subject has located and hit the *target dot* is calculated by finding the Euclidean distance between the *drawn location* where the test subject positioned the pen and the *actual location* – the coordinates of the centre of the target dot stored in a model of each overlay. If the calculated distance is outside the radius of the dot (7.5mm) then the patient failed to hit the target. The same measure is used for calculating the error distance on the

starting dot. If both results are within the dot radius the test subject has successfully completed the overlay.

$$\text{Target Distance Error} = \sqrt{(x_{\text{drawn}} - x_{\text{actual}})^2 + (y_{\text{drawn}} - y_{\text{actual}})^2} \quad (5.6)$$

where :

$x_{\text{drawn}}$  = the  $x$  coordinate of the pen when placed on the tablet

$x_{\text{actual}}$  = the model centre  $x$  coordinate of the target

$y_{\text{drawn}}$  = the  $y$  coordinate of the pen when placed on the tablet

$y_{\text{actual}}$  = the model centre  $y$  coordinate of the target

### 5.3.2.2 Quadrant Comparison

Having extracted the timing and distance error for all four overlays, a comparison between feature measurements from the two left hand side and right hand side targets and, furthermore, on a quadrant basis, may provide performance discrimination between test subject groups. Expected performance from severe right CVA neglect subjects will be identified by the inability to locate (or locate with a large distance error) the left hand side target dots.

### 5.3.3 Line Bisection Features

As with the point location task, the number of features that can be extracted from the line bisection overlays is limited because of the simple task composition. Completion of the task only requires a single bisection mark to be made on the overlay. The task completion time is therefore insignificant as this feature is directly related to the length of bisection mark *drawn by the test subject*, which is not an important measure in assessing the outcome of the task. Because two different line lengths are used as bisection targets, an intra-subject measure can be established concerning performance differences between short and long lines, in addition to the quadrant positioning of the bisection line.

### 5.3.3.1 Deviation from Actual Midpoint

The main accuracy measure extracted from the line bisection task is to assess the distance between the test subject's midpoint estimate and the true midpoint coordinate. Figure 5.13 is an example of the deviation calculation from a hypothetical neglect response. By vectorising the result, the direction (left or right of the midpoint) of the error can be obtained. A neglect response should theoretically bisect the target line to the right of the midpoint.

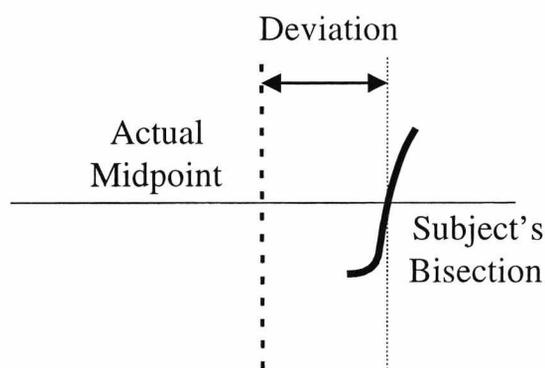


Figure 5.13 : Horizontal midpoint deviation calculation

The actual deviation from midpoint calculation is given by:

$$\text{Midpoint Deviation} = x_{\text{drawn}} - x_{\text{actual}} \quad (5.7)$$

where :

$x_{\text{drawn}}$  = test subject's  $x$  coordinate at model  $y$  coordinate

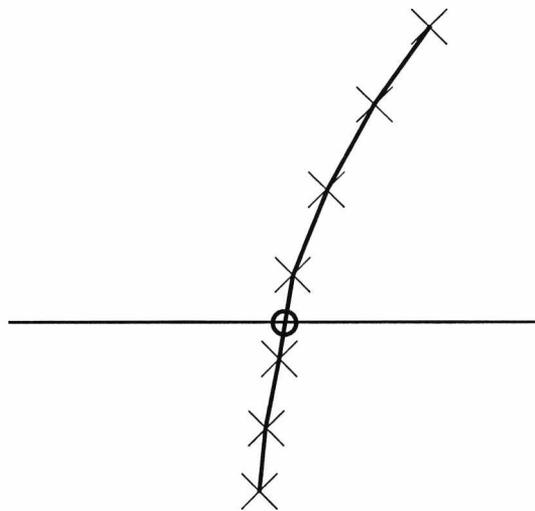
$x_{\text{actual}}$  = model midpoint  $x$  coordinate

A negative result indicates that the test subject has bisected the line to the left of the actual midpoint.

The extraction routine for horizontal target lines initially uses the  $y$  coordinate of the true midpoint position ( $y_{\text{actual}}$ ) to assess if the test subject has moved the pen through this location

on  $y$  axis. A secondary check assesses if the drawn  $x$  coordinate at this point ( $x_{drawn}$ ) is between the length limits of the target line, thus determining whether the line has been bisected. The model information for each target line, containing the true midpoint coordinates and length are stored in a separate text file to facilitate modification.

If direct correspondence of  $y$  coordinate (i.e.  $y_{actual} = y_{drawn}$ ) cannot be found in the test response file, then an attempt is made to find a pair of  $y$  coordinates in the test subject's response either side of  $y_{actual}$ , within the  $x$  coordinate line length limits. Interpolation is used to obtain the drawn  $x$  coordinate at  $y_{actual}$ . In Figure 5.14 the crosses show the sampled coordinates which pass either side of the target line. The circle shows the interpolated crossing point.



*Figure 5.14 : Interpolated crossing point*

The interpolation routine uses the two points vertically either side of the target line to calculate the  $x$  coordinate ( $x_{drawn}$ ) at  $y_{actual}$ :  $(x_1, y_1)$  above the target line and  $(x_2, y_2)$  below the target line (Figure 5.15).

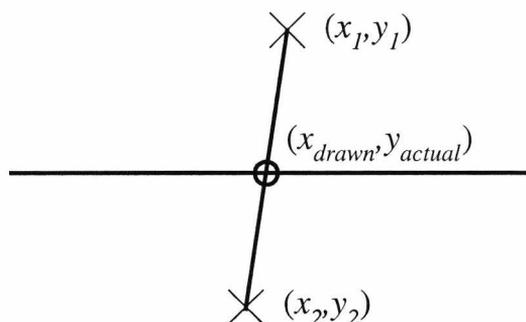


Figure 5.15 : Interpolation calculation points

The equation of the crossing line is calculated by finding the line equation common to the two points:

$$y_1 = mx_1 + c \quad \text{and} \quad y_2 = mx_2 + c \quad (5.8)$$

giving

$$m = \frac{(y_1 - y_2)}{(x_1 - x_2)} \quad \text{and} \quad c = y_1 - mx_1 \quad (5.9)$$

The  $x$  coordinate interpolated crossing point is given by:

$$x_{drawn} = \frac{(y_{actual} - c)}{m} \quad (5.10)$$

The bisection error is normalised by length of target line. The midpoint deviation error is represented as the percentage of the half-length of the target line:

$$Bisection \text{ Error } \% = \left( \frac{Midpoint \text{ Deviation}}{\left( \frac{Target \text{ Line Length}}{2} \right)} \right) \times 100 \quad (5.11)$$

### 5.3.3.2 Direction of Bisection

The direction of bisection is calculated from the y coordinates of the test patient's cancellation stroke. As the zero origin for the tablet is in the top left hand corner, an increasing y coordinate trend indicates a downwards (towards the test subject) movement. Trends are calculated by comparing pairs of y coordinates representing the pen cancellation mark and accumulating the number of y value step increments and decrements. For example the following list of y coordinates contains 3 increment steps and 1 decrement:

1345, 1347, 1350, 1352, 1348

The highest accumulator value indicates the direction trend, in this case downwards, towards the test subject.

### 5.3.3.3 Quadrant and Line Length Comparison

The 8 overlays used in the bisection task provide opportunities for performance comparisons relative to the vertical side and quadrant position, as well as line length. Expected performance from a neglect population is for a greater bisection error to the right of the midpoint for lines positioned on the left of the overlay. Mean bisection error should increase proportionally for all population groups when bisecting the longer target lines, with the neglect group accordingly still producing the largest error [75].

### 5.3.4 Cancellation Features

The features implemented for the analysis of the cancellation task include many new dynamic assessments of sequence, timings and construction. These can be used to supplement or, indeed, enhance the traditional static assessment of the number of targets cancelled on a single overlay. In particular, work has concentrated on the detailed assessment of timing within the cancellation process. Care must be taken when using the generic overall task time feature (as defined in Section 5.3.1.1) on the cancellation overlays, as although it will

accurately represent the execution time of the completed task, the result is dependent on the number of cancellations made. Thus a neglect patient cancelling four targets will probably have a smaller task execution time than a control patient cancelling all forty targets on the Albert's task. Because of this, features have been devised which are normalised with the number of cancellations made in the particular area of the overlay under investigation.

### 5.3.4.1 Total Number of Cancellations

To assess the number of cancellations made by a test subject, a *mask file* for each overlay, containing the position and identification of each target, is compared against the *test response* file. The following is an extract from the mask file for the first OX cancellation task:

```
1, 4490, 3200,  
1, 5310, 1700,  
1, 5310, 2430,  
2, 1170, 2430,  
2, 1170, 3200,  
2, 2000, 2070,
```

The first column contains the type of target (1 is a 'O' target character, 2 is a 'X' distractor character) followed by the coordinates of the centre of the target. The target's identifier is assigned by the position of the record within the mask file.

In analysing the test response file, if the pen has been positioned within the cancellation area of the target, it is recorded within the *cancellation array* (a record of the number of cancellations per target). Figure 5.16 shows the cancellation area for the two types of task. In the OX task, the area is a 7x7mm zone enclosing the target character. In the Albert's task, where the orientation of the line is pseudo-random, the cancellation area forms a box with the line ends forming opposite corners of the active area.

Multiple cancellations of a single target 'Z' are recorded only if the pen has left the tablet surface *or* another target has been cancelled subsequent to 'Z' within the cancellation sequence. This prevents erroneous multiple recognition as the test subject moves the pen through the cancellation area.

The sequence in which the cancellations occurred is stored in a separate *sequence response* file which contains a list of target identifiers and the time at which the cancellation occurred. This file is also used to extract further timing data which is described in Section 5.3.4.10.

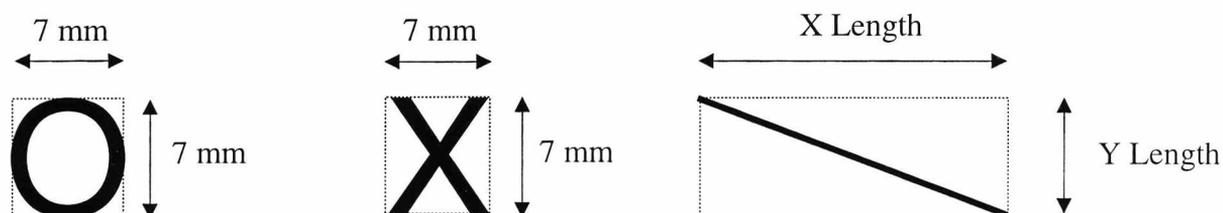


Figure 5.16 : Target areas for cancellation tasks

Calculation of the number of cancellations on the overlay is obtained by counting individual entries in the cancellation array. For the OX tasks, the number of incorrect cancellations can be obtained from observing the array entries for the ‘X’ distractor characters.

### 5.3.4.2 Number of Cancellations per Quadrant

Using the known centre  $y$  and  $x$  coordinates of the overlay, the target positions can be divided into quadrants. The same detection routines as defined in Section 5.3.4.1. are utilised which calculate the number of correct (and incorrect) cancellations. Table 5.5 shows the number of targets per quadrant:

Task	Top Left	Top Right	Bottom Left	Bottom Right
<b>OX 1</b>	3 (2)	2 (2)	3 (4)	4 (4)
<b>OX 2</b>	3 (3)	3 (3)	3 (3)	3 (3)
<b>Albert's</b>	7	10	11	8

Table 5.5 : Targets per quadrant in cancellation task. Figures in brackets denote number of distractors

The second OX task was devised to overcome the quadrant target imbalance of the other two tasks. The analysis in this second OX task does not include the centrally placed targets on each axis as described in Section 5.2.3.

### 5.3.4.3 Time per Cancellation

The overall time is calculated as for all other tasks by subtracting the time the pen was first placed on the tablet from the time the pen is removed at the end of the final cancellation. This measure, however, does not account for the number of cancellations within a sequence. A shorter global time may be caused by fewer cancellations on an overlay. Conversely, a test subject cancelling all targets will take longer than a subject only locating a few targets and therefore a raw time measurement will not reflect the patient's ability in location, planning and motor aspects of the cancellation.

The time per cancellation feature provides a more accurate overview of performance speed giving a mean time per cancellation. Timing features relating to individual phases of cancellations are extracted through other measures from these tasks (See Section 5.3.4.10)

The time per cancellation measure is calculated by :

$$\text{Time per cancellation} = \frac{\text{Overall Execution Time}}{nt (+ nd)} \quad (5.12)$$

where:

$nt$  = Number of target cancellations.

$nd$  = Number of distractor cancellations (only on OX task).

### 5.3.4.4 Processing Time per Quadrant

The time spent cancelling within each quadrant is calculated from the sequence response file. Each quadrant time is calculated separately. The processing time,  $PT_n$ , associated with the  $n^{\text{th}}$  target cancelled by the test subject is calculated by :

$$PT_n = time_n - time_{n-1} \quad (5.13)$$

where :

$time_n$  = time when  $n^{\text{th}}$  target was cancelled.

$time_{n-1}$  = time when  $n-1^{\text{th}}$  target was cancelled.

The processing time per quadrant is the sum of the processing times associated with targets cancelled within the quadrant under investigation. For example, if the 5<sup>th</sup> target cancelled is in the upper right quadrant, the time  $PT_5$  is added to the accumulator for this quadrant. Positional data of targets is obtained by cross-referencing with the mask file to obtain the coordinates from which the quadrant can be computed.

#### 5.3.4.5 Processing Time per Cancellation in Quadrant

This measure uses the processing time per quadrant feature calculated in the previous section. Dividing by the number of cancellations obtained from the sequence response file, the time spent in each quadrant is normalised according to the number of cancellations the test subject makes.

Separate features are calculated for each quadrant by the following formula:

$$PTCQ_{[x]} = \frac{PTQ_{[x]}}{NCQ_{[x]}} \quad (5.14)$$

where :

$PTCQ_{[x]}$  = Processing time per cancellation within quadrant  $x$

$PTQ_{[x]}$  = Processing time within quadrant  $x$

$NCQ_{[x]}$  = Number of cancellations within quadrant  $x$

### 5.3.4.6 Sequence Starting Location (and Quadrant)

The first entry in the sequence response file gives the initial target (or distractor) cancelled by the test subject. By again cross-referencing the mask file, the coordinates of the initial target can be extracted and the starting quadrant found. Right CVA neglect patients with a left side affected visual field would be expected to start a sequence towards the right hand side of the overlay whereas normal subjects tend to start to the left of the overlay [61].

### 5.3.4.7 Sequence Analysis

Directly extracted from the sequence response file is the order in which the test subject cancelled the targets. Assessment can be made about how the drawn sequence differs for a series of standard predefined sequences for each overlay. These model sequences have been categorised as either *raster* or *snake* patterns and have been defined for each corner target (top left, bottom right etc.) and movement direction (left to right, top to bottom movement etc.) Figure 5.17 shows the two types of traversal method. Whereas the snake pattern systematically moves up and then down columns or rows, the raster pattern moves down a column (or across a row) and then 'jumps back' to the beginning of the next. Table 5.6 details the 16 archetypal sequences defined for the OX1 task. The numbers in each sequence correspond to the target identifier which are shown in Figure 5.18.

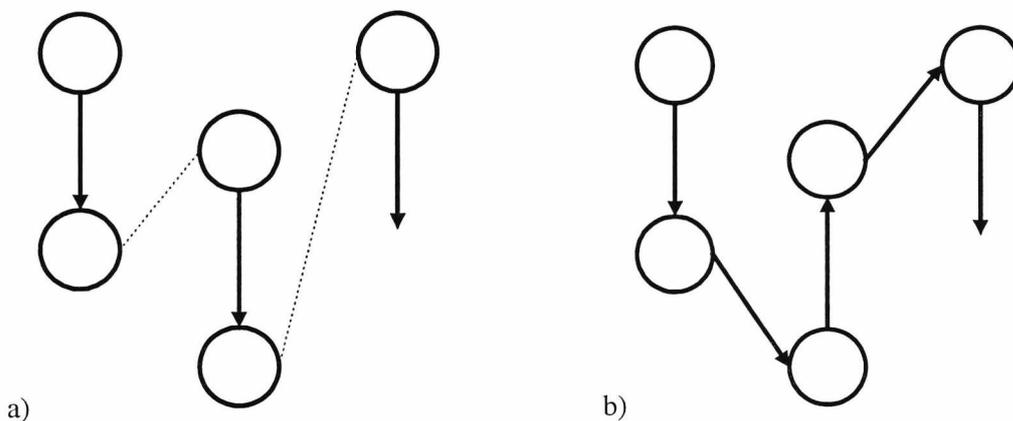


Figure 5.17 : (a) Raster and (b) snake cancellation traversal method

Sequence ID	Traversal Method	Start Side	Start Position in Column	Traversal Direction	Sequence											
					1	2	3	4	5	6	7	8	9	10	11	12
1	Raster	Left	Top	Vert	1	2	3	4	5	6	7	8	9	10	11	12
2	Raster	Left	Bottom	Vert	2	1	4	3	6	5	8	7	10	9	12	11
3	Raster	Right	Top	Vert	11	12	9	10	7	8	5	6	3	4	1	2
4	Raster	Right	Bottom	Vert	12	11	10	9	8	7	6	5	4	3	2	1
5	Raster	Left	Top	Horiz	1	5	11	2	9	3	7	12	6	8	4	10
6	Raster	Left	Bottom	Horiz	4	10	6	8	3	7	12	2	9	1	5	11
7	Raster	Right	Top	Horiz	11	5	1	9	2	12	7	3	8	6	10	4
8	Raster	Right	Bottom	Horiz	10	4	8	6	12	7	3	9	2	11	5	1
9	Snake	Left	Top	Vert	1	2	4	3	5	6	8	7	9	10	12	11
10	Snake	Left	Bottom	Vert	2	1	3	4	6	5	7	8	10	9	11	12
11	Snake	Right	Top	Vert	11	12	10	9	7	8	6	5	3	4	2	1
12	Snake	Right	Bottom	Vert	12	11	9	10	8	7	5	6	4	3	1	2
13	Snake	Left	Top	Horiz	1	5	11	9	2	3	7	12	8	6	4	10
14	Snake	Left	Bottom	Horiz	4	10	8	6	3	7	12	9	2	1	5	11
15	Snake	Right	Top	Horiz	11	5	1	2	9	12	7	3	6	8	10	4
16	Snake	Right	Bottom	Horiz	10	4	6	8	12	7	3	2	9	11	5	1

Table 5.6 : Archetypal sequences defined for OX1 cancellation task

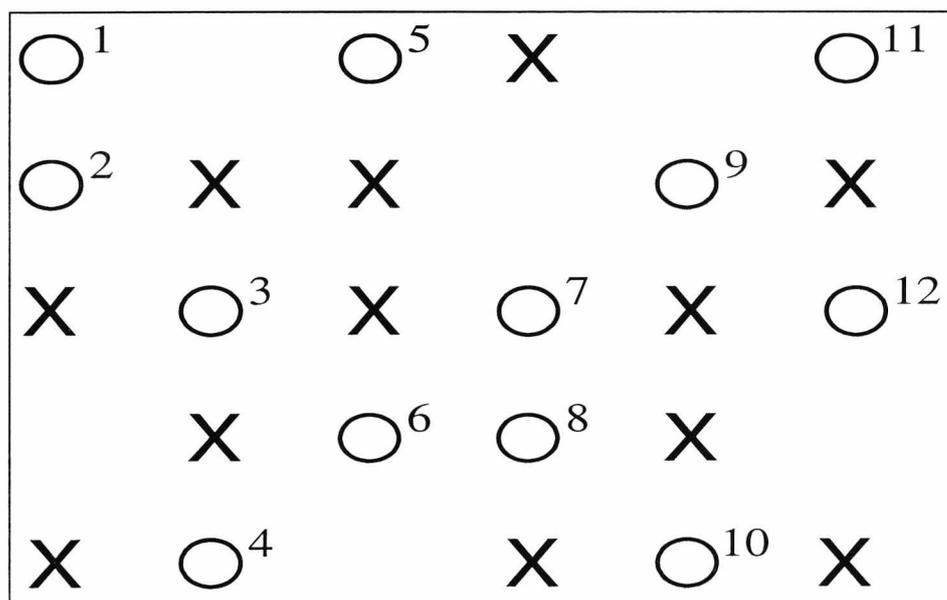


Figure 5.18 : Location identification numbers (OX1 task)

Due to the small number of targets on the OX1 task, this is the only cancellation overlay to be assessed by sequence.

The drawn sequence is classified by finding the best match against the model sequences. Before classification, the drawn sequence is pre-processed to remove all 'X' character (distractor) identification references, as only 'correct' cancellation locations (targets) are included in the predefined sequences. Sequence duplications within a sequence are also removed, forming a list describing the order in which the 'O' characters were located, allowing a direct comparison against the reference patterns.

In an initial implementation, classification was determined by finding the largest correlation between all of the model patterns and the test attempt sequence. Correlation reflects the linear relationship between two data sets and, as shown in Figure 5.19(a), will correctly classify a sequence even with a data item missing. The items within the sequence are, however, identifiers rather than weighted or scoring values and so, when a single item is misaligned with the model sequence (shown in Figure 5.19(b)) then the correlation score does not accurately represent the similarity between the drawn and model sequences.

1	2	3	4	5	6	7	8	9	10	11	12	<i>Model sequence</i>
2	3	4	5	6	7	8	9	10	11	12	-	<i>a) Correlation = 1</i>
2	3	4	5	6	7	8	9	10	11	12	1	<i>b) Correlation = 0.54</i>

*Figure 5.19 : Sequence correlation error example*

To overcome this problem, a direct correctness-of-match between the each model and test sequence is calculated which assesses the match in each of the sequence positions. The number of correct matches is recorded. The test attempt sequence is then barrel shifted one position to the right (the location identifier in the extreme right hand position is moved to the start of the sequence) and again a correctness of match score is calculated. This procedure is repeated until the drawn sequence 'returns' to its original position. Figure 5.20 shows this process. As can be seen, the number of matches varies as the pattern is shifted.

Recording the highest number of matches (a measure of the conformity of the completion sequence) produces two other performance results:

- The number of shifts required to obtain the highest match. This result also represents the position within the model sequence where the test attempt started to conform to a predefined sequence.
- The pre-defined sequence producing the highest match.

1	2	3	4	5	6	7	8	9	10	11	12	<i>Model sequence</i>
2	3	4	5	6	7	9	11	10	12	8	-	<i>Actual sequence (0)</i>
-	2	3	4	5	6	7	9	11	10	12	8	<i>Right shift 1 (6)</i>
8	-	2	3	4	5	6	7	9	11	10	12	<i>Right shift 2 (2)</i>
12	8	-	2	3	4	5	6	7	9	11	10	<i>Right shift 3 (1)</i>

Figure 5.20 : Barrel shift correctness of match operation. The figure in the parenthesis is the correctness-of-match score

### 5.3.4.8 Duplications

Interrogating the cancellation array, this feature indicates the number of duplicate cancellations that are made within a sequence. Each entry in the array, corresponding to the individual cancellation targets, is analysed to establish if the target has been cancelled more than once. If this is the case, then the number of cancellations greater than the single case is summed over the sequence.

$$duplications = \sum_{i=0}^n (((can[i]) - 1) \mathbb{1}(can[i] > 1)) \quad (5.15)$$

where:

$n$  = number of targets on overlay

$can[i]$  = number of cancellations made at cancellation position  $i$

### 5.3.4.9 Path Crossings

An extension of the completion sequence analysis is an assessment of the number of ‘path crossings’ within the strategy. Figure 5.21 is an example from the OX1 task of a strategy with a single crossing point. The figure also shows a classic neglect response for this type of cancellation task in that the cancellation sequence is started over at the right hand side of the overlay (target 9) and targets to the left were not cancelled.

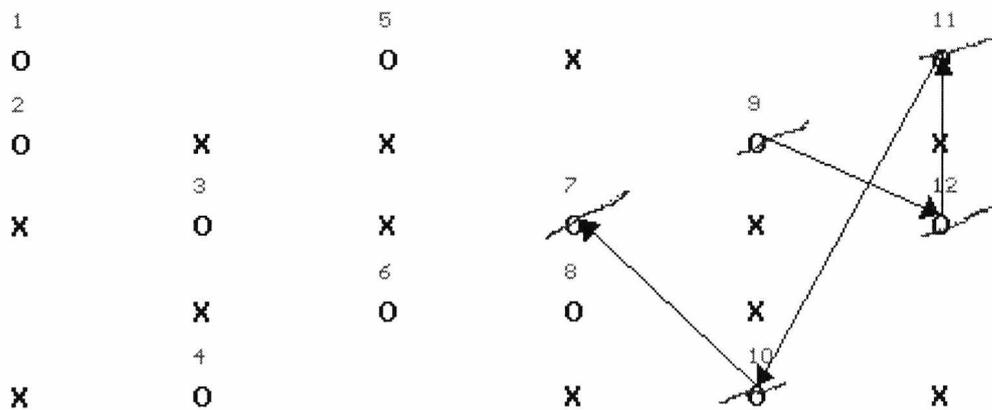


Figure 5.21 : Path with single crossing point

Calculation of path crossings again uses the sequence list with reference to the list of target coordinates contained in the mask file. Consecutive pairs of cancellation points within the sequence form *virtual lines* indicating the sequence path. Each virtual line is compared against others in the sequence to test for crossings using a line intersection routine [181]. The case where one line starts and another ends at the same point (i.e. the next line in sequence) is ignored.

Line crossings are detected by the following line intersection method: the two line segments for inspection are defined by the coordinates of end-points: Line 1 from  $(x_1, y_1)$  to  $(x_2, y_2)$  and Line 2 from  $(x_3, y_3)$  to  $(x_4, y_4)$ . Using the standard equation for a straight line:

$$0 = ax + by + c \quad (5.16)$$

The following equations were used to define each virtual line giving:

$$0 = a_1 x + b_1 y + c_1 \quad \text{for Line 1} \quad (5.17)$$

$$0 = a_2 x + b_2 y + c_2 \quad \text{for Line 2} \quad (5.18)$$

where :

$$a_1 = (y_2 - y_1), b_1 = (x_1 - x_2) \text{ and } c_1 = (x_2 \times y_1) - (x_1 \times y_2)$$

$$a_2 = (y_4 - y_3), b_2 = (x_3 - x_4) \text{ and } c_2 = (x_4 \times y_3) - (x_3 \times y_4)$$

using standard trigonometric principles.

A simple check to see if the segments intersected can be performed by separately substituting the end location points of Line 2 into equation of the first line (Equation 5.17). If the result of both equations were of the same sign then both end points of Line 2 were above (both negative) or below (both positive) Line 1 and the lines did not intersect. A result of 0 from either equation signifies that an end point of Line 2 lies on Line 1. For verification, the same procedure is repeated for Line 1 substituted into Equation 5.18.

The point of intersection  $(x_i, y_i)$  was defined as:

$$0 = a_1 x_i + b_1 y_i + c_1 = a_2 x_i + b_2 y_i + c_2 \quad (5.19)$$

which rearranged gives :

$$y_i = \frac{-(a_1 x_i + c_1)}{b_1} = \frac{-(a_2 x_i + c_2)}{b_2} \quad (5.20)$$

This leads to the definition of two formulae to derive the point of intersection :

$$x_i = \frac{-c_1 b_2 + c_2 b_1}{-a_2 b_1 + a_1 b_2} \quad (5.21)$$

$$y_i = \frac{-(a_1 x_i + c_1)}{b_1} \quad (5.22)$$

To calculate the number of path crossings an accumulation is maintained of all instances found in the sequence. The case where an identical path crossing is detected caused simply by the swapping of line segments assigned to Line 1 and Line 2 is ignored.

#### 5.3.4.10 Inter-Cancellation Timings and Regression Analysis

The inter-cancellation times (the time interval between target cancellation points) are separated into four features for assessment purposes: *overall*, *drawing*, *movement* and *premovement*. In traditional testing (i.e. not using computer-based tests), the overall time between components has been studied, usually by videoed or other observational analysis. However, by dividing the timings further, investigations into the constructional aspects of a cancellation sequence are enabled. With reference to Figure 5.22, the following is a description of the implemented timing measurements:

- The *overall time* is defined as the period between the start of the first cancellation (Point A) to the start of second cancellation (Point C). This is the conventional timing feature as defined in Section 5.3.4.3
- The *drawing time* is defined as the timing period between the start of a cancellation (Point A) to when the pen becomes stationary at the end of the cancellation drawing (Point B).
- The *movement time* is defined as the timing period between the point at which the pen is removed from the tablet after the first cancellation to when the pen is replaced on the tablet for the next cancellation (Point C).
- The *premovement time* is the amount of time that the pen is stationary at the end of the cancellation movement (time pen stationary at Point B), indicating the time during which a decision is taken about which cancellation point to move to next. This is generally a relatively short duration in comparison with the movement time.

These times are related by the expression :

$$\text{Overall} = \text{Drawing} + \text{Movement} + \text{Premeovement} \quad (5.23)$$

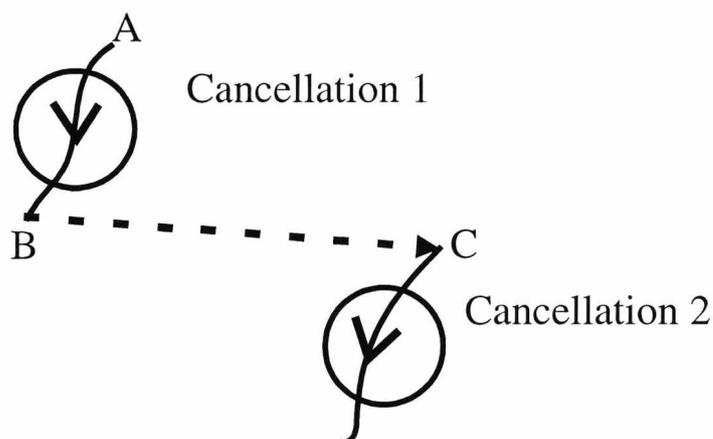


Figure 5.22 : Cancellation task assessment - timing definitions

Mean values of these measures are taken over the entire overlay and also on a quadrant basis. Again, a more accurate measure is obtained by dividing accumulated times by the number of cancellations within the area of interest, thus normalising for performance.

Having obtained a series of timings across the sequence, the linear performance trend can be calculated. This gives an indication of a speed-up (negative slope) or slow-down (positive slope) of any of the timing measures between cancellations throughout the sequence. An increase in speed (or timing reduction) indicates a performance improvement. Observing this data, a correlation can be formed between timing performance and spatial location noting the target position on the overlay. It is expected that neglect subjects will slow as they enter their neglected visual field which usually occurs towards the end of the cancellation sequence.

The implemented linear regression algorithm uses the standard least-squares method for calculating the best fit between a linear line equation ( $y = mx + c$ ) and the supplied data [182]. The slope of the calculated line ( $m$ ) gives the trend for the supplied values. The equations for calculation of  $m$  and  $c$  are as given as:

$$m = \frac{n(\sum xy)(\sum x)(\sum y)}{n(\sum (x^2)) - (\sum x)^2} \quad (5.24)$$

and

$$c = \frac{(\sum y)(\sum (x^2)) - (\sum x)(\sum xy)}{n(\sum (x^2)) - (\sum x)^2} \quad (5.25)$$

where :

$n$  = number of cancellations in sequence

$x$  = cancellation sequence position

$y$  = intercancellation (and other timing) values

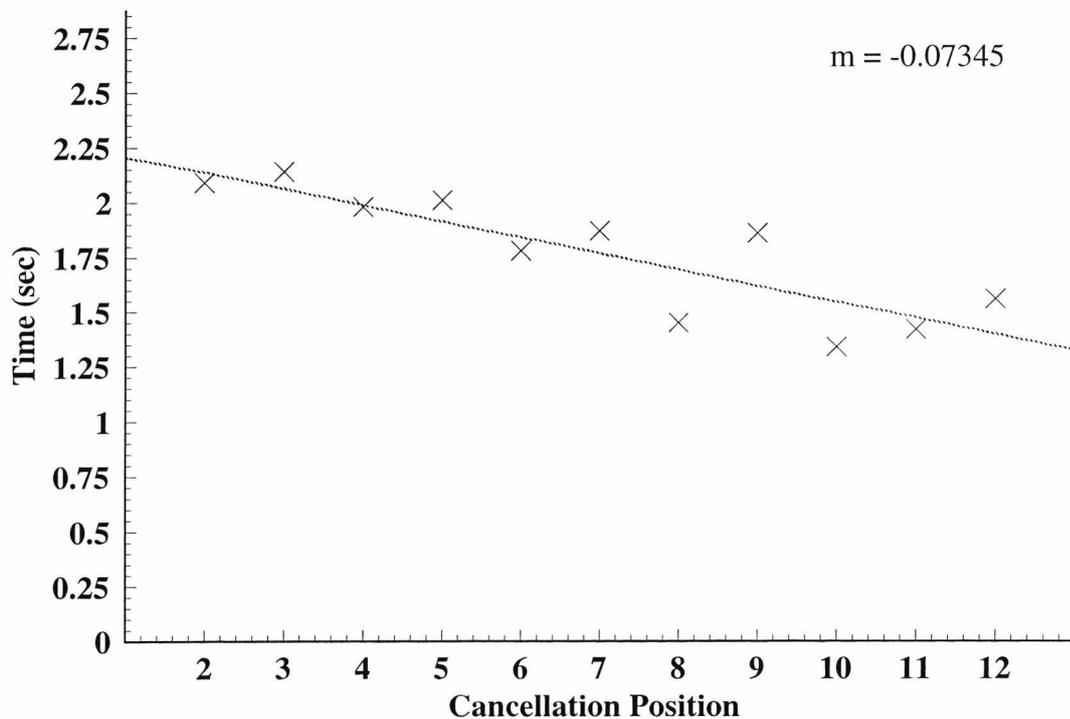


Figure 5.23 : Timing regression for overall timing

Figure 5.23 shows example data from the first OX cancellation task. The position in the sequence is represented on the  $x$  axis while the overall time between cancellations is on the  $y$  axis. As intercancellation times signify the time *between* cancellations, no time value is attributed to the first cancellation. The negative slope in the linear regression confirms an increase in performance (reduction in timing) as the test subject progresses through the

cancellation sequence. The slope value ( $m$ ) indicates the performance trend while the  $y$ -axis intercept ( $c$ ) provides the offset. This intercept value signifies the qualitative timing data within the sequence. A larger offset value indicates a mean increase in intercancellation time.

### 5.3.5 Generic Drawing Based Features

Many of the features extracted from the drawing based assessments are computationally generic and can be applied to drawings made in the figure completion, copying and drawing from memory tasks. Following the segmentation of images into *components*, a range of standard static assessments, for example, length, curvature and corner formation and a series of dynamic features, such as pen velocity and time taken to draw a component, can be applied. This section describes the standard segmentation routine used to extract individual components from drawing data and then defines the generic features that have been implemented to analyse these segments.

Task specific features such as component labelling and unique accuracy measurements are defined in the sections following the description of the generic routines.

#### 5.3.5.1 Segmentation and Component Count

Drawings are assessed on a component basis following a velocity thresholded segmentation process. Shape models, detailing component breakdown, size and angular correspondence, are defined for each of the shapes used within each of the tasks in Sections 5.3.6.1 to 5.3.6.3. For the simpler shapes, such as the square copying and diamond completion, each side is designated a component of the drawing, but for the more complex shapes, components are related to major elements of the drawing, such as a leg in the "man" or window in the "house". If any side is missing from these components, for example if only three sides of a window were drawn, this constitutes an omission of a particular component. In counting the number of components present, spatial organisation is ignored. Orientation, however, is observed and in the case of the figure completion task, only components that are of the correct vertical inversion of the given shape are counted.

Segments are extracted from a test response file by analysing and thresholding the pen velocity profile. A ‘dip’ in the profile indicates a slowing of the pen movement, which is usually found at a corner or junction between components. Using the Euclidean displacement calculation defined in Section 5.3.5.4, the velocity profile can be obtained. Figures 5.24 and 5.25 show two examples of (a) the drawing response and (b) the corresponding velocity profiles extracted from a square drawn from memory. In the first example (Figure 5.24), the segments, corresponding to individual sides of the square, are clearly separated. However, in the second example (Figure 5.25), the separation is less clear, requiring the use of a threshold to obtain the timing bounds of the segments. Calculation of features where a single component comprises more than one segment in the velocity profile is obtained by summing the features from the separate segments.

Segmentation of the drawing into individual components is performed by scanning horizontally across the profile at the  $y=0$  level and counting the number of segment areas (a segment is an area under the profile, enclosed between two boundaries defined by the profile). From the shape model of the square, it is expected that the drawing is comprised of at least four segments. If the requisite number of segments are found at the  $y=0$  scan (as is the case in Figure 5.24), then the positions at which the profile rises and falls to the  $y=0$  level defines the start and end timings of the segments. The corresponding pen coordinates at these times can be extracted from the pen response file.

If less than the expected number of segments are found at the  $y=0$  position, then scans are taken at incremental points on the  $y$  axis until a position is found with the required number of segments. In Figure 5.25, a value of  $y=2.8$  (labelled  $T$ ) is the lowest value of  $y$  to give four segments.

The case where, after scanning across all values on the  $y$  axis, the required number of segments is not found, the  $y$  position which resulted in the maximum number of segment is used as the threshold level.

This segmentation method produces the best results on smoothed profiles, eliminating errors caused by data spikes.

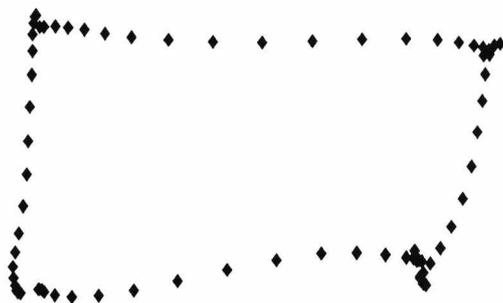


Figure 5.24a : Velocity segmentation drawing 1

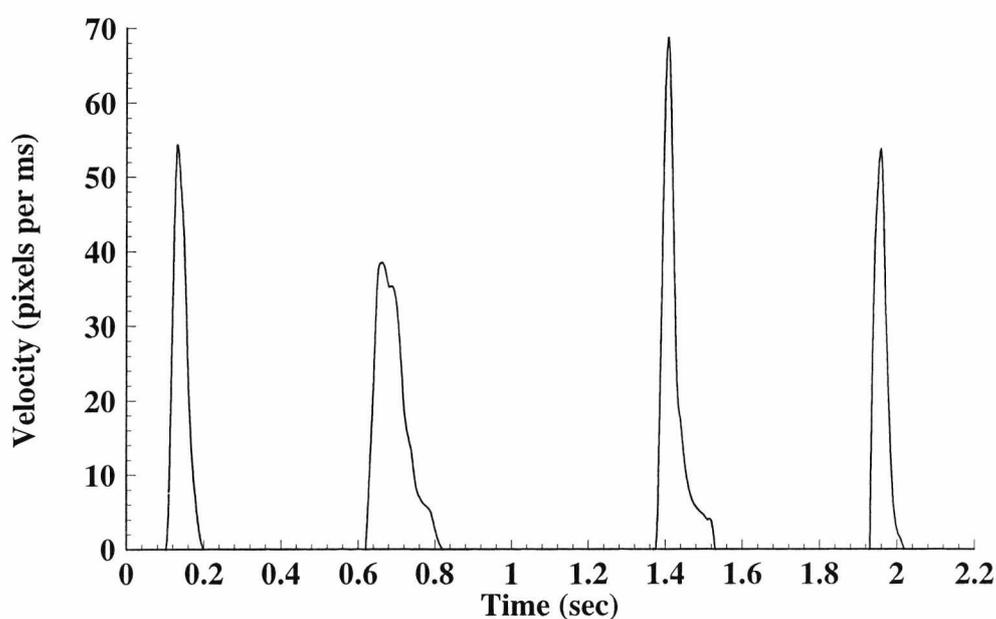


Figure 5.24b : Velocity segmentation profile 1

Labelling of the components is achieved with reference to the shape model file. Figures 5.26 to 5.29 show an example of this procedure using a drawing of a square. The drawing of this shape is expected to consist of four sides, two vertical and two horizontal, with right angle at each vertex. As the example is taken from a drawing from memory task, no size regulations are imposed on the drawing. Figure 5.26 shows the original drawing.



Figure 5.25a : Velocity segmentation drawing 1

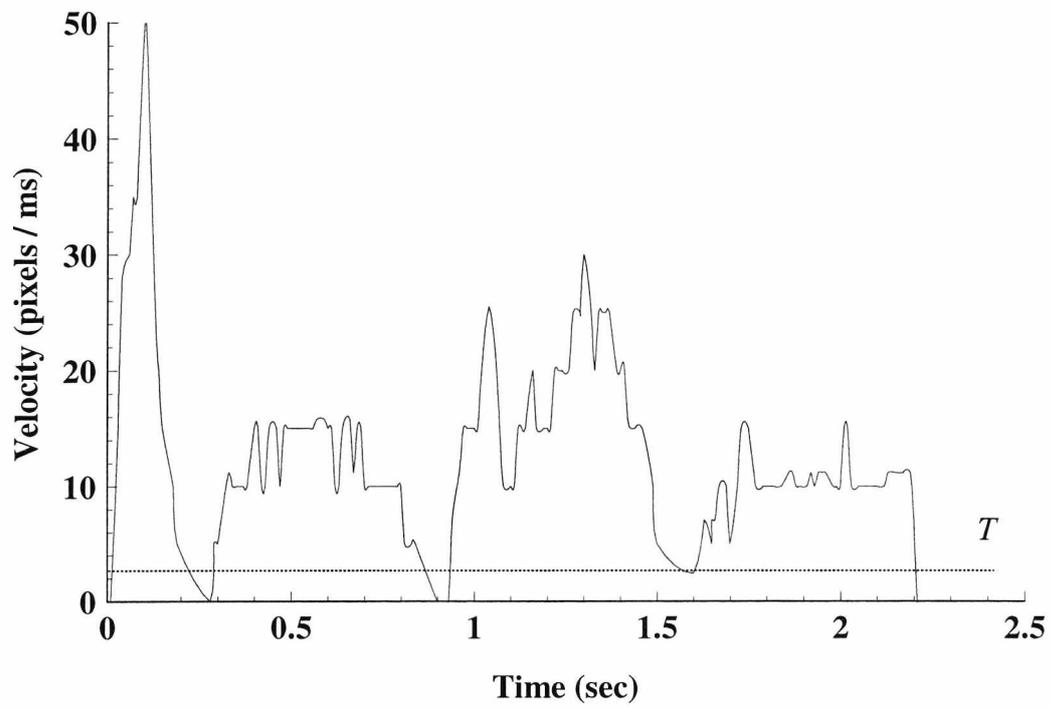
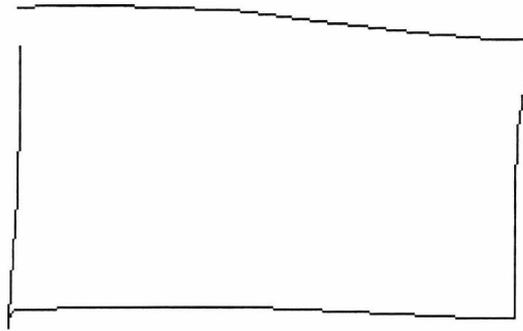
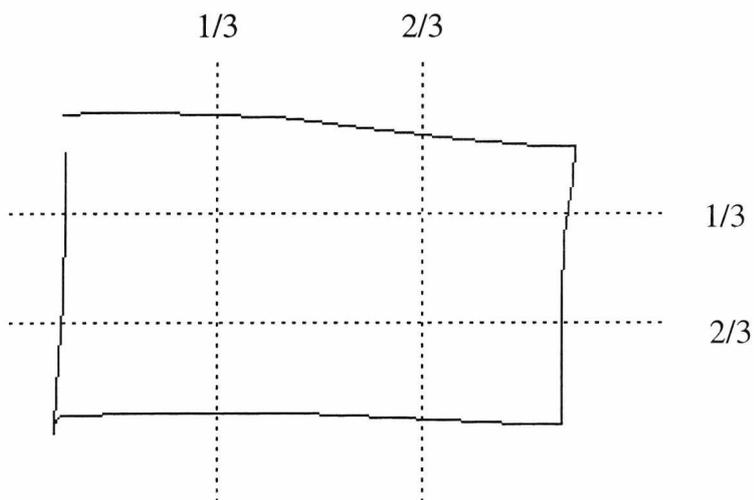


Figure 5.25b : Velocity segmentation profile 2



*Figure 5.26 : Original drawing*

From this drawing, the one-third distance points are calculated separately for each axis, dividing the drawing into a 3x3 matrix (Figure 5.7). The positions within the matrix are labelled as shown in Figure 5.8.



*Figure 5.27 : 1/3 point axis locations*

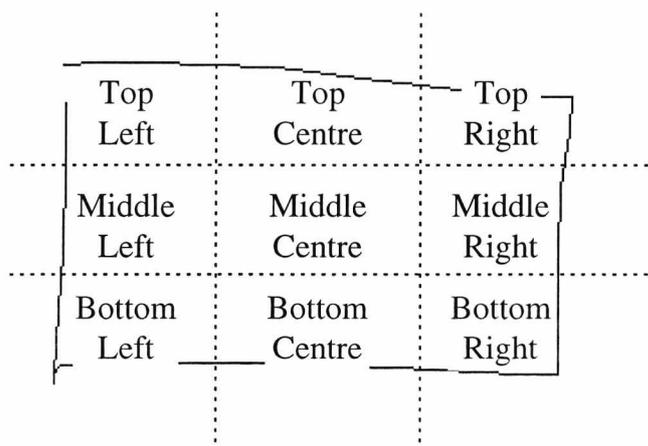


Figure 5.28 : 9 point sector definitions

Figure 5.29 shows an exploded view of the original drawing highlighting the individual components as identified by thresholding. Arbitrary labels are first assigned to each component (A to D). Using the pair of end-point coordinates (as shown by the circles on the diagram), two features can be established for each component. Firstly, the dominant drawing direction can be obtained by independently calculating the difference between  $x$  and  $y$  end-point coordinates. The largest difference indicates the dominant direction, for example in Figure 5.29, components A's dominant direction is in the  $y$  axis (there is very little difference in the  $x$  axis end-point positions). The second feature, is the spatial positioning. Using the matrix shown in Figure 5.28, both end points of component A are located in left sided sectors. From these two features we can establish that component A is a left hand vertical side of the square. Table 5.7 details the other components within the drawing.

Component	Dominant Direction	End-point Positioning
A	$y$ axis	Left-hand side
B	$x$ axis	Bottom
C	$y$ axis	Right-hand side
D	$x$ axis	Top

Table 5.7 : Automatic component labelling

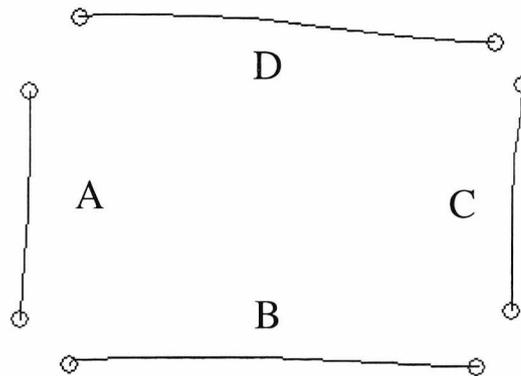


Figure 5.29 : Exploded view with segment labels

To ensure correct assignment of component identifiers, labels can be manually verified. All other processes from these assigned labels, such as size, drawing order and angle calculation are automatically processed.

### 5.3.5.2 Component Drawing, Premovement and Movement Times.

The time taken to draw the individual components is divided into the drawing, movement and premovement timing phases as defined in Section 5.3.4.10 enabling the accurate analysis of the motor and planning phases of drawing construction. With more complex shapes, a particular component may be drawn in more than one segment. Timing features are summed for the segments comprising the individual components of the drawing.

### 5.3.5.3 Component Length and Euclidean Component Distance

The component length in pixels is the sum of Euclidean distances between pairs of points drawn within a segment.

$$\text{Component Length} = \sum_{i=1}^{n-1} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \quad (5.26)$$

where :

$n$  = number of coordinates in segment

Taking a Euclidean between start and end coordinates of the segment gives the shortest travel distance of the lines but fails to account for curvature of the line.

$$\text{Euclidean Component Distance} = \sqrt{(x_{start} - x_{end})^2 + (y_{start} - y_{end})^2} \quad (5.27)$$

Analysis of the relative positioning of the start and end coordinates enables the  $x$  and  $y$  segment movement trend (left/right, up/down) to be obtained. Paired coordinate sample trend analysis as defined in Section 5.3.3.2 can be used to provide a more accurate directional assessment.

#### 5.3.5.4 Pen Velocity

Pen velocity across the surface of the tablet was calculated by taking the first derivative of the coordinate pair displacement against time. Third order, four coefficient polynomial modelling was used to obtain a derivative of displacement at each coordinate point [180]. Using a constant sampling time-base, the following approximation uses displacement values of four sets of coordinates at times  $t-2$ ,  $t-1$ ,  $t+1$  and  $t+2$ .

$$\frac{ds}{dt} \approx \frac{1}{12}(-s_{t+2} + 8s_{t+1} - 8s_{t-1} + s_{t-2}) \quad (5.28)$$

The displacements used within the calculation can be extracted on an axis-component-basis obtaining the separate horizontal and vertical velocity features or by calculating the Euclidean displacement using both the  $x$  and  $y$  components. The mean velocity is obtained by summing the velocities at individual points within the segment and dividing by the number of samples taken.

$$\text{Mean velocity} = \frac{\sum_{i=0}^n \left( \frac{ds}{dt} \right)}{n} \quad (5.29)$$

where

$n$  = the number of samples.

The peak velocity is the highest recorded velocity value within the segment analysis.

### 5.3.5.5 Velocity Profiling

A velocity profile, plotting the obtained velocity values against time for an individual segment, can provide additional information into the kinematic aspects of drawing. Typical asymptomatic velocity response from a straight line segment produces a 'bell-shaped' profile [47]. Figure 5.30 models this normal performance in drawing a straight line segment. The profile skew to the start of the segment indicates a shorter acceleration phase in comparison to the deceleration phase following the peak velocity.

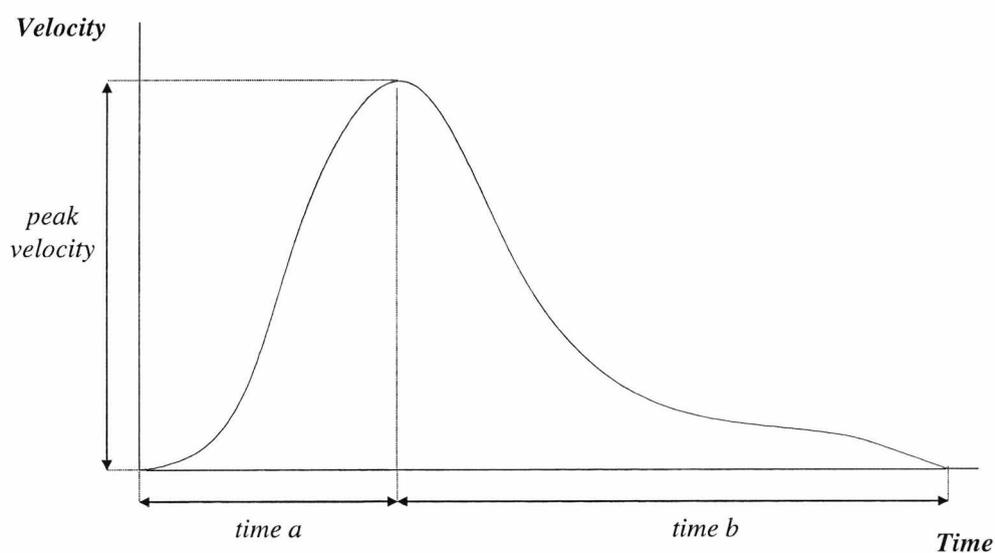


Figure 5.30 : Velocity profile model of single line component

Several timing features can be obtained from the profile. The time to reach the peak velocity (acceleration phase) and the time from peak to zero velocity (deceleration phase) have been used as performance indication features [2]. The segment drawing time is thus the sum of these two phases:

$$\begin{aligned} \text{Segment Drawing Time} = & \text{Time to reach peak velocity (acceleration phase)} \\ & + \text{Time from peak to zero velocity (deceleration phase)} \end{aligned} \quad (5.30)$$

A comparison cannot be made across freeform drawings using these defined timing features as the results are dependent on the distance drawn in the line segment. Only if this distance is known or standardised, for example movement analysis between two targets of known separation, is application of these features valid. To overcome this problem, the velocity skew percentage is an intra-profile assessment which does not account for distance drawn or variations in drawing time. This measure indicates the position in the profile in terms of total segment drawing time where the peak velocity occurs. The calculation (with reference to Figure 5.30) is made:

$$\text{velocity skew \%} = \left( \frac{\text{time to reach peak velocity (time a)}}{\text{total drawing time of the component (time a + time b)}} \right) \times 100 \quad (5.31)$$

### 5.3.5.6 Pen Acceleration

Pen acceleration is the second derivative of the coordinate displacement. As with the velocity profiling, either separate axis acceleration features or the combined Euclidean displacement can be used in the calculation.

The third order, four coefficient polynomial approximation for acceleration is given by:

$$\frac{d^2s}{dt^2} \approx \frac{1}{3}(s_{t+2} - s_{t+1} - s_{t-1} + s_{t-2}) \quad (5.32)$$

Peak acceleration, time to reach peak acceleration are calculated using an identical method to the velocity feature calculation.

### 5.3.5.7 Pen Hesitation Percentage

Pen hesitation measures the percentage of time that the pen is below a velocity threshold *and* on the tablet during the segment drawing. Using the polynomial differentiation velocity calculation defined in Section 5.3.5.4, the following formula is derived:

$$Hesitation \% = \left( \frac{\sum_{t=0}^n \left( \left[ \frac{(-s_{t+2} + 8s_{t+1} - 8s_{t-1} + s_{t-2}))}{12} \right] \leq \alpha \right)}{n} \right) \times 100 \quad (5.33)$$

where :

$\alpha$  = hesitation threshold

$n$  = number of samples

An accumulation is maintained of the number of samples that have a calculated velocity below a variable hesitation threshold. Many studies, such as Mattingley, Phillips and Bradshaw [2], use a threshold of 0, which means that hesitation is only recorded when the pen is stationary on the tablet (i.e. a pause time).

### 5.3.5.8 Image Width and Height

These simple features measure the width and height of the drawn image and can be combined to determine the area occupied by the drawing. For tasks such as figure copying and completion, the accuracy of the drawn image can be assessed by comparing the dimensions of the test subject's drawing against the width and height of the model shape. These features can be compared between patients for the drawing from memory task where no model image dimensions are supplied.

Image width is calculated by subtracting the minimum drawn  $x$  coordinate contained in the test response file from the maximum  $x$  drawn coordinate.

$$\text{Image Width} = \text{MAX}(x) - \text{MIN}(x) \quad (5.34)$$

Image height is calculated in a similar way but using the maximum and minimum drawn  $y$  coordinates.

$$\text{Image Height} = \text{MAX}(y) - \text{MIN}(y) \quad (5.35)$$

Image Area is calculated in the standard way by multiplying Image Width by Image Height.

### 5.3.5.9 Line Curvature Deviation

All drawings (with the exception of the man's head in the figure completion task) consist of straight line components. Line curvature measures the positional deviation from the best linear fit straight line for a particular segment. Curvature is calculated by two separate operations:

- Calculation of least squares best fit 'ideal' line.
- Calculation of perpendicular error distance between ideal and drawn position.

Firstly, the best linear line fit is calculated for the line segment by the least-squares method as described in Section 5.3.4.10. Figure 5.31 shows a best fit linear regression of the line segment drawn by a test subject. Points  $a$  and  $b$  signify the extremes of the best fit line, located perpendicular to the ends of the drawn line segment (Points  $e$  and  $f$ ).

The position of the points  $a$  and  $b$  are calculated by finding the equation of the line perpendicular to the best fit line. Given the equation of best fit line as:

$$y = m_{best} x + c_{best} \quad (5.36)$$

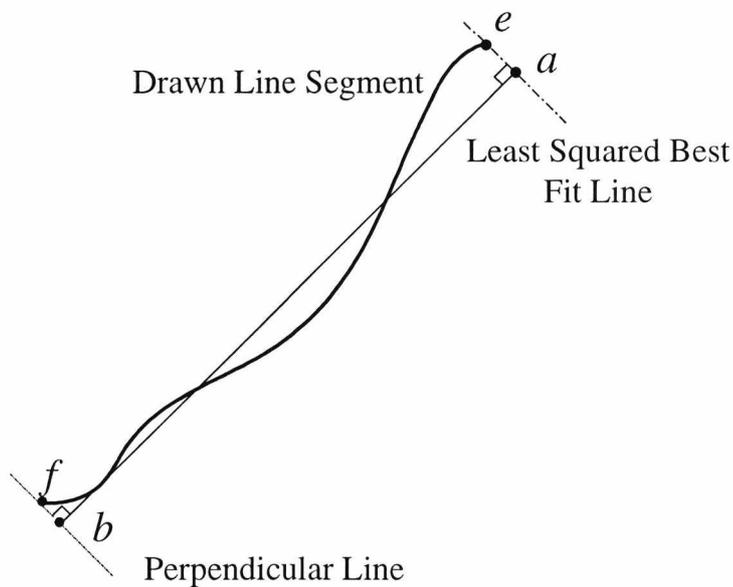


Figure 5.31 : Best line fit to drawn segment

the equation for the perpendicular line at points  $e$  and  $f$  is:

$$y = m_{\text{perpendicular}} x + c_{\text{perpendicular}} \quad (5.37)$$

where :

$$m_{\text{perpendicular}} = \left( -\frac{1}{m_{\text{best}}} \right) \quad (5.38)$$

$$c_{\text{perpendicular}} = y_e - m_{\text{perpendicular}} x_e \quad \text{for point } e \quad (5.39)$$

$$c_{\text{perpendicular}} = y_f - m_{\text{perpendicular}} x_f \quad \text{for point } f \quad (5.40)$$

Locations of the end points  $a$  and  $b$  are calculated by finding the intersection between the best fit line and the appropriate perpendicular line equation.

The error distances between the best fit line and the test subject's drawn line are calculated by taking sample points along the drawn line. Figure 5.32 shows the sample points along the line. The error distance is calculated at each of these points using the following method. With

reference to Figure 5.33,  $a$  and  $b$  are the end points of the best fit line,  $L$  is the Euclidean length of line  $ab$ ,  $p$  is a point on the drawn line,  $e$  is the perpendicular location to  $p$  on the best fit line and  $dist$  is error distance between points  $p$  and  $e$ .

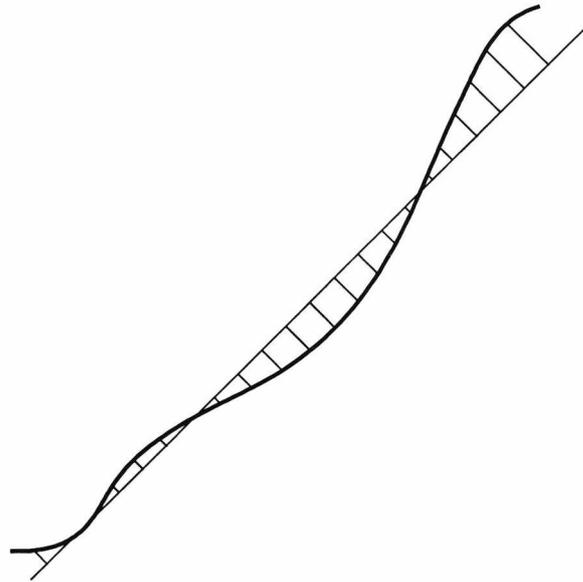


Figure 5.32 : Error distance sample points

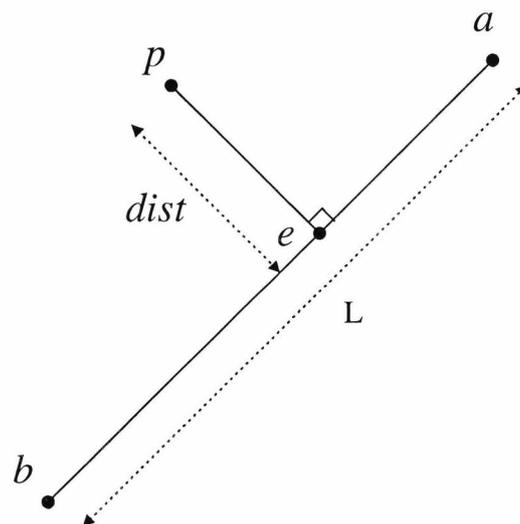


Figure 5.33 : Distance between best fit line and drawn line

The area of the triangle formed between  $a, b$  and  $p$  is given by :

$$area = \frac{dist \times L}{2} \quad (5.41)$$

where:

$$L = \sqrt{(b_x - a_x)^2 + (b_y - a_y)^2} \quad (5.42)$$

Therefore rearranging:

$$dist = \frac{2 \times area}{L} \quad (5.43)$$

The area of the triangle  $abp$  is also calculated by the following standard formula:

$$area = \frac{1}{2} |a_x b_y + b_x p_y + p_x a_y - a_y b_x - b_y p_x - p_y a_x| \quad (5.44)$$

The distance  $dist$  can therefore be computed.

The overall curvature deviation is calculated by summing the error distances and then dividing by the number of sample points :

$$Curvature\ Deviation = \frac{\sum Error\ Distances}{n} \quad (5.45)$$

where :

$n$  = number of sample point along drawn segment.



### 5.3.5.11 Component Corner Formations and Intersection Angle

Having obtained the start and end coordinates for individual components, the accuracy in corner formation (how it 'joins' to other components) can be established. A Euclidean distance is again used to calculate the error in formation. If the corner is perfectly drawn then no error distance is recorded. This measure is applied to each corner between components within a drawing and summed to produce a total component error measure.

Given two components  $Q$  and  $R$  which form a corner at their start and end coordinates respectively, the formation error is calculated :

$$\text{Formation Error} = \sqrt{(Q_{xstart} - R_{xend})^2 + (Q_{ystart} - R_{yend})^2} \quad (5.46)$$

Figure 5.28 shows two examples of corner formation error. In Figure 5.35a the two lines forming the vertex do not intersect and thus a virtual intersection point is extrapolated using the method described in Section 5.3.4.9. In the second of the examples (Figure 5.35b), the end points intersect and 'overshoot'. The same corner formation error calculation is used in both cases.

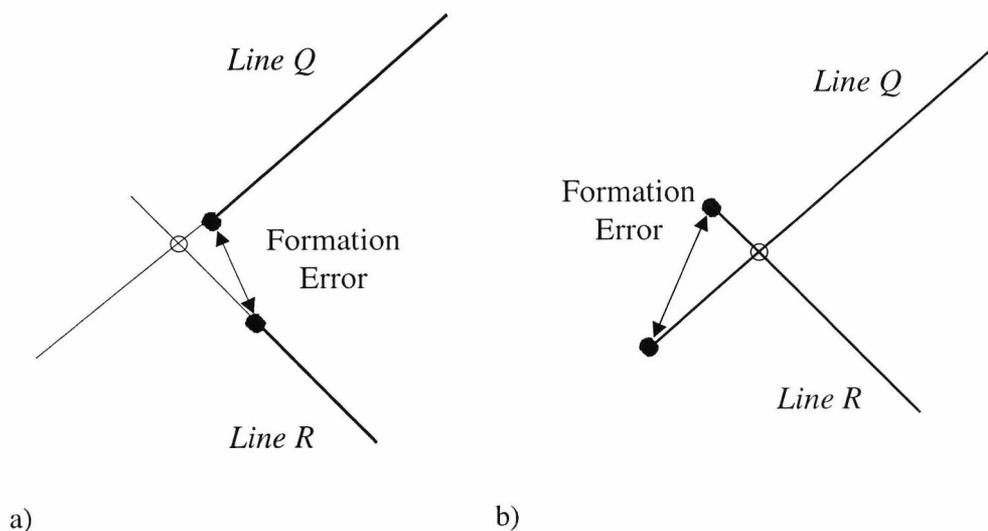


Figure 5.35 : Corner formation distance error.

The angle between two components can be calculated by using the intersection point (interpolated or otherwise). Figure 5.36 shows the two components,  $Q$  and  $R$  intersecting at a point  $S$ . The angle at the intersection is calculated using the Equation 5.47. In model based copying tasks, the calculated angle is compared against the ideal angle to assess accuracy.

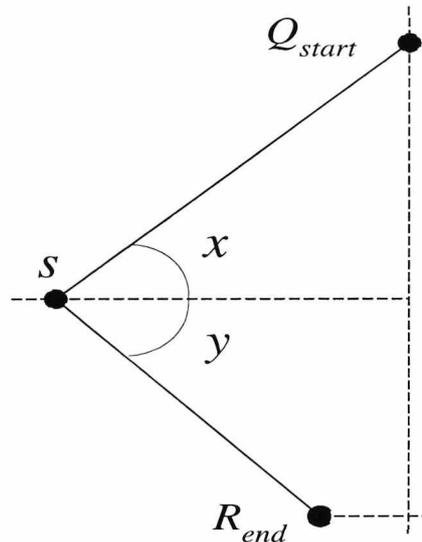


Figure 5.36 : Intersection angle calculation

intersection angle =  $x + y$

$$= \tan^{-1}\left(\frac{|Qy - Sy|}{|Qx - Sx|}\right) + \tan^{-1}\left(\frac{|Sy - Ry|}{|Sx - Rx|}\right) \quad (5.47)$$

### 5.3.5.12 Spatial Deviation

Defined within the shape models for each task are the location of the angle intersection points. This feature is the error distance between designated points within the test subject's drawing and the ideal model. It is calculated by the summing of all distance errors contained within the drawing divided by the number of intersection points. This feature is not applied to the drawing from memory task which has no distance based model.

$$Spatial\ Deviation = \frac{\sum_{i=0}^n \sqrt{(x_{drawn\_i} - x_{actual\_i})^2 + (y_{drawn\_i} - y_{actual\_i})^2}}{n} \quad (5.48)$$

where :

$n$  = number of intersection points

$actual\_i$  = model location of  $i$ th intersection point

$drawn\_i$  = drawn location of  $i$ th intersection point

### 5.3.6 Drawing Task Specific Feature Extraction

Having defined generic analysis procedures for application with the drawing tasks, this section details task specific features and component models for the individual shapes used within each task. These models contain information concerning the component count, composition, size and expected angles between components which are used by the generic routines to assess accuracy.

#### 5.3.6.1 Figure Completion Features

The components in each of the representational shapes used in the Figure Completion task are defined in Figure 5.37. Components are labelled alphabetically with corners of interest, at which angular accuracy between components is assessed, labelled numerically. Table 5.8 details the components, while Table 5.9 defines the expected angle and vertex positions.

The positional data detailed in Table 5.9 are for the overlays requiring copying to right hand side (i.e. the left hand side of the overlay is printed on the overlay). The locations for the mirror image of each shape are transformed by reflecting in the  $x$  axis around the line  $x = 3190$ .

Shape	Component	Description	Length
Diamond	a	Upper Diagonal	62 mm
	b	Lower Diagonal	62 mm
Man	a	Head	34 mm
	b	Arm	80 mm
	c	Body	23 mm
	d	Leg	66 mm
House	a	Roof	130 mm
	b	Walls	122 mm
	c	Window 1	28 mm
	d	Window 2	28 mm
	e	Door	20 mm

Table 5.8 : Figure completion components

Shape	Vertex	Location	Angle
Diamond	1	3190, 1170	-
	2	2170, 2470	90
	3	3190, 2470	-
Man	1	3190, 1930	-
	2	3070, 2250	85
	3	3110, 2390	53
	4	3110, 2720	45
	5	3190, 2520	-
House	1	1870, 2130	90
	2	2260, 2840	90

Table 5.9 : Figure completion intersection and spatial measurement points

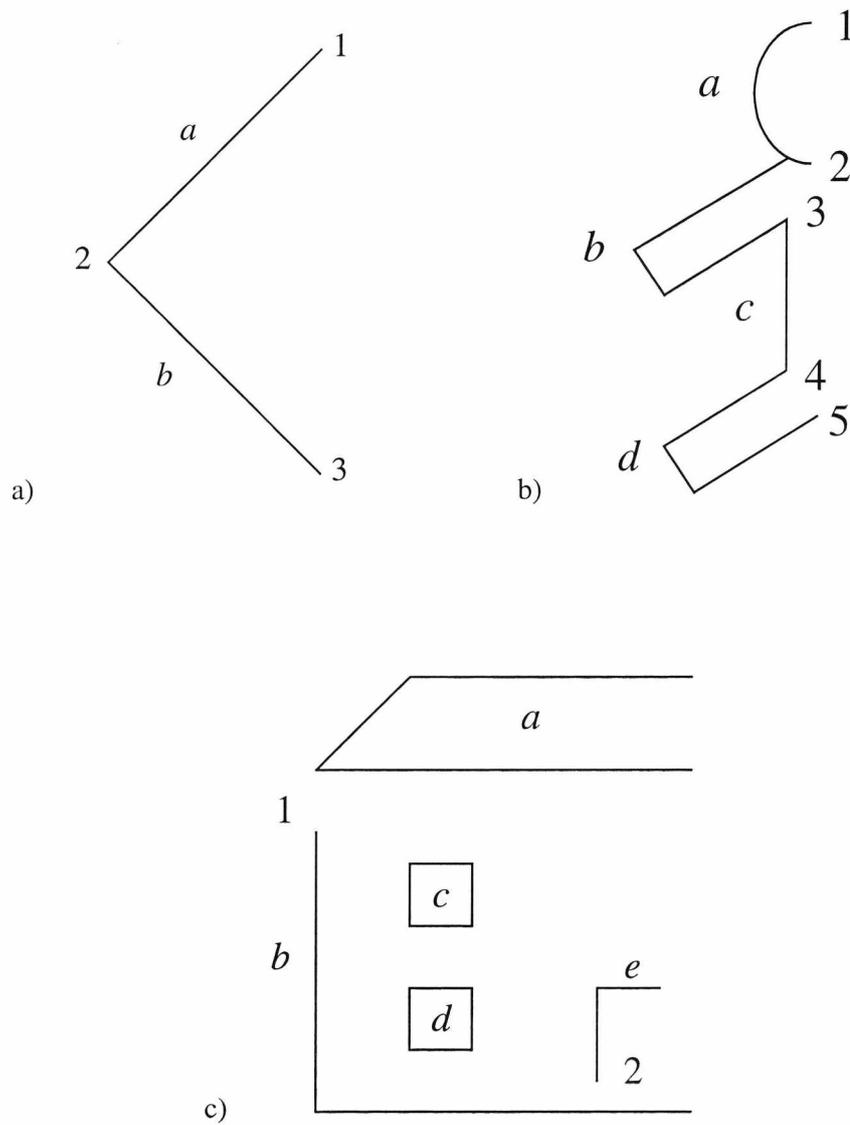


Figure 5.37 : Component analysis of figure completion task

### 5.3.6.2 Figure Copying Features

The components and intersection points for the four shapes drawn in the figure copying task are detailed in Tables 5.10 and 5.11 (referencing Figure 5.38). The component definitions for the square and cube shape are also used in the drawing from memory task. However, as these shapes have no spatial model, the intersection point locations are not used in assessment for this task.

Shape	Component	Description	Length
Square	a	Left	25 mm
	b	Bottom	25 mm
	c	Right	25 mm
	d	Top	25 mm
Cross	a	Horizontal	112 mm
	b	Vertical	112 mm
Star	a	Horizontal	28 mm
	b	Left Vertical 1	28 mm
	c	Right Vertical 1	28 mm
	d	Left Vertical 2	28 mm
	e	Right Vertical 2	28 mm
Cube	a	Left	30 mm
	b	Bottom	30 mm
	c	Right	30 mm
	d	Top	30 mm
	e	Perspective	84 mm

Table 5.10 : Figure copying components

Shape	Component	Location	Angle
Square	1	2910, 1840	90
	2	2910, 2210	90
	3	3460, 2210	90
	4	3460, 1840	90
Cross	1	3000, 1850	90
	2	3000, 2050	90
	3	3300, 1850	90
	4	3300, 2050	90
Star	1	3170, 1810	35
	2	2890, 1960	35
	3	3470, 1960	35
	4	2990, 2200	35
	5	3390, 2200	35
Cube	1	2730, 1790	90
	2	2730, 2240	90
	3	3400, 2240	90
	4	3400, 1790	90
	5	2850, 1700	135
	6	3520, 2150	135

Table 5.11 : Figure copying intersection and spatial measurement points

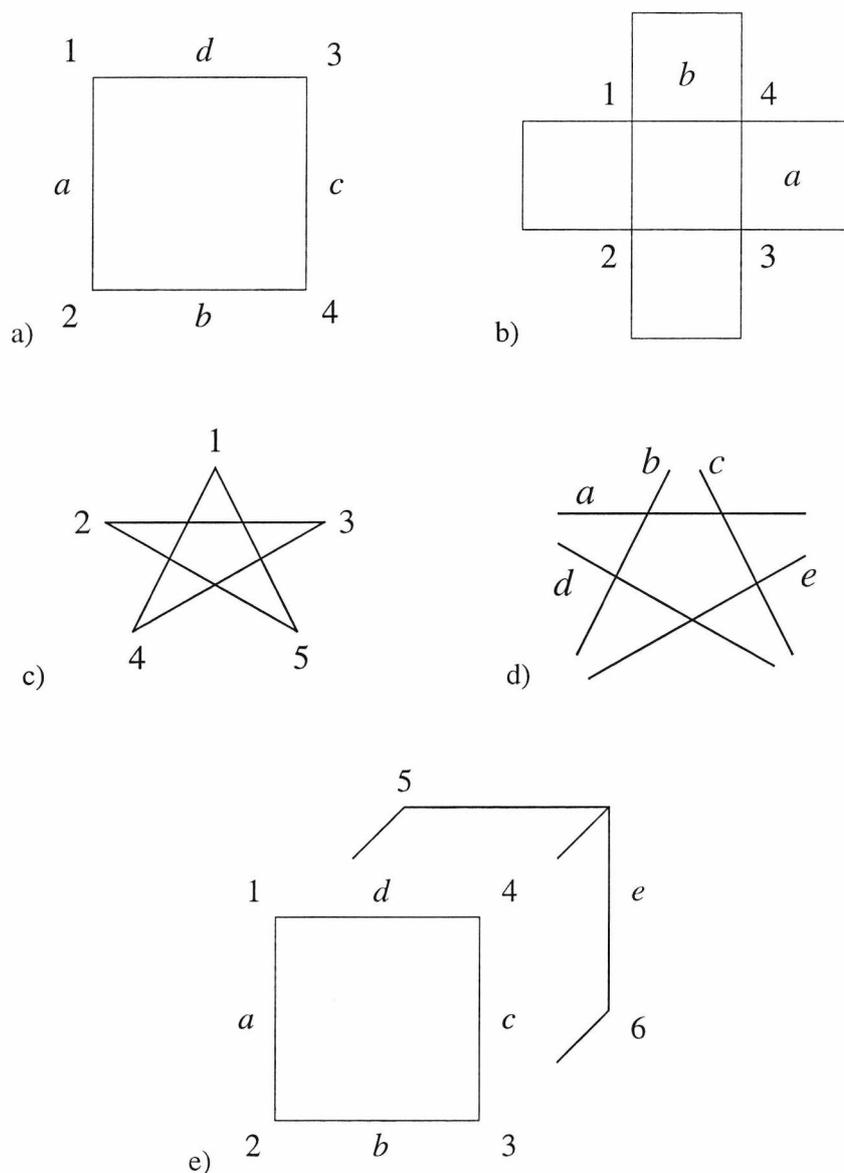
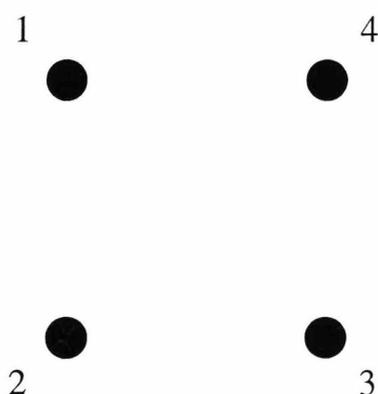


Figure 5.38 : Component analysis of figure drawing tasks

### 5.3.6.3 Drawing Profile Features

While the majority of the features extracted from these tasks concern the comparison between left and rightward kinematic movements and the relationship to spatial positioning of targets, a number of sequence based features can be extracted from the last three overlays comprising this task. Overlays 6 and 7 (Figure 5.11) require the drawing of a square between four unlabelled dots located to a particular side of the visual field. Using the labelling model

shown in Figure 5.39, a sequence analysis can be obtained and compared between the two overlays.



*Figure 5.39 : Visual field based square drawing task labels*

A similar sequence analysis can be performed with the clockface drawing task (Figure 5.10). Assigning identifiers corresponding to the clockface position of the dots, a comparison between location and spatial performance, as well as the sequence in which the dots were located, can be obtained.

## 5.4 Summary

This chapter has defined a series of pencil and paper based tasks which have been designed specifically for the assessment of visuo-spatial neglect. Many of the tasks have been implemented or modified from traditional neuropsychological assessments facilitating the validation of existing scoring mechanisms with the new test battery.

The computer based performance features that are extracted from the test responses have been implemented to satisfy two main developmental criteria. Firstly, to provide accurate, consistent and objective automated assessment of features known to detect the presence of neglect within a test subject. Secondly, a range of dynamic based features which assess timing, sequencing, kinematic and constructional aspects of the performance have been

implemented to establish if these features can distinguish between test subjects. Most importantly, the spatial aspects of neglect which are apparent using the conventional static assessments, noticeably the performance deficits within the left hand visual field for right CVA test subjects, can be investigated using intra-subject side and quadrant assessment. The implementation of features include many generic extraction algorithms that can be applied to any pen-based test response.

While automated implementations of conventional assessment techniques will provide accuracy and consistency to a range of recognised marking methodologies (such as the number of cancellations made or the bisection error distance), the novel dynamic features extracted from the test response can be used to enhance understanding of test performance and improve the sensitivity of the test. Consistent application of marking criteria enables performance monitoring over time, indicating recovery rates and rehabilitation scheme effectiveness.

The next chapter presents the results from these features using the responses collected from a population of neglect, stroke control and age matched control subjects.

## Chapter 6

# Analysis of Diagnostic Tasks

### 6.1 Introduction

This chapter presents an analysis of subjects' performance based on the features defined in Chapter 5. Feature significance is assessed using two separate patient grouping criteria, the first based on the results of the Rivermead Conventional Behavioural Inattention Test battery (*BIT based*) and a second which introduces several grades of patient performance indicating severity of neglect (*grade-based*). This second grouping reflects the fact that neglect performance cannot be treated as a homogeneous condition, an observation demonstrated by the range of results from identified neglect patients obtained from the BIT test. A significant difference between test populations for a particular feature establishes a performance deficit segregation which can be utilised to classify and assess patient test responses. Investigations are also made into the effects of age and gender on each task. Insignificant differences between test groups formed on these two criteria indicates the task suitability for use across a range of test subjects.

Following a description of the patient group demographics and the trial methodology, each task in the computer based test battery is analysed separately. Performance rules for the inclusion of an individual test response in a task analysis are defined. These rules are used primarily in the drawing based tasks (figure copying, completion and drawing from memory) to exclude a test response that does not attain a defined performance criteria such as containing the requisite number of sides or components. Responses are also excluded from other tasks where no contact has been made with the tablet surface (e.g. no cancellations/bisections have been registered).

The purpose of using inclusion criteria is threefold:

- Firstly, it provides an instant assessment of the quality of the response (by imposing a quality threshold). The percentage of a particular test group to attain a level of performance can be instantly assessed.

- Secondly, by assessing the exclusion rate for a particular test group, the difficulty in completing the task and furthermore the suitability of each task to assess all groups of test subjects can be identified. A high exclusion rate for a particular group indicates that the task is too difficult for the test subjects and therefore is unable to provide features which can be consistently measured across all subject groups. The effects of shape complexity within the drawing tasks on response exclusion are investigated
- Thirdly, it simplifies the computer-based analysis of the drawings as the assessment algorithms are all presented with a response which contained a pre-defined number of components. It also serves to demonstrate the diagnostic ability of the dynamic features. All responses included in the drawing analysis attain a standard which statically would be marked universally high. Differences detected dynamically indicate constructional properties unobtainable by static performance assessment.

The significant features from each task in the test battery are identified with both the BIT defined and grade-based grouping criteria. In particular, using the latter of these groupings, the features which produce significant results can be used to detect a finer resolution of test performance through the identification of significant differences between sub-groups formed from neglect and stroke control subjects.

The chapter concludes with an automatic classification feasibility study using the significant results identified from the computer-based assessment. The study uses a number of standard classification techniques and assesses each on the ability to identify performance characteristics from the supplied features. The aim of the study is to identify suitable areas for further exploration into result diagnosis.

## 6.2 Trial Demographics

Test data from right CVA stroke subjects was collected by a single Occupational Therapist assessor from four different geriatric hospital centres in East Kent. Each of the stroke subjects were initially assessed using the Conventional battery of the Rivermead BIT. In addition, a series of healthy age matched control subjects with no known history of vascular disease were tested by the same assessor. Identification of neglect within a stroke subject was defined by scoring below the BIT cut-off score of 130 (as defined in the BIT assessment manual [80] -

see Section 3.2.2). A list of stroke test subjects participating in the trial is contained in Appendix A.

A total of 30 neglect test subjects were identified, 17 male, 13 female, with a mean age of 74.06 (SD=7.95, range 90-59). The mean BIT score obtained by the neglect group was 87.03 (range 41-129) and the mean number of days post-stroke when subjects were tested was 103. Figure 6.1 shows the cumulative BIT scores obtained by the neglect group showing an even spread of results across the range of scores obtained.

The neglect subject group was divided into three performance related sub-groups for the second of the marking criteria, formed on the basis of performance bands within the conventional BIT score: severe neglect (group 1), moderate neglect (group 2) and mild neglect (group 3). Table 6.1 details the group membership. The inclusion scores for each performance band were selected to divide the neglect population by approximately three equal groupings at decade score division points.

<b>Group</b>	<b>Number in Group</b>	<b>BIT Score Inclusion</b>	<b>Mean Age</b>	<b>Mean BIT Score</b>
<b>Severe (1)</b>	8	<70	73.38	57.13
<b>Moderate (2)</b>	11	>69 and <100	73.27	79.27
<b>Mild (3)</b>	11	>99 and <139	75.5	116.54

*Table 6.1 : Neglect grade based analysis groupings*

Stroke subjects scoring above 129 in the Conventional BIT battery were included in the stroke control test group (SC). 58 subjects (33 male, 25 female) were included in this group with a mean age of 74.45 (SD=8.39, range 92-57) and a mean BIT score of 143.55. Testing occurred on an average of 116 days post stroke. Of the 58 subjects, 35 scored full marks on the BIT battery, the other 23 subjects scoring between 130 and 145 marks.

Figure 6.2 shows the cumulative test scores for the stroke control group. It shows that the majority of group members (60%) scored the maximum 146 marks awarded from the conventional battery.

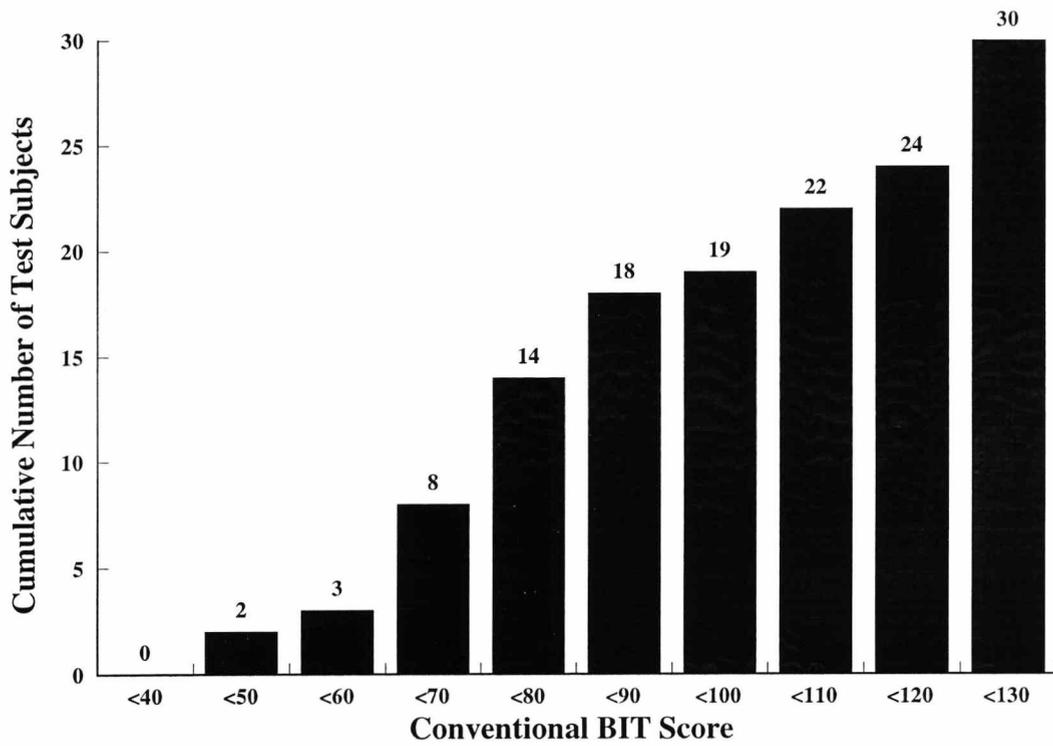


Figure 6.1 : Cumulative BIT scored obtained by neglect group

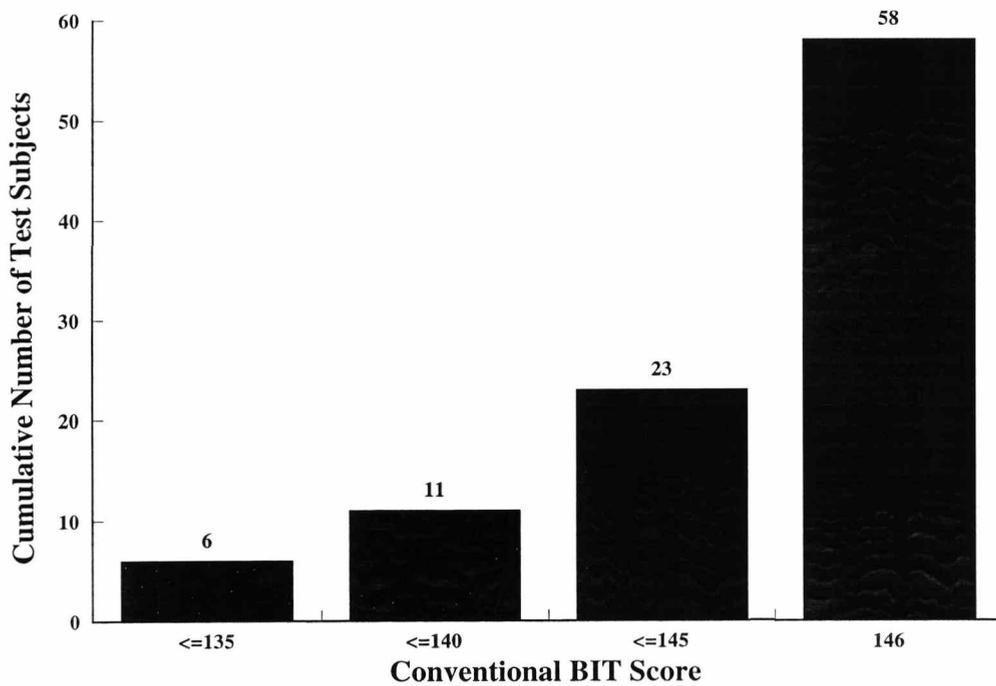


Figure 6.2 : Cumulative BIT scored obtained by stroke control group

For the grade-based grouping criteria, two sub-groups were formed from the stroke control patients: The group of 35 subjects scoring a maximum 146 on the BIT test (Group 5), and a second sub-group of patients over the BIT pass threshold, but failing to score maximum marks (Group 4). Table 6.2 details these sub-groupings. An interesting analysis from the feature results is to note the performance similarity between the AMC group and Group 5. Differences between these groups will identify effects due to the CVA.

<b>Group</b>	<b>Number in Group</b>	<b>BIT Score Inclusion</b>	<b>Mean Age</b>	<b>Mean BIT Score</b>
<b>Stroke Control Moderate (4)</b>	23	>129 and <146	75.91	139.83
<b>Stroke Control No Errors (5)</b>	35	146	73.41	146

*Table 6.2 : Stroke control grade-based analysis groupings*

The age matched control group (AMC) were not assessed with the BIT and so were included in the same group for both BIT and grade-based marking criteria. 13 age matched subjects were assessed (5 male, 8 female) with a mean age of 72.77 (SD = 4.07, range 79-63).

To analyse any performance characteristic due to age, three groups were defined which divided the neglect test population in equally sized groups (Table 6.3).

<b>Age Group</b>	<b>Range</b>
<b>1</b>	<70
<b>2</b>	>69 and <76
<b>3</b>	>75

*Table 6.3 : Geriatric age analysis groupings*

No significant differences were found in gender, age and the number of days post-stroke that testing was performed between test subjects within both the BIT and grade-based marking criteria groupings. This indicates that the groups were drawn from a similar population with

the identification (and severity) of neglect the significant difference between group membership.

### 6.3 Data Analysis

Continuous data features were analysed using an Analysis of Variance Test (ANOVA) [183][184] with a Bonferroni post-hoc multiple comparison test [185]. Once an ANOVA has identified that differences do exist in a feature, assessing all subject groups, a post-hoc multiple comparison test identifies *which* groups differ significantly. The Bonferroni test is a standard post-hoc test which adjusts the significance obtained from the ANOVA to allow for the fact the multiple comparisons are made. A detailed explanation of these procedures are contained in many statistical textbooks [166][186][187]. Significance was defined at the 5% level. Discrete features (such as starting corner on shape, number of sides drawn etc.) were assessed using a Chi-Squared test of significance [188]. Again, the significance was defined at 5%.

A feature that identifies significant performance differences between pairwise comparisons of the mean result of *each* of the test subject groups is diagnostically capable of identifying neglect, stroke control and age matched control performance. A feature producing a significant difference between the stroke control and neglect groups identifies a performance characteristic specific to neglect which differs significantly from stroke patients without neglect. A feature that does not produce a significant difference between the AMC and SC groups indicates *normal asymptomatic* performance by the SC group.

When using the grade-based test groupings, significantly different results between groups formed from neglect subjects (groups 1, 2 and 3) or stroke control subjects (4 and 5) indicate that the feature is sensitive to performance scale.

### 6.4 Point Location Results

All test subjects attempted this simple screening test. Test responses were removed from the analysis if the test subject failed to replace the pen at any point on the overlay (neither of the targets were located by the test subject) or failed to remove the pen from the overlay surface between the two dots (i.e. the test subject joined the two dots on the overlay) preventing the

calculation of timing data. Table 6.4 details the percentage of each group excluded from the analysis. While the AMC (Age Matched Control) group had very few exclusions, a quarter of the neglect group failed to complete the overlay on the left hand side, while the performance was normal for the right hand side of target overlays. Interestingly, the SC (Stroke Control) group also had almost a quarter of the group excluded for the top left target overlay, but performed normally on the other overlays.

With the removal of the overlays with no or an incorrect 'dot-joining' response, the accuracy of each subject can be analysed. Table 6.5 details the mean error distances from the centre of target dot to the subject's response. While not significant, it can be observed that normal performance is obtained for both stroke groups, apart from in the top left quadrant. Analysis of the timing intervals between the starting and target pen positions (Table 6.6) shows significant differences in the top left quadrant between SC and neglect groups ( $p=0.013$ ) and AMC and neglect groups ( $p=0.022$ ). This demonstrates the diagnostic ability of a dynamic feature which can identify a neglect characteristic from a response which would be considered 'normal' by static assessment.

No significant differences were identified between grade-based test groupings or between age groupings and gender.

<b>Group</b>	<b>Top Left</b>	<b>Top Right</b>	<b>Bottom Left</b>	<b>Bottom Right</b>
<b>AMC</b>	0	7.7	15.4	0
<b>Neglect</b>	23.3	0	23.3	0
<b>SC</b>	22.4	3.5	6.9	0

*AMC = Age Matched Controls, SC = Stroke Controls*

*Table 6.4 : Percentage of each test group excluded in point location analysis*

<b>Group</b>	<b>Top Left</b>	<b>Top Right</b>	<b>Bottom Left</b>	<b>Bottom Right</b>
<b>AMC</b>	4.70	4.04	3.96	4.15
<b>Neglect</b>	5.34	4.60	4.35	4.66
<b>SC</b>	5.51	4.62	4.42	4.91

*Table 6.5 : Mean error distances (mm) to centre of target dot*

<b>Group</b>	<b>Top Left</b>	<b>Top Right</b>	<b>Bottom Left</b>	<b>Bottom Right</b>
<b>AMC</b>	1.89	1.06	0.97	0.82
<b>Neglect</b>	4.35	2.61	1.45	3.11
<b>SC</b>	3.41	2.06	1.64	1.34

*Table 6.6 : Mean movement timing (sec) between dots*

## 6.5 Line Bisection Results

As with the Point Location Task, all test subjects attempted the Line Bisection overlays. No exclusions were made from the task. Table 6.7 shows the percentage of each of the BIT defined test groups on which side of the true midpoint the test subject bisected the eight target lines (ignoring the severity of the error). It can be observed that neglect patients on average make more rightward errors than the other two groups.

<b>Group</b>	<b>Bisection Error</b>	<b>TL Short</b>	<b>TR Short</b>	<b>BL Short</b>	<b>BR Short</b>	<b>TL Long</b>	<b>TR Long</b>	<b>BL Long</b>	<b>BR Long</b>
<b>AMC</b>	Left	46%	58%	54%	15%	23%	46%	38%	46%
	Right	54%	42%	46%	85%	77%	54%	62%	54%
<b>Neglect</b>	Left	24%	41%	17%	27%	8%	11%	4%	21%
	Right	76%	59%	83%	73%	92%	89%	96%	79%
<b>SC</b>	Left	23%	29%	21%	25%	21%	37%	22%	42%
	Right	77%	71%	79%	75%	79%	63%	78%	58%

*Table 6.7 : Horizontal bisection error direction from midpoint*

Table 6.8 details the mean bisection error (as a percentage of the half-line length) across the eight overlays. The inaccuracy of the neglect group is apparent in comparison with the other two groupings.

<b>Group</b>	<b>Bisection Error % (SD)</b>
<b>AMC</b>	0.87 (0.04)
<b>Neglect</b>	11.04 (0.15)
<b>SC</b>	3.99 (0.08)

*Table 6.8 : Horizontal bisection error over 8 overlays for BIT defined groups*

Applying an ANOVA to this data, there are significant performance differences between the AMC and the neglect groups and also between the SC and neglect groups (both  $p < 0.001$ ) indicating both the diagnostic capability of the task and (as the results correlate with conventional testing performance) the computer based implementation and testing infrastructure.

To assess whether all overlays within the bisection test battery are diagnostically relevant, the effect of length and location on bisection error can be investigated. Table 6.9 shows the bisection errors for each group across the four overlays for each line length. The greater error associated with increase in line length has been widely reported in the literature [75] and, as can be observed, results from the study support this claim, particularly with the neglect group.

<b>Group</b>	<b>Long Line (140mm) Bisection Error % (SD)</b>	<b>Short Line (50mm) Bisection Error % (SD)</b>
<b>AMC</b>	0.88 (0.03)	0.85 (0.05)
<b>Neglect</b>	14.71 (0.15)	6.83 (0.13)
<b>SC</b>	4.70 (0.09)	3.27 (0.07)

*Table 6.9 : Horizontal bisection error by line lengths for BIT defined groups*

The 50mm line overlays do not produce any significant differences between the test groups. The 140mm line, however, produces significant results between the neglect group and both AMC and SC groups (again, both  $p < 0.001$ ). This indicates that the significance in the overall test results relies on the 140mm lines and that the diagnostic capability of the bisection task can be maintained solely from the use of overlays with this length of target line.

As the bisection target line is located in each quadrant over the four overlays of a particular line length, the significance of line position can be assessed. Table 6.10 records the mean bisection errors from both short *and* long lines in a particular quadrant.

The left handed quadrants of the page are located in the RCVA neglect patient's inattentive visual field and so the observed effects are expected. Examination of the mean bisection errors indicates that the bottom left hand quadrant produces a greater error in neglect patients. These effects are examined further by the assessment of bisection errors from the longer 140mm lines. The data are contained in Table 6.11.

<b>Group</b>	<b>Top Left*</b>	<b>Top Right</b>	<b>Bottom Left*</b>	<b>Bottom Right+</b>
<b>AMC</b>	1.39 (0.04)	-0.084 (0.04)	0.59 (0.05)	1.56 (0.05)
<b>Neglect</b>	11.97 (0.14)	6.10 (0.16)	13.05 (0.13)	10.04 (0.15)
<b>SC</b>	4.83 (0.09)	2.62 (0.08)	5.47 (0.07)	3.08 (0.1)

*Table 6.10 : Horizontal bisection error by quadrant for BIT defined groups*

\* ANOVA of both left hand side quadrants produced a significance of  $p < 0.001$  between neglect and both control groups.

+ ANOVA in the bottom right quadrant produced a significance of  $p = 0.026$  between age matched controls and neglects and  $p = 0.006$  between stroke controls and neglects.

Results in the top right hand quadrant were not significant.

<b>Group</b>	<b>Top Left*</b>	<b>Top Right+</b>	<b>Bottom Left*</b>	<b>Bottom Right#</b>
<b>AMC</b>	1.67 (0.03)	0.07 (0.03)	1.60 (0.04)	0.05 (0.03)
<b>Neglect</b>	15.31 (0.14)	11.56 (0.17)	16.32 (0.13)	10.61 (0.16)
<b>SC</b>	5.01 (0.08)	1.99 (0.05)	5.68 (0.08)	0.22 (0.07)

*Table 6.11 : Horizontal bisection error by quadrant for 140mm bisection lines for BIT defined groups*

ANOVA significances:

\*  $p < 0.001$  both between neglect and age matched and also between neglect and stroke control.

+  $p = 0.002$  between neglect and age matched.  $p < 0.001$  between neglect and stroke control.

#  $p = 0.006$  between neglect and age matched.  $p < 0.001$  between neglect and stroke control.

Again, the most significant results occur to the left side of the overlay, with the bottom left quadrant producing a larger mean error for the neglect group. Significant results are obtained in all four quadrants when bisecting the longer 140mm target lines, confirming their ability to separate test groupings.

From these data it can be concluded that, though greater mean errors are produced by the neglect group in comparison with the two control subject groups in bisecting shorter lines, these differences are not significant between test populations. The longer lines, however, produce significant differences in all four quadrants of the page, with the greatest separation between groups produced when bisecting in the bottom left quadrant of the overlay.

No significant results were found relating gender or patient age to BIT group and bisection error - overall, by line length or by quadrant. This supports the claims of Roig [71] and proves the universality of test application.

Having proved that the bisection error produces significant differences on particular line positions and lengths between BIT defined groups, the next investigation is the division of the test population into grade-based groupings (as defined in Section 6.3). If significant differences can be detected it can be inferred that bisection is sensitive to neglect severity.

As an initial investigation of the differences between these groups, the mean bisection error was calculated across all eight overlays (Table 6.12). Secondly, using the findings of the initial bisection study, the analysis concentrated on the 140mm line lengths both globally (also Table 6.12) and by quadrant (Table 6.13).

<b>Group</b>	<b>Bisection Error % (SD) across 8 overlays</b>	<b>Bisection Error % (SD) 140 mm length lines only</b>
<b>1</b>	12.93 (0.17)	17.58 (0.18)
<b>2</b>	5.68 (0.15)	21.06 (0.14)
<b>3</b>	5.35 (0.12)	6.79 (0.11)
<b>4</b>	4.35 (0.08)	3.62 (0.08)
<b>5</b>	3.75 (0.09)	3.03 (0.07)
<b>AMC</b>	0.88 (0.03)	0.85 (0.05)

*Table 6.12 : Horizontal bisection error across 8 overlay (both line lengths) and for 140mm lines using severity marking scheme defined groups*

ANOVA significances:

8 Overlays: Grp 1 vs. Grp 3 :  $p=0.025$ , Grp 1 vs. Grp 4 :  $p=0.002$ , Grp 1 vs. Grp 5 :  $p<0.001$ . All other comparisons not significant.

140 mm lines: Grp 1 vs. Grp 4 :  $p=0.005$ , Grp 1 vs. Grp 5 :  $p=0.002$ , Grp 2 vs. Grp 3 :  $p<0.001$ , Grp 2 vs. Grp 4 :  $p<0.000$ , Grp 2 vs. Grp 5 :  $p<0.000$ . All other comparisons not significant.

Group	Top Left	Top Right	Bottom Left	Bottom Right
<b>1</b>	9.79 (0.15)	10.60 (0.25)	22.68 (0.10)	18.12 (0.19)
<b>2</b>	23.12 (0.13)	17.21 (0.16)	21.35 (0.13)	17.16 (0.15)
<b>3</b>	11.43 (0.12)	6.43 (0.11)	9.32 (0.13)	-0.03 (0.07)
<b>4</b>	4.24 (0.07)	1.94 (0.06)	7.34 (0.09)	1.38 (0.07)
<b>5</b>	5.56 (0.09)	2.03 (0.05)	4.57 (0.06)	-0.57 (0.06)
<b>AMC</b>	1.67 (0.03)	0.07 (0.03)	1.60 (0.04)	0.05 (0.03)

*Table 6.13 : Horizontal bisection error by quadrant for 140mm bisection lines using severity marking scheme defined groups*

ANOVA significances:

Top Left

Grp 2 vs. Grp 4 :  $p < 0.001$ , Grp 2 vs. Grp 5 :  $p < 0.001$ .

Top Right

Grp 2 vs. Grp 4 :  $p = 0.001$ , Grp 2 vs. Grp 5 :  $p < 0.001$ .

Bottom Left

Grp 1 vs. Grp 4 :  $p = 0.012$ , Grp 1 vs. Grp 5 :  $p = 0.001$ , Grp 2 vs. Grp 3 :  $p = 0.048$ , Grp 2 vs. Grp 4 :  $p = 0.003$ , Grp 2 vs. Grp 5 :  $p < 0.001$ .

Bottom Right

Grp 1 vs. Grp 3 :  $p = 0.003$ , Grp 1 vs. Grp 4 :  $p = 0.002$ , Grp 1 vs. Grp 5 :  $p < 0.000$ , Grp 2 vs. Grp 3 :  $p = 0.001$ , Grp 2 vs. Grp 4 :  $p < 0.001$ , Grp 2 vs. Grp 5 :  $p < 0.001$ .

All other comparisons not significant.

Using an ANOVA analysis, significant differences can be observed within groups 1,2 and 3, the neglect grade-based sub-groups (significant between moderate and mild neglect groups), on both the bisection error across all eight overlays and using the 140mm length target lines. Analysis of the effects of line positioning on the performance sensitivity indicates a neglect group separation in targets located in the bottom half of the overlay. No differences could be found between the two SC groups (4 and 5).

No differences or significant effects were found for the direction of cancellation

## 6.6 Cancellation Results

The three cancellation tasks (OX1, OX2 and Albert's) have been analysed by a wide range of features which have been defined in Section 5.3.4. As well as the conventional assessment of the correct number of cancellations on the overlay, novel timing and constructional features have been assessed. While all test subjects performed the tasks, a single stroke control patient failed to cancel any of the targets on the three overlays and therefore was removed from the analysis. No significant effects were found due to age or gender in any of the three tasks.

### 6.6.1 OX1 Results

Table 6.14 shows the significant features extracted from the first OX overlay. Mean and standard deviation values have been calculated for each of the BIT based test groups and the ANOVA calculated significance between groupings are shown. As described in Section 6.3, the features that produce a significant difference between the SC and neglect groups (but not between the SC and AMC groups, indicating that normal performance is obtained from the SC group) are the most diagnostically relevant in detecting the condition of neglect.

Analysing the results, the number of cancellations made on the overlay provides a clear significant difference between the neglect and other two groups. Quadrant based cancellation assessment indicates that targets located to the left hand side of the overlay are sensitive to neglect. A number of other dynamic based features also indicate diagnostic ability, in particular the intercancellation and pen movement time regressions across the cancellation sequence. Examination of the mean values from these features show that the normal performance trend is a timing decrease ('speed-up' between cancellations) as the sequence progresses (-ve slope) while the neglect group tend to 'slow-down'. The ratio of left to right hand sided target intercancellation and movement times shows that the times obtained are significantly greater on the left hand side for the neglect group. Cancellation sequence analysis indicates that the neglect group cancel with significantly less regularity (matched against a series of archetypal sequences) than the other two subject groups.

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neg	AMC vs SC	SC vs Neg
<b>Total Time (sec)</b>	<b>22.360</b> 12.130	<b>22.730</b> 9.850	<b>12.640</b> 4.920	0.012	0.03	
<b>Time per Cancellation (sec)</b>	<b>2.440</b> 1.390	<b>1.950</b> 0.786	<b>1.070</b> 0.408	<0.001	0.013	
<b>Num. Of Cancellations (Top Left)</b>	<b>1.960</b> 1.240	<b>2.960</b> 0.264	<b>3</b> 0	0.001		<0.001
<b>Num. Of Cancellations (Bottom Left)</b>	<b>2.030</b> 1.180	<b>2.940</b> 0.225	<b>3</b> 0	<0.001		<0.001
<b>Num. Of Cancellations (Total)</b>	<b>9.600</b> 2.950	<b>11.840</b> 0.520	<b>11.920</b> 0.277	<0.001		<0.001
<b>Sequence Matches</b>	<b>6.460</b> 2.780	<b>7.940</b> 2.420	<b>10.530</b> 2.020	<0.001		0.003
<b>Av. Intercancellation Time, Top Right (Sec)</b>	<b>2.070</b> 1.220	<b>1.850</b> 0.780	<b>1.230</b> 0.430	0.018		
<b>Av. Movement Time, Top Right (Sec)</b>	<b>1.410</b> 0.947	<b>1.230</b> 0.569	<b>0.835</b> 0.332	0.037		
<b>Av. Intercancellation Time, Bottom Left (Sec)</b>	<b>2.440</b> 2.400	<b>2.030</b> 1.040	<b>1.115</b> 0.459	0.033		
<b>Av. Intercancellation Time, Bottom Rt (Sec)</b>	<b>2.240</b> 1.620	<b>1.970</b> 0.894	<b>1.040</b> 0.409	0.006	0.026	
<b>Av. Movement Time, Bottom Rt (Sec)</b>	<b>1.590</b> 1.570	<b>1.340</b> 0.680	<b>0.679</b> 0.210	0.022		
<b>Intercancellation Time Regression</b>	<b>0.220</b> 0.564	<b>-0.002</b> 0.128	<b>-0.018</b> 0.032			0.009
<b>Movement Time Regression</b>	<b>0.228</b> 0.570	<b>0.009</b> 0.115	<b>-0.014</b> 0.029			0.01
<b>Left Side Intercancel. Time (sec)</b>	<b>9.842</b> 9.264	<b>11.393</b> 5.450	<b>5.931</b> 2.273		0.025	
<b>Right Side Intercancel. Time (sec)</b>	<b>11.470</b> 6.615	<b>10.886</b> 4.492	<b>6.571</b> 2.459	0.013	0.02	
<b>Right Side Movement Time (sec)</b>	<b>8.040</b> 6.040	<b>7.380</b> 2.970	<b>4.330</b> 1.440	0.02	0.047	
<b>Left Intercancel. Time per Cancellation (sec)</b>	<b>2.720</b> 2.350	<b>1.930</b> 0.910	<b>0.989</b> 0.378	0.001		
<b>Left Movement Time per Cancellation (sec)</b>	<b>2.120</b> 2.350	<b>1.300</b> 0.742	<b>0.663</b> 0.196	0.006		
<b>Right Intercancel. Time per Cancellation (sec)</b>	<b>2.010</b> 1.170	<b>1.830</b> 0.739	<b>1.110</b> 0.412	0.007	0.024	
<b>L/R Intercancel. Time Per Cancel. Ratio</b>	<b>1.607</b> 1.259	<b>1.065</b> 0.310	<b>0.915</b> 0.237	0.013		0.005
<b>L/R Movement Time per Cancel. Ratio</b>	<b>1.866</b> 1.611	<b>1.071</b> 0.393	<b>0.944</b> 0.264	0.009		0.001

Table 6.14 : Significant BIT grouped OX1 feature results

These results show that the spatial deficits that exist in target cancellation are also demonstrated in constructional aspects of task execution such as timing and task completion strategy.

The significance of the sequence match feature, indicating the ‘normality’ of the cancellation sequence matched against a series archetypal sequences, has led to further investigations into the order of cancellation targets made by test subjects. Table 6.15 details the starting target position of the cancellation sequence made by the test subjects. The positions in the table refer to the target identifiers shown in Figure 5.18.

Position												
	1	2	3	4	5	6	7	8	9	10	11	12
Group												
Neglect	5	0	0	1	3	2	0	0	2	4	7	6
SC	26	1	2	2	2	0	0	1	0	8	6	9
AMC	12	0	0	0	0	0	0	0	0	0	1	0

*Table 6.15 : OX1 task cancellation sequence starting target position*

Significance : AMC vs Neglect:  $p < 0.01$ , Neglect vs Stroke Control:  $p < 0.05$

While the majority of AMC and SC subjects start at targets located in the top left corner of the overlay (normal Western language reading and writing position), neglect patients tend to begin cancellation to the right of the overlay. These differences are significant and are highlighted further by clustering the cancellation targets into the overlay quadrants where they are located. Table 6.16 shows the starting cancellation quadrant results.

Quadrant				
	Top Left	Top Right	Bottom Left	Bottom Right
Group				
Neglect	8	9	3	10
SC	29	6	4	18
AMC	12	1	0	0

*Table 6.16 : OX1 task cancellation sequence starting quadrant*

Significance : AMC vs Neglect:  $p < 0.001$ , AMC vs Stroke Control:  $p < 0.01$ ,  
Neglect vs Stroke Control:  $p < 0.05$

Using the predefined sequences, as described in Section 5.3.4.7, Table 6.17 shows the accumulation of the nearest matched sequence by the three test groups. While no clear pattern emerges from these results (other than a modal number of test subjects in the SC and AMC groups cancelling using sequence number 1, top left hand start, raster vertical sequence), the differences between the AMC and neglect grouping are significant.

		Sequence															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Group																	
Neg		2	2	7	2	0	0	0	1	3	2	6	4	0	1	0	0
SC		17	3	5	4	1	0	0	0	4	3	7	7	2	3	1	0
AMC		6	0	1	0	3	0	0	0	2	0	0	0	1	0	0	0

Table 6.17 : OX1 task best matched model cancellation sequence

Significance : AMC vs Neglect:  $p < 0.01$

A clearer understanding of cancellation strategy can be obtained by grouping the predefined sequences by prominent direction of movement (vertical and horizontal) and traversal method (raster and snake). Table 6.18 shows that a vertical movement based sequence is used by the majority of test subjects in all groups. The AMC group, however, are significantly different to the other groups, mainly because a larger proportion of the group use horizontal based cancellation movement.

		Pattern			
		Raster Vertical	Raster Horizontal	Snake Vertical	Snake Horizontal
Group					
Neglect		13	1	15	1
SC		29	1	21	6
AMC		7	3	2	1

Table 6.18 : OX1 task best matched model cancellation sequence type

Significance : AMC vs Neglect:  $p < 0.01$ , AMC vs Stroke Control:  $p < 0.01$

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Total Time (sec)</b>	<b>20.869</b> 7.494	<b>17.421</b> 9.466	<b>29.006</b> 14.747	<b>24.046</b> 8.342	<b>22.522</b> 10.224	<b>12.640</b> 4.920	6vs4=0.017 6v3=0.001, 5v6=0.041
<b>Time per Cancellation (sec)</b>	<b>2.306</b> 1.270	<b>2.460</b> 1.451	<b>2.534</b> 1.518	<b>2.009</b> 0.691	<b>1.915</b> 0.854	<b>1.070</b> 0.408	2vs6=0.014, 3vs6=0.008
<b>Num. of Cancellations (Top Left)</b>	<b>1.875</b> 1.246	<b>1.182</b> 1.251	<b>2.818</b> 0.603	<b>3</b> 0	<b>2.941</b> 0.343	<b>3</b> 0	1vs3=0.016, 1vs4<0.001 1vs5<0.001, 1vs6=0.001 2vs3<0.001, 2vs4<0.001, 2vs5<0.001, 2vs6<0.001
<b>Num. of Cancellations (Bottom Left)</b>	<b>2.125</b> 1.246	<b>1.091</b> 1.045	<b>2.909</b> 0.302	<b>2.957</b> 0.209	<b>2.941</b> 0.239	<b>3</b> 0	1vs2=0.001, 1vs3=0.025 1vs4=0.003, 1vs5=0.002 1vs6=0.005, 2vs3<0.001 2vs4<0.001, 2vs5<0.001 2vs6<0.001
<b>Num. of Cancellations (Top Right)</b>	<b>4</b> 0	<b>3.182</b> 1.079	<b>4</b> 0	<b>4</b> 0	<b>3.912</b> 0.288	<b>3.923</b> 0.277	1vs2<0.001, 2vs3<0.001 2vs4<0.001, 2vs5<0.001 2vs6<0.001
<b>Num. Of Cancellations (Total)</b>	<b>10</b> 2.391	<b>7.364</b> 3.202	<b>11.727</b> 0.647	<b>11.957</b> 0.209	<b>11.765</b> 0.654	<b>11.923</b> 0.277	1vs2=0.001 1vs4=0.007, 1vs5=0.014 1vs6=0.024, 2vs3<0.001 2vs4<0.001, 2vs5<0.001 2vs6<0.001
<b>Sequence Matches</b>	<b>7.625</b> 2.615	<b>4.546</b> 2.162	<b>7.546</b> 2.583	<b>8.261</b> 2.598	<b>7.735</b> 2.313	<b>10.530</b> 2.020	2vs4=0.001, 2vs5=0.003 2vs6<0.001, 3v6=0.043 5v6=0.008
<b>Av. Intercancellation Time, Bottom Rt (Sec)</b>	<b>2.268</b> 1.831	<b>2.564</b> 2.136	<b>1.903</b> 0.706	<b>1.997</b> 0.841	<b>1.961</b> 0.941	<b>1.040</b> 0.409	2vs6=0.023
<b>Av. Movement Time, Bottom Rt (Sec)</b>	<b>1.639</b> 1.425	<b>1.964</b> 2.234	<b>1.192</b> 0.685	<b>1.291</b> 0.477	<b>1.374</b> 0.797	<b>0.679</b> 0.210	2vs6=0.037
<b>Av. L/R Intercancel. Time Ratio</b>	<b>1.992</b> 1.645	<b>1.353</b> 0.589	<b>1.423</b> 1.278	<b>1.105</b> 0.244	<b>1.092</b> 0.281	<b>1.042</b> 0.237	1vs4=0.039, 1vs5=0.023 1vs6=0.045
<b>Av. LR Move Ratio</b>	<b>2.423</b> 2.096	<b>1.574</b> 0.782	<b>1.687</b> 1.728	<b>1.108</b> 0.328	<b>1.105</b> 0.396	<b>1.074</b> 0.286	1vs4=0.012, 1vs5=0.007 1vs6=0.022
<b>Left Side Intercancel. Time (sec)</b>	<b>8.286</b> 3.819	<b>5.114</b> 5.051	<b>15.702</b> 12.202	<b>11.695</b> 4.282	<b>11.189</b> 6.170	<b>5.931</b> 2.273	2vs3=0.002, 3vs6=0.003
<b>Left Side Movement Time (sec)</b>	<b>6.534</b> 3.676	<b>3.838</b> 3.801	<b>11.683</b> 12.485	<b>7.620</b> 2.733	<b>7.817</b> 5.378	<b>3.980</b> 1.178	2vs3=0.021, 3vs6=0.017
<b>Left Side Drawing Time (sec)</b>	<b>1.470</b> 1.083	<b>1.077</b> 1.710	<b>3.306</b> 1.699	<b>3.480</b> 2.475	<b>2.634</b> 2.208	<b>1.678</b> 1.785	2vs4=0.03
<b>Left Intercancel. Time Per Cancellation (sec)</b>	<b>2.659</b> 1.863	<b>2.310</b> 0.714	<b>3.034</b> 3.275	<b>1.958</b> 0.703	<b>1.912</b> 1.037	<b>0.989</b> 0.378	3vs6=0.01
<b>L/R Intercancel. Time Per Cancel. Ratio</b>	<b>2.062</b> 1.771	<b>1.342</b> 0.552	<b>1.487</b> 1.237	<b>1.062</b> 0.260	<b>1.081</b> 0.340	<b>0.915</b> 0.237	1vs4=0.016, 1vs5=0.013 1vs6=0.006
<b>L/R Movement Time Per Cancel. Ratio</b>	<b>2.317</b> 1.994	<b>1.518</b> 0.731	<b>1.801</b> 1.813	<b>1.062</b> 0.338	<b>1.087</b> 0.433	<b>0.944</b> 0.264	1vs4=0.021, 1vs5=0.018 1vs6=0.015

Table 6.19 : Significant grade-based grouped OX1 feature results

Table 6.19 details the significant features using the grade-based test groupings. Significant differences between groups 1, 2 and 3 (neglect grade-based groupings) indicate the detection of the severity of neglect which is not evident when treating the neglect group as a homogenous performance standard. Differences between groups 4 and 5 indicate performance standards within the stroke control group. Applying these grade-based grouping, the features that are sensitive to neglect severity are the number of cancellations made in the top left and bottom left hand quadrants, the number of cancellations in total and the left side intercancellation and movement times. This indicates the sensitivity of the left hand side targets on the cancellation task to varying levels of neglect. No differences were found between the two SC groupings.

Grade-based neglect levels can also be detected by the cancellation sequence starting position (Table 6.20) and cancellation sequence (Table 6.21). In analysing the cancellation strategy used to complete the overlay, differences were found between the severe and mild neglect groupings. In the Table 6.19 and other tables detailing the grade-based assessment results, Group 6 refers to the AMC group.

Position												
Group	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	1	0	0	0	0	0	2	2	3
2	1	0	0	0	2	0	0	0	2	1	4	1
3	3	0	0	0	1	2	0	0	0	1	1	3
4	12	0	0	1	0	0	0	0	0	2	3	4
5	14	1	1	1	2	0	0	1	0	6	3	4
AMC	12	0	0	0	0	0	0	0	0	0	1	0

Table 6.20 : OX1 task cancellation sequence starting target position – grade-based grouping

Significance: Grp 1 vs 2:  $p < 0.001$ , Grp 1 vs 3:  $p < 0.01$ , Grp 1 vs 6:  $p < 0.02$ ,  
 Grp 2 vs 3:  $p < 0.001$ , Grp 2 vs 4:  $p < 0.01$ , Grp 2 vs 5:  $p < 0.001$ , Grp 2 vs 6:  $p < 0.001$   
 Grp 3 vs 4:  $p < 0.05$ , Grp 3 vs 5:  $p < 0.02$ , Grp 3 vs 6:  $p < 0.01$ , Grp 5 vs 6:  $p < 0.05$

		Sequence															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Group																	
<b>1</b>	0	0	2	0	0	0	0	0	1	0	3	2	0	0	0	0	
<b>2</b>	2	1	4	0	0	0	0	0	1	1	0	2	0	0	0	0	
<b>3</b>	0	1	1	2	0	0	0	1	1	1	2	1	0	1	0	0	
<b>4</b>	7	0	2	1	1	0	0	0	4	0	2	3	1	2	0	0	
<b>5</b>	10	3	3	3	0	0	0	0	0	3	5	4	1	1	1	0	
<b>AMC</b>	6	0	1	0	3	0	0	0	2	0	0	0	1	0	0	0	

Table 6.21 : OX1 task best matched model cancellation sequence – grade-based grouping

Significance: Grp 1 vs 3:  $p < 0.05$ , Grp 1 vs 6:  $p < 0.01$ , Grp 2 vs 6:  $p < 0.01$   
 Grp 3 vs 6:  $p < 0.01$ , Grp 4 vs 6:  $p < 0.02$ , Grp 5 vs 6:  $p < 0.01$

## 6.6.2 OX2 Results

The second OX overlay was introduced to the test battery at a later stage of the trial programme to overcome imbalances in the number of targets positioned in each of the quadrants that are present on the first OX overlay. A total of 15 BIT defined neglect subjects and 25 stroke control patients were included in the analysis of this task. No age matched subjects performed this task.

The significant feature results obtained from the BIT defined groupings are displayed in Table 6.22. A significance value in the table refers to the difference between the neglect and SC groupings. As with the OX1 overlay, differences are obtained in features assessing the number of cancellations made in total and on the left hand side of the overlay. Differences are also obtained from dynamic timing based features extracted from performance on the left hand side of the overlay. Neglect patients also take longer than the SC group to cancel individual targets (Time per cancellation feature) and take a greater time on average to move between cancellations in all quadrants of the overlay (significant in the bottom left and top right quadrants).

No significant results were found from the analysis of the cancellation sequence start position or the sequence starting quadrant.

Analysing the results from the OX2 task with test subjects grouped by grade-based criteria, there are no differences within the neglect groups (1,2 and 3) or the stroke control groups (4

and 5) (Table 6.23). This indicates that the overlay is not sensitive to performance related levels of neglect.

As with the BIT based groups, no significant results were found in either the cancellation sequence start position or the sequence starting quadrant using grade-based groupings.

<b>Feature</b>	<b>Neglect Mean (SC)</b>	<b>SC Mean (SC)</b>	<b>Significance</b>
<b>Time per Cancellation (sec)</b>	<b>3.550</b> 1.746	<b>2.426</b> 1.003	0.013
<b>Num. of Cancellations (Top Left)</b>	<b>2.067</b> 1.335	<b>2.88</b> 0.44	0.008
<b>Num. of Cancellations (Bottom Left)</b>	<b>1.933</b> 1.280	<b>2.92</b> 0.4	0.001
<b>Num. Of Cancellations (Total)</b>	<b>9.4</b> 3.542	<b>11.6</b> 0.957	0.005
<b>Av. Movement Time, Top Rt (Sec)</b>	<b>5.154</b> 3.459	<b>3.415</b> 1.683	0.041
<b>Av. Intercancel. Time, Bottom Left (Sec)</b>	<b>7.346</b> 4.653	<b>4.854</b> 2.085	0.034
<b>Av. Movement Time, Bottom Left (Sec)</b>	<b>5.843</b> 4.278	<b>3.053</b> 1.336	0.005
<b>Left Side Drawing Time (sec)</b>	<b>1.345</b> 1.168	<b>2.603</b> 1.773	0.019
<b>Left Movement Time Per Cancellation (sec)</b>	<b>1.817</b> 1.567	<b>1.031</b> 0.427	0.024

*Table 6.22 : Significant BIT grouped OX2 feature results*

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	Significance
Time per Cancellation (sec)	<b>2.996</b> 1.628	<b>5.070</b> 2.388	<b>2.999</b> 0.685	<b>2.456</b> 0.614	<b>2.426</b> 1.003	2v4=0.013,2v5=0.005
Num. of Cancellations (Bottom Left)	<b>1.6</b> 1.517	<b>1.25</b> 1.5	<b>2.667</b> 0.516	<b>3</b> 0	<b>2.920</b> 0.4	1v4=0.026,1v5=0.28 2v4=0.006,2v5=0.006
Num. of Cancellations (Bottom Right)	<b>2.4</b> 0.894	<b>2.75</b> 0.5	<b>2.833</b> 0.408	<b>3</b> 0	<b>2.813</b> 0.403	1v4=0.024
Number of Path Intersections	<b>2.4</b> 3.912	<b>8</b> 12.329	<b>1</b> 2.450	<b>1.778</b> 2.819	<b>0.688</b> 1.078	2vs5=0.039
Number of Cancellation Duplications	<b>0.2</b> 0.447	<b>1.25</b> 1.893	<b>0.167</b> 0.408	<b>0.333</b> 0.707	<b>0</b> 0	2vs5=0.024
Av. Movement Time, Bottom Left (Sec)	<b>6.943</b> 8.051	<b>7.47</b> 0	<b>5.022</b> 2.225	<b>3.391</b> 0.904	<b>3.053</b> 1.336	1v4=0.046,1v5=0.017
Pre-movement Timing Regression	<b>0.001</b> 0.023	<b>0.01</b> 0.013	<b>-0.003</b> 0.006	<b>0.001</b> 0.005	<b>-0.01</b> 0.01	2vs5=0.048
Av. LR Move Ratio	<b>2.982</b> 3.987	<b>0.431</b> 0.248	<b>1.204</b> 0.501	<b>1.262</b> 0.486	<b>0.874</b> 0.324	1v5=0.043
L/R Movement Time Per Cancel. Ratio	<b>3.074</b> 3.905	<b>0.495</b> 0.157	<b>1.357</b> 0.591	<b>1.246</b> 0.503	<b>0.879</b> 0.287	2v5=0.027

Table 6.23 : Significant grade-based grouped OX2 feature results

### 6.6.3 Albert's Cancellation Task

Unlike the OX2 overlay, all test subjects performed this task. Table 6.24 details the significant results from the this task. While there were significant differences between the AMC and neglect groups in a number of intercancellation based timing features in both the left and right hand sides of the overlays, the only significant differences between the SC and neglect groups occur with features analysing the number of cancellations made on the overlay. Assessment of the number of cancellations made on each side of the overlay reveals a more significant difference on the left hand side, supporting the results from the other cancellation-based tasks.

A further significant feature is the cancellation sequence starting quadrant. The results contained in Table 6.25 show the all AMC subjects start cancelling in the top left hand quadrant of the overlay, whereas neglect performance is biased to starting cancellation at the bottom right hand quadrant, as found in the OX1 overlay. Significant differences between the AMC and the two stroke groups are obtained. The cancellation sequence starting quadrant is also sensitive to levels of neglect (Table 6.26)

Assessment of the results when using grade-based subject groupings, as defined in Section 6.3 shows, that a number of features exist that are sensitive to differences within the neglect groups (1,2 and 3) and the stroke control groups (4 and 5) (Table 6.27). The number of cancellations made in the bottom right hand quadrant of the overlay causes a separation between members of the neglect groupings, as does a series of dynamic movement and intercancellation time features on the left hand side of the overlay, including the ratio of intercancellation, movement and drawing times between the left and right sides of the overlay.

Quadrant				
	Top Left	Top Right	Bottom Left	Bottom Right
<b>Group</b>				
<b>Neglect</b>	8	7	0	15
<b>SC</b>	29	8	1	19
<b>AMC</b>	13	0	0	0

*Table 6.25 : Albert's task cancellation sequence starting quadrant*

Significance : AMC vs Neglect:  $p < 0.001$ , AMC vs Stroke Control:  $p < 0.01$

Quadrant				
	Top Left	Top Right	Bottom Left	Bottom Right
<b>Group</b>				
<b>1</b>	2	1	0	5
<b>2</b>	4	4	0	3
<b>3</b>	2	2	0	7
<b>4</b>	12	3	0	7
<b>5</b>	17	5	1	12
<b>AMC</b>	13	0	0	0

*Table 6.26 : Albert's task cancellation sequence starting quadrant – grade-based grouping*

Significance: Grp 1 vs 3:  $p < 0.05$ , Grp 1 vs 6:  $p < 0.01$ , Grp 2 vs 3:  $p < 0.05$ , Grp 2 vs 6:  $p < 0.001$ , Grp 3 vs 6:  $p < 0.001$ , Grp 4 vs 6:  $p < 0.01$ , Grp 5 vs 6:  $p < 0.01$

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neglect	AMC vs SC	SC vs Neglect
<b>Total Time (sec)</b>	<b>51.95</b> 29.19	<b>53.41</b> 21.63	<b>28.54</b> 6.93	0.09	0.02	
<b>Time per Cancellation (sec)</b>	<b>2.01</b> 1.12	<b>1.6</b> 0.73	<b>0.79</b> 0.202	<.001	0.07	
<b>Num. of Cancellations (Top Left)</b>	<b>4.867</b> 3.04	<b>6.710</b> 1.09	<b>7</b> 0	0.03		<.001
<b>Num. of Cancellations (Top Right)</b>	<b>8.63</b> 2.09	<b>9.61</b> 1.39	<b>10</b> 0	0.029		0.019
<b>Num. of Cancellations (Bottom Left)</b>	<b>7.16</b> 4.74	<b>10.28</b> 1.8	<b>11</b> 0	<.001		<.001
<b>Num. of Cancellations (Bottom Right)</b>	<b>7</b> 1.53	<b>7.66</b> 1.1	<b>7.76</b> 0.438			0.046
<b>Num. of Cancellations (Total)</b>	<b>27.66</b> 10.2	<b>34.38</b> 4.26	<b>35.79</b> 0.438	0.001		<.001
<b>Av. Intercancellation Time, Top Left (Sec)</b>	<b>1.13</b> 1.08	<b>1.42</b> 0.899	<b>0.683</b> 0.23		0.028	
<b>Av. Drawing Time, Top Left (Sec)</b>	<b>0.236</b> 0.19	<b>0.304</b> 0.2	<b>0.137</b> 0.065		0.016	
<b>Av. Intercancellation Time, Top Right (Sec)</b>	<b>1.48</b> 0.766	<b>1.280</b> 0.684	<b>0.635</b> 0.153	0.01	0.07	
<b>Av. Movement Time, Top Right (Sec)</b>	<b>1.042</b> 0.66	<b>0.944</b> 0.588	<b>0.453</b> 0.150	0.008	0.02	
<b>Av. Drawing Time, Top Right (Sec)</b>	<b>0.4</b> 0.277	<b>0.304</b> 0.193	<b>0.153</b> 0.077	0.002		
<b>Av. Intercancellation Time, Bottom Left (Sec)</b>	<b>1.32</b> 1.24	<b>1.04</b> 0.381	<b>0.582</b> 0.129	0.01		
<b>Av. Movement Time, Bottom Left (Sec)</b>	<b>1.03</b> 1.11	<b>0.745</b> 0.292	<b>0.407</b> 0.137	0.014		
<b>Av. Intercancellation Time, Bottom Rt (Sec)</b>	<b>1.53</b> 0.812	<b>1.27</b> 0.699	<b>0.72</b> 0.201	0.002	0.032	
<b>Av. Movement Time, Bottom Rt (Sec)</b>	<b>1.120</b> 0.702	<b>0.994</b> 0.638	<b>0.53</b> 0.162	0.016		
<b>Left Side Intercancel. Time (sec)</b>	<b>9.842</b> 9.264	<b>11.393</b> 5.45	<b>5.931</b> 2.273	0.025		
<b>Right Side Intercancel. Time (sec)</b>	<b>11.470</b> 6.615	<b>10.886</b> 4.492	<b>6.571</b> 2.459	0.013	0.02	
<b>Right Side Movement Time (sec)</b>	<b>8.047</b> 6.049	<b>7.387</b> 2.979	<b>4.332</b> 1.445	0.02	0.047	
<b>Right Intercancel. Time Per Cancellation (sec)</b>	<b>0.765</b> 0.486	<b>0.625</b> 0.264	<b>0.373</b> 0.151	0.002		
<b>Right Movement Time Per Cancellation (sec)</b>	<b>0.536</b> 0.451	<b>0.423</b> 0.177	<b>0.245</b> 0.087	0.003	0.044	

Table 6.24 : Significant BIT grouped Albert's feature results

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Total Time (sec)</b>	<b>59.970</b> 42.878	<b>42.727</b> 23.271	<b>55.350</b> 21.996	<b>53.118</b> 15.935	<b>53.595</b> 24.778	<b>28.540</b> 6.930	1v6=0.047,4v6=0.045 5v6=0.018
<b>Time per Cancellation (sec)</b>	<b>2.213</b> 1.221	<b>1.850</b> 0.718	<b>2.045</b> 1.424	<b>1.625</b> 0.69	<b>1.59</b> 0.77	<b>0.79</b> 0.202	1v6=0.005,2v6=0.045 3v6=0.007
<b>Num. of Cancellations (Top Left)</b>	<b>5</b> 2.976	<b>4.364</b> 3.472	<b>5.273</b> 2.867	<b>6.909</b> 0.426	<b>6.6</b> 1.355	<b>7</b> 0	2v4=0.006,2v5=0.013 2v6=0.013
<b>Num. of Cancellations (Top Right)</b>	<b>8.625</b> 2.504	<b>7.727</b> 2.370	<b>9.546</b> 0.934	<b>9.409</b> 2.131	<b>9.743</b> 0.611	<b>10</b> 0	2v5=0.003,2v6=0.006
<b>Num. of Cancellations (Bottom Left)</b>	<b>6.875</b> 4.486	<b>6.364</b> 5.259	<b>8.182</b> 4.644	<b>10.591</b> 1.141	<b>10.086</b> 2.106	<b>11</b> 0	1v4=0.043, 1v6=0.035 2v4=0.003,2v5=0.006, 2v6=0.003
<b>Num. of Cancellations (Bottom Right)</b>	<b>7.125</b> 1.727	<b>6.091</b> 1.700	<b>7.818</b> 0.405	<b>7.409</b> 1.709	<b>7.829</b> 0.382	<b>7.760</b> 0.438	2v3=0.008,2v4=0.032 2v5<0.001,2v6=0.007
<b>Num. of Cancellations (Total)</b>	<b>27.625</b> 10.809	<b>24.546</b> 11.776	<b>30.818</b> 7.731	<b>34.318</b> 5.277	<b>34.257</b> 3.567	<b>35.790</b> 0.438	2v4=0.001,2v5<0.001 2v6=0.001
<b>Av. Intercancellation Time, Top Left (Sec)</b>	<b>1.759</b> 1.621	<b>0.899</b> 0.914	<b>0.907</b> 0.569	<b>1.629</b> 1.234	<b>1.297</b> 0.589	<b>0.683</b> 0.230	4v6=0.046
<b>Av. Intercancellation Time, Top Right (Sec)</b>	<b>1.455</b> 0.901	<b>1.413</b> 0.531	<b>1.576</b> 0.917	<b>1.229</b> 0.726	<b>1.313</b> 0.666	<b>0.635</b> 0.153	5v6=0.016
<b>Av. Movement Time, Top Right (Sec)</b>	<b>1.086</b> 0.839	<b>0.863</b> 0.318	<b>1.188</b> 0.781	<b>0.892</b> 0.584	<b>0.976</b> 0.596	<b>0.453</b> 0.150	3v6=0.039
<b>Av. Intercancellation Time, Bottom Left (Sec)</b>	<b>2.028</b> 1.815	<b>0.847</b> 0.820	<b>1.298</b> 0.936	<b>1.074</b> 0.392	<b>1.018</b> 0.379	<b>0.582</b> 0.129	1v2=0.008,1v4=0.022 1v5=0.006,1v6<0.001
<b>Av. Movement Time, Bottom Left (Sec)</b>	<b>1.685</b> 1.650	<b>0.616</b> 0.681	<b>0.969</b> 0.818	<b>0.771</b> 0.294	<b>0.729</b> 0.295	<b>0.407</b> 0.137	1v2=0.005,1v4=0.007 1v5=0.002,1v6<0.001
<b>Av. Drawing Time, Bottom Rt (Sec)</b>	<b>0.330</b> 0.137	<b>0.447</b> 0.377	<b>0.322</b> 0.132	<b>0.244</b> 0.204	<b>0.256</b> 0.145	<b>0.155</b> 0.113	2v6=0.005
<b>Left Side Intercancel. Time (sec)</b>	<b>8.286</b> 3.818	<b>5.113</b> 5.050	<b>15.701</b> 12.201	<b>11.754</b> 4.372	<b>11.166</b> 6.079	<b>5.931</b> 2.273	2v3=0.002,3v6=0.003
<b>Left Side Movement Time (sec)</b>	<b>6.533</b> 3.675	<b>3.838</b> 3.800	<b>11.682</b> 12.484	<b>7.597</b> 2.795	<b>7.825</b> 5.298	<b>3.980</b> 1.178	2v3=0.021, 3v6=0.017
<b>Left Side Drawing Time (sec)</b>	<b>1.470</b> 1.082	<b>1.077</b> 1.709	<b>3.306</b> 1.699	<b>3.556</b> 2.505	<b>2.610</b> 2.179	<b>1.677</b> 1.784	2v4=0.023
<b>Left Intercancel. Time Per Cancellation (sec)</b>	<b>0.852</b> 0.706	<b>0.435</b> 0.280	<b>1.479</b> 2.267	<b>0.671</b> 0.237	<b>0.673</b> 0.415	<b>0.329</b> 0.126	3v6=0.016
<b>Left Drawing Time Per Cancellation (sec)</b>	<b>0.176</b> 0.214	<b>0.098</b> 0.115	<b>0.290</b> 0.206	<b>0.201</b> 0.139	<b>0.157</b> 0.146	<b>0.093</b> 0.099	3v6=0.037
<b>Right Movement Time Per Cancellation (sec)</b>	<b>0.558</b> 0.378	<b>0.617</b> 0.646	<b>0.438</b> 0.234	<b>0.432</b> 0.124	<b>0.417</b> 0.203	<b>0.245</b> 0.087	2v6=0.029
<b>L/R Intercancel. Time Per Cancel. Ratio</b>	<b>1.728</b> 2.034	<b>0.394</b> 0.445	<b>1.770</b> 2.167	<b>1.025</b> 0.258	<b>1.114</b> 0.483	<b>0.912</b> 0.227	2v3=0.019
<b>L/R Movement Time Per Cancel. Ratio</b>	<b>2.150</b> 3.100	<b>0.428</b> 0.494	<b>1.911</b> 2.164	<b>1.017</b> 0.331	<b>1.136</b> 0.576	<b>0.942</b> 0.258	1v2=0.036
<b>L/R Drawing Time Per Cancel. Ratio</b>	<b>1.156</b> 1.457	<b>0.398</b> 0.621	<b>1.663</b> 2.190	<b>1.122</b> 0.419	<b>1.127</b> 0.728	<b>0.917</b> 0.290	2v3=0.048

Table 6.27 : Significant grade-based grouping Albert's feature results

## 6.7 Figure Completion Results

Drawing tasks are assessed on a component basis which are defined for each of the shapes or representational drawings. The analysis of these tasks involves the pre-processing of drawing responses and removing those which do not meet the required inclusion criteria. Performance based features are only extracted from responses which meet the inclusion criteria (defined in Section 5.3.6) with the exception of the component count assessment which is applied to both included and excluded responses. Removal of unconventional responses simplifies the analysis and allows consistent and unambiguous identification of individual components and enables feature extraction from specific areas of the drawings.

This section details the significant feature results from the figure completion task. The two grouping criteria (BIT and graded-based) are analysed individually. A summary of the component-based inclusion criteria is provided.

As with the bisection and cancellation tasks, no significant differences were found between age and gender based assessments.

### 6.7.1 Overlay 1 (Diamond) BIT based patient groupings

In the first pair of overlays a diamond shape is required to be completed, first drawing into the right and then, on the second overlay, the left hand side of the page. The inclusion criteria for test responses of this simple shape are the two edge components in the correct vertical orientation (i.e. a mirror image is drawn). Intersection is not necessary as the corner distance error is calculated as a performance feature. Table 6.28 shows the percentages of each group that do not meet the required drawing performance. Whereas all of the AMC group drew responses which comprised the correct number of components, approximately a quarter of the stroke groups (SC and neglect) are excluded from the analysis.

The significant features from the first completion overlay are detailed in Table 6.29. A number of dynamic based features separate the SC and neglect groups: the increased mean number of pen lifts made by neglect patients indicate that the neglect group treat the drawing of individual components separately, rather than as an entire shape drawn in one continuous movement.

Group	% of Group Removed
Neglect	23.3
Stroke Control	23.7
Age Matched Control	0

Table 6.28 : Percentage of group excluded from analysis, figure completion, overlay 1 (diamond)

The pen movement timings between components are also significantly greater for the neglect group. One explanation for this is that neglect patients spend a greater time in a planning phase between drawing individual components. The pressure used by neglect subjects is significantly less, as is the height of the image drawn.

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neglect	AMC vs SC	SC vs Neglect
Pen Lifts	4.087 6.178	1.777 1.881	2.166 1.114			0.045
Mean Pressure (range -128 to 128)	-2.801 46.742	31.714 41.254	3.910 54.509			0.011
Image Height (mm)	87.357 6.386	91.737 4.070	91.628 2.299	0.038		0.002
Mean Pen Velocity (mm/sec)	4.744 3.223	4.204 2.316	6.870 5.230		0.033	
Movement Time Between Components (sec)	2.407 1.255	1.510 0.525	1.185 0.250	0.006		0.02

Table 6.29 : Significant BIT grouped overlay 1 (diamond) figure completion feature results

### 6.7.2 Overlay 2 (Diamond) BIT based patient groupings

In this overlay, the test subject is presented with the right hand side of the image and has to draw into the left side of the visual field (the inattentive side of the visual field for right CVA neglect subjects). Using the same response inclusion criteria as the first overlay in the figure completion task, Table 6.30 shows the percentage of group responses removed from the analysis. While the number of exclusions from the SC and AMC groups are similar to the first overlay, a larger percentage of the neglect responses are inaccurate. The fact that almost half of the neglect population are unable to produce an accurate response indicates that either the

diamond completion is not a particularly good task for the diagnosis of neglect or that the inclusion criteria are too strictly defined.

From the group percentage inclusion, it is possible to conclude that neglect patients produce a more accurate response when copying from their inattentive field (overlay 1), than when drawing into it (overlay 2).

<b>Group</b>	<b>% of Group Removed</b>
<b>Neglect</b>	46.6
<b>Stroke Control</b>	18.9
<b>Age Matched Control</b>	0

*Table 6.30 : Percentage of group excluded from analysis, figure completion, overlay 2 (diamond)*

Examination of the results shows that there are no significant differences between the SC and neglect groupings, indicating that those neglect subjects that are able to produce drawings with the correct number of components perform normally (as per the control groups) across the range of features (Table 6.31).

<b>Feature</b>	<b>Neglect Mean (SD)</b>	<b>SC Mean (SD)</b>	<b>AMC Mean (SD)</b>	<b>AMC vs Neglect</b>	<b>AMC vs SC</b>	<b>SC vs Neglect</b>
<b>Mean Point 2 Positional Displacement (mm)</b>	<b>11.979</b> 6.982	<b>8.553</b> 5.487	<b>3.765</b> 4.113	0.001	0.032	
<b>Mean Pen Velocity (mm/sec)</b>	<b>5.542</b> 2.488	<b>5.282</b> 3.219	<b>9.781</b> 12.031		0.040	
<b>Movement Time Between Components (sec)</b>	<b>2.268</b> 0.820	<b>1.620</b> 0.478	<b>1.020</b> 0.353	0.018		

*Table 6.31 : Significant BIT grouped overlay 2 (diamond) figure completion feature results*

### 6.7.3 Overlay 1/Overlay 2 Result Ratio

By forming a ratio from pairs of feature results from the first two completion overlays, an analysis can be formed to identify significant differences between performance within the different halves of the visual field. Because the ratio analyses performance *within* a test subject, measuring differences between overlay pairs, all ratios are normalised and therefore not affected by the values of the results in comparison with other test subjects. Using two results, the first from a feature extracted from overlay 1 (drawing on the right hand side) and the second of the same feature extracted from overlay 2 (drawing on the left hand side) a ratio is formed by dividing the overlay 1 result by the overlay 2 result. A ratio of less than 1.0 indicates that the performance result for drawing on the right hand side of the overlay was less than that when drawing to the left hand side. Performance differences between the two overlays are shown by the deviation away from 1.0. Table 6.32 show the features with significant differences between left and right overlays.

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neglect	AMC vs SC	SC vs Neglect
<b>Image Width Ratio</b>	<b>0.847</b> 0.317	<b>1.096</b> 0.355	<b>1.029</b> 0.189			0.047
<b>Angle Between Components Ratio</b>	<b>1.252</b> 0.394	<b>0.979</b> 0.198	<b>1.002</b> 0.128	0.029		0.002
<b>Total length drawn Ratio</b>	<b>0.882</b> 0.185	<b>1.036</b> 0.160	<b>1.038</b> 0.136	0.041		0.008

Table 6.32 : Significant BIT grouped overlay 1/overlay 2 figure completion feature results ratios (diamond)

The results show that when drawing in the right hand visual field (the attentive field) the neglect group on average draw less (in terms of total length drawn) and compress the width of the drawn image.

This width compression of a drawing copied *from* the inattentive field is noted in the majority of the neglect responses and is clearly a characteristic of neglect performance. The performance ratio between overlays for the two control groups is consistent.

### 6.7.4 Overlay 3 (Man) BIT based patient groupings

The significant results from the third completion overlay, requiring the right hand completion of a representational drawing of a man, are contained in Table 6.34. The significant features are similar to overlay 1 (Pen lifts, pressure, movement time) thus indicating a common set of results which can be used to characterise groupings on the basis of the completions drawn in the right hand visual field.

The inclusion criteria for overlays 3 and 4 are for the drawn image to contain four recognisable components: a head, an arm, a body and a leg. These are defined in Section 5.3.6. The components must be drawn as a horizontal inversion of the presented image. The number of components drawn by all group members (prior to response exclusion) is significant, with the neglect group producing fewer recognisable defined components. This is also shown in Table 6.33 in that 70% of the neglect group produce responses that are excluded from the test analysis. Again, the high exclusion rates indicate that this drawing task is too difficult for the neglect and stroke control groups. An evaluation of shape complexity is provided in Section 6.10.

The diagnostic power of dynamic features is demonstrated by the data contained in Table 6.34. Analysing responses that all contain the same number of components and therefore on a static analysis level all 'pass' the drawing assessment, three dynamic based features detect differences in performance.

<b>Group</b>	<b>% of Group Removed</b>
<b>Neglect</b>	70
<b>Stroke Control</b>	49.1
<b>Age Matched Control</b>	7.6

*Table 6.33 : Percentage of group excluded from analysis, figure completion, overlay 3 (man)*

<b>Feature</b>	<b>Neglect Mean (SD)</b>	<b>SC Mean (SD)</b>	<b>AMC Mean (SD)</b>	<b>AMC vs Neglect</b>	<b>AMC vs SC</b>	<b>SC vs Neglect</b>
<b>Pen Lifts</b>	<b>13.777</b> 9.896	<b>6.931</b> 5.188	<b>8.454</b> 5.027			0.018
<b>Pen Movement Time (sec)</b>	<b>17.450</b> 12.086	<b>8.554</b> 7.655	<b>7.944</b> 7.726			0.028
<b>Mean Pressure (range –128 to 128)</b>	<b>-37.129</b> 36.910	<b>27.091</b> 45.317	<b>20.752</b> 59.422	0.028		0.003
<b>Number of Components</b>	<b>2.366</b> 1.938	<b>3.438</b> 1.742	<b>4.750</b> 0.866	<0.001		0.022

*Table 6.34 : Significant BIT grouped overlay 3 (man) figure completion feature results*

### 6.7.5 Overlay 4 (Man) BIT based patient groupings

Table 6.35 details the number of incorrect responses completing the representational shape of man to the left of the overlay. In comparison with overlay 3, the results show that neglect patients have greater difficulty in producing an accurate response when drawing the shape in the inattentive field. The very high exclusion rate from the neglect groups again indicates that this shape is too difficult for both the neglect and stroke control groups. Assessment against the results obtained from the simple diamond completion task show that these effects are more pronounced when copying a more complex (an increased number of components) shape.

A similar number of accurate responses to those produced for overlay 3 were obtained from the two control groups indicating that control performance within a task using the same shapes does not vary according to which side is being drawn. These findings are supported by the performances obtained from overlays 1 and 2.

<b>Group</b>	<b>% of Group Removed</b>
<b>Neglect</b>	83.3
<b>Stroke Control</b>	52.6
<b>Age Matched Control</b>	15.2

*Table 6.35 : Percentage of group excluded from analysis, figure completion, overlay 4 (man)*

Significant results from overlay 4 are detailed in Table 6.36. Along with a difference in the number of components drawn (analysed with the entire test population), two dynamic features (mean pressure and peak pen acceleration) produce differences between SC and neglect test groups when the correct number of components are drawn in a response.

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neglect	AMC vs SC	SC vs Neglect
<b>Pen Movement Time (sec)</b>	<b>20.060</b> 13.760	<b>10.822</b> 8.679	<b>6.567</b> 4.056	0.024		
<b>Mean Pressure</b>	<b>-48.135</b> 24.040	<b>19.550</b> 57.819	<b>5.915</b> 52.111			0.042
<b>Mean Corner Positional Displacement (mm)</b>	<b>43.432</b> 20.669	<b>47.184</b> 18.765	<b>67.772</b> 12.035		0.015	
<b>Number of Components</b>	<b>1.666</b> 1.768	<b>3.175</b> 1.909	<b>4</b> 1.809	<0.001		0.001
<b>Mean Point 1 Positional Displacement (mm)</b>	<b>9.238</b> 2.882	<b>8.238</b> 4.472	<b>12.382</b> 2.277		0.030	
<b>Mean Point 2 Positional Displacement (mm)</b>	<b>9.571</b> 3.673	<b>9.263</b> 4.621	<b>15.372</b> 2.560		0.002	
<b>Mean Point 3 Positional Displacement (mm)</b>	<b>7.883</b> 5.086	<b>9.366</b> 5.134	<b>13.956</b> 2.655		0.048	
<b>Mean Point 4 Positional Displacement (mm)</b>	<b>9.276</b> 6.254	<b>10.207</b> 4.254	<b>14.566</b> 2.810		0.036	
<b>Peak Pen Acceleration (mm/sec/sec)</b>	<b>13.536</b> 6.680	<b>2.261</b> 8.823	<b>0.752</b> 8.833	0.035		0.033

*Table 6.36 : Significant BIT grouped overlay 4 (man) figure completion feature results*

### 6.7.6 Overlay 3/Overlay 4 Result Ratio

In forming ratios between results extracted from overlays 3 and 4, no significant differences were noted. An explanation for this is that the high neglect group exclusion rate, caused by the use of a more complex shape, removes all patients except those with mild neglect. These patients, while producing differences detected by a particular dynamic feature within a single overlay, draw with performance characteristics that do not differ between responses made to the left and right sides of the visual field and therefore form ratios that do not differ significantly from the two control groups.

### 6.7.7 Overlay 5 (House) BIT based patient groupings

As with the representational shape of the man, the two overlays requiring the image of a house to be completed causes a high a number of image exclusions (Table 6.37) indicating that as the number of components within a shape increases, a less accurate response will be drawn (see Section 6.10). With overlays 5 and 6, five components are required to be drawn to meet the inclusion criteria: a wall/floor section, window 1, window 2, a door and a roof section.

The high error rate for overlay 5, requiring the completion of the house to the left of the overlay, is shown in Table 6.38 where the neglect group draw significantly fewer components than the two control groups. Again, the neglect patients who successfully draw the correct number of components in the house image perform as per the control group. The neglect in this task is investigated further in Sections 6.16 and 6.17 by the dividing the test population by grade-based groupings.

Group	% of Group Removed
Neglect	76.6
Stroke Control	47.3
Age Matched Control	15.3

Table 6.37 : Percentage of group excluded from analysis, figure completion, overlay 5 (house)

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neglect	AMC vs SC	SC vs Neglect
Pen Movement Time (sec)	28.357 15.426	19.329 15.304	9.706 4.082	0.021		
Total Execution Time (sec)	52.945 21.698	38.25 20.939	23.210 10.809	0.008		
Number of Components	2.2 1.972	3.754 1.672	4.833 0.577	<0.001		<0.001

Table 6.38 : Significant BIT grouped overlay 5 (house) figure completion feature results

### 6.7.8 Overlay 6 (House) BIT based patient groupings

The number of response exclusions when drawing the house image to the left hand side is almost identical to overlay 5 for all subject groups (Table 6.39). Again the number of components drawn by the neglect group is significantly fewer. The overall drawing time in the inattentive field is the only feature which is sensitive to detecting cases of neglect that have successfully drawn the correct number of components (Table 6.40).

Group	% of Group Removed
Neglect	76.6
Stroke Control	43.8
Age Matched Control	15.3

*Table 6.39 : Percentage of group excluded from analysis, figure completion, overlay 6 (house)*

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neglect	AMC vs SC	SC vs Neglect
<b>Pen Movement Time (sec)</b>	<b>24.642</b> 16.229	<b>20.240</b> 11.362	<b>10.609</b> 4.646	0.036		
<b>Total Execution Time (sec)</b>	<b>44.18</b> 24.406	<b>37.697</b> 17.963	<b>22.12</b> 8.359	0.036		0.042
<b>Number of Components</b>	<b>2.4</b> 1.975	<b>3.649</b> 1.846	<b>4.916</b> 0.288	<0.001		0.007

*Table 6.40 : Significant BIT grouped overlay 6 (house) figure completion feature results*

These results show that the image is too complex for testing the entire performance range of neglect, as it is only copied successfully by the neglect subjects with mild symptoms.

### 6.7.9 Overlay 5/Overlay 6 Result Ratio

As with the left/right hand drawing response ratios formed from features extracted from overlays 3 and 4, no significant differences were noted for overlays 5 and 6. This again may

be attributed to the complexity of the representational shape used and the high component-based exclusion rate.

### 6.7.10 Overlay 1 (Diamond) grade-based patient groupings

The significant results from grade-based grouping as defined in Section 6.3 for the first figure completion overlay are shown in Table 6.41. As with the cancellation analysis, differences within the neglect groups (1, 2 and 3) and SC groups (4 and 5) indicate sensitivity to performance grading.

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
Pen Lifts	<b>4.333</b> 5.391	<b>7.285</b> 9.586	<b>1.7</b> 1.251	<b>2.533</b> 2.416	<b>1.379</b> 1.473	<b>2.166</b> 1.114	2v3=0.022, 2v5=0.002 2v6=0.031
Total Execution Time (sec)	<b>17</b> 22.708	<b>10.814</b> 10.688	<b>6.858</b> 2.390	<b>7.879</b> 2.703	<b>6.931</b> 4.892	<b>5.7</b> 2.565	1v6=0.047
Image Height (mm)	<b>83.986</b> 9.886	<b>88.061</b> 3.891	<b>89.135</b> 4.352	<b>91.917</b> 2.933	<b>91.645</b> 4.658	<b>91.628</b> 2.299	1v4=0.009, 1v5=0.006 1v6=0.017
Point 1 and 3 X- Axis Displacement (mm)	<b>7.09</b> 8.568	<b>1.439</b> 1.066	<b>3.11</b> 3.984	<b>0.966</b> 0.585	<b>2.808</b> 2.798	<b>1.192</b> 1.127	1v2=0.047, 1v4=0.003 1v6=0.006
Movement Time Between Components (sec)	- -	<b>1.336</b> 0.803	<b>2.943</b> 1.109	<b>1.358</b> 0.411	<b>1.645</b> 0.599	<b>1.185</b> 0.250	2v3=0.014, 3v4=0.001 3v5=0.006, 3v6<0.001

Table 6.41 : Significant grade-based grouped overlay 1 (diamond) figure completion results

The results show that the dynamic features of the number of pen lifts and movement time between components along with the static horizontal displacement between reference points 1 and 3 are able to grade performance within the neglect population.

### 6.7.11 Overlay 2 (Diamond) grade-based patient groupings

No significant differences were found within the neglect groups or SC groups using the grade-based groupings for the second completion overlay. Differences did exist between AMC and the moderate neglect group (group 2) on two separate features as detailed in Table 6.42.

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Mean Point 2 Positional Displacement (mm)</b>	<b>15.687</b> 11.419	<b>14.734</b> 4.982	<b>10.135</b> 6.948	<b>10.059</b> 7.298	<b>7.845</b> 4.029	<b>3.765</b> 4.113	2v6=0.014
<b>Movement Time Between Components (sec)</b>	<b>1.92</b> 1.216	<b>3.155</b> 0.374	<b>2</b> 0.572	<b>1.695</b> 0.561	<b>1.671</b> 0.334	<b>1.020</b> 0.353	2v4=0.049, 2v5=0.043 2v6=0.002

*Table 6.42 : Significant grade-based grouped overlay 2 (diamond) figure completion feature results*

### 6.7.12 Overlay 1/Overlay 2 grade-based result ratio

As with the analysis of this task using standard BIT defined groupings, a ratio is formed between feature results obtained from the left and right hand side drawing overlays (right hand drawing/left hand drawing). This enables performance differences between the two visual fields to be assessed. A ratio of greater than 1.0 indicates a larger result was obtained for the drawing made on the right hand side.

The results in Table 6.43 show that when the severe and moderate neglect group are able to draw both halves of the image, their movement and total execution time on the right hand side (their intact field) is greater. This indicates that when drawing in their intact field, neglect subjects are not able to respond as quickly to the positional feedback of the pen.

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Movement Time Ratio</b>	<b>6.617</b> 9.964	<b>2.627</b> 1.604	<b>0.919</b> 0.801	<b>1.437</b> 1.187	<b>1.049</b> 0.778	<b>1.133</b> 0.465	1v3=0.008, 1v4=0.018 1v5=0.004, 1v6=0.007
<b>Total Execution Time Ratio</b>	<b>4.288</b> 5.656	<b>1.706</b> 0.930	<b>1.077</b> 0.314	<b>1.313</b> 0.902	<b>1.154</b> 0.417	<b>1.313</b> 0.376	1v3=0.001, 1v4=0.002 1v5=0.001, 1v6=0.002
<b>Number of Components Drawn Ratio</b>	<b>1.722</b> 1.678	<b>1.018</b> 0.412	<b>1</b> 0.223	<b>0.990</b> 0.228	<b>1.123</b> 0.447	<b>1</b> -	1v4=0.041

*Table 6.43 : Significant grade-based grouped overlay 1/overlay 2 (diamond) figure completion feature results ratios*

### 6.7.13 Overlay 3 (Man) grade-based patient groupings

Table 6.44 shows the results of the right hand side drawing of the man. The number of components drawn is significantly different between all three of the neglect based groups. This confirms the difficulty that neglect patients (especially severe neglect cases) have in constructing an accurate drawing of more complex shapes.

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
Mean Pressure (range -128 to 128)	-73.327 -	-25.940 41.814	-34.826 38.879	15.394 54.194	29.467 40.174	20.752 59.422	3v5=0.044
Number of Components	1.375 1.598	1.636 1.804	3.818 1.470	3.136 1.780	3.588 1.725	4.75 0.866	1v3=0.025, 1v5=0.012 1v6<0.001, 2v3=0.033 2v5=0.012, 2v6<0.001

\*ANOVA Significance not calculated on Group 1 due to single group membership.

*Table 6.44 : Significant grade-based grouped overlay 3 (man) figure completion feature results*

### 6.7.14 Overlay 4 (Man) grade-based patient groupings

As with the left hand side diamond drawing task (overlay 2), no differences exist within the neglect (groups 1,2 and 3) or SC (4 and 5) groups (Table 6.45). This indicates that when drawing into the inattentive field all neglect patients who have the ability to complete the task perform similarly.

### 6.7.15 Overlay 3/Overlay 4 grade-based result ratio

A single performance feature was found to be significant in comparing the left and right hand responses. The mean displacement error of point 2 (the junction between the arm and the head components of the man) is able to provide a performance differential between the two performance graded SC groups (Table 6.46).

Significant differences of groups 1 and 2 against the other grade-based groups could not be calculated as only a single subject in each of these groups completed both overlays.

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
Total Execution Time (sec)	11.76 -	51.75 -	28.84 13.227	28.399 11.589	20.977 11.827	14.999 6.704	4v6=0.044
Mean Corner Displacement Error (mm)	53.395 -	75.058 -	29.570 4.055	45.601 16.886	48.106 21.035	67.772 12.035	3v6=0.019
Number of Components	1.375 1.685	1.090 1.640	2.454 1.809	3.318 1.835	3.029 1.976	4 1.809	1v6=0.023, 2v4=0.023 2v5=0.047, 2v6=0.002
Mean Point 2 Positional Displacement (mm)	9.320 -	15.750 -	7.595 1.418	7.877 4.643	10.193 4.688	15.372 2.56	4v6=0.005
Mean Point 3 Positional Displacement (mm)	10.116 -	15.705 -	4.532 1.343	9.001 4.240	9.741 5.915	13.936 2.655	3v6=0.046
Mean Point 4 Positional Displacement (mm)	12.217 -	18.567 -	5.198 2.407	10.137 2.178	10.112 5.378	14.566 2.81	3v6=0.007
Peak Pen Acceleration (mm/sec/sec)	8.56 -	11.7 -	15.806 8.213	4.018 5.170	2.326 9.750	0.752 8.833	3v6=0.037

\*ANOVA Significance not calculated on Group 1 and Group 2 due to single group membership.

Table 6.45 : Significant grade-based grouped overlay 4 (man) figure completion feature results

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
Mean Point 2 Positional Displacement Ratio	1.634 -	0.728 -	1.368 1.004	1.535 0.435	0.916 0.390	0.817 0.238	4v5=0.026, 4v6=0.01

\*ANOVA Significance not calculated on Group 1 and Group 2 due to single group membership.

Table 6.46 : Significant grade-based grouped overlay 3/overlay 4 (man) figure completion feature results ratios

### 6.7.16 Overlay 5 (House) grade-based patient groupings

The significant results from the right hand side drawing of the house are contained in Table 6.47. Unlike the other two right hand drawings tasks (Overlays 1 and 3) the high complexity of the image to be copied causes no significant differences within the performance graded neglect or SC groups. None of the severe neglect subjects and only a single moderate neglect subject are able to complete this task.

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Total Execution Time (sec)</b>	- -	<b>77.62</b> -	<b>48.833</b> 20.565	<b>42.802</b> 17.119	<b>36.478</b> 23.169	<b>23.21</b> 10.809	3v6=0.044
<b>Number of Components</b>	<b>1.5</b> 1.603	<b>1.454</b> 1.634	<b>3.454</b> 2.018	<b>3.545</b> 1.969	<b>3.852</b> 1.479	<b>4.833</b> 0.577	1v4=0.043, 1v5=0.005 1v6<0.000, 2v4=0.011 2v5=0.001, 2v6<0.001
<b>Mean Pen Velocity (mm/sec)</b>	- -	<b>1.930</b> -	<b>2.625</b> 0.858	<b>2.585</b> 1.297	<b>3.407</b> 1.591	<b>5.969</b> 4.842	4v6=0.024

\*ANOVA Significance not calculated on Group 2 due to single group membership. No successful test attempts from Group 1 were obtained

*Table 6.47 : Significant grade-based grouped overlay 5 (house) figure completion feature results*

### 6.7.17 Overlay 6 (House) grade-based patient groupings

Again no differences exist when drawing in the inattentive field within the neglect or SC groups (Table 6.48).

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Pen Movement Time (sec)</b>	<b>10.45</b> -	<b>43.99</b> -	<b>23.612</b> 15.805	<b>25.726</b> 9.796	<b>17.445</b> 11.243	<b>10.609</b> 4.646	4v6=0.004
<b>Total Drawing Time (sec)</b>	<b>6.55</b> -	<b>20.48</b> -	<b>21.964</b> 11.463	<b>20.503</b> 6.281	<b>16.047</b> 9.637	<b>11.008</b> 5.117	4v6=0.039
<b>Total Execution Time (sec)</b>	<b>16.99</b> -	<b>64.46</b> -	<b>45.562</b> 24.565	<b>46.211</b> 12.767	<b>33.585</b> 18.917	<b>22.12</b> 8.359	3v6=0.043, 4v6=0.003
<b>Mean Pen Velocity (mm/sec)</b>	<b>5.51</b> -	<b>2.78</b> -	<b>3.066</b> 1.001	<b>2.819</b> 1.129	<b>3.685</b> 1.710	<b>5.430</b> 2.893	4v6=0.011
<b>Number of Components</b>	<b>1.625</b> 1.767	<b>1.727</b> 1.902	<b>3.636</b> 1.689	<b>3.409</b> 2.039	<b>3.764</b> 1.741	<b>4.916</b> 0.288	1v5=0.031, 1v6=0.001 2v5=0.014, 2v6<0.001

\*ANOVA Significance not calculated on Group 1 and Group 2 due to single group membership.

*Table 6.48 : Significant grade-based grouped overlay 6 (house) figure completion feature results*

### 6.7.18 Overlay 5/Overlay 6 grade-based result ratio

Due to the complexity of the shape used in this task and the uniform feature results it produces – if a test subject can complete this task, they will perform as per the age matched control group - the only significant difference between left and right hand drawing performance is in the number of components drawn by each group. Table 6.49 shows a difference between the severe (group 1) and moderate (group 2) neglect groups in the ratio of components drawn on the right to those drawn on the left. On average the severe group draw more components on the left, while the moderate group conform to the expected neglect behaviour and draw more on the right hand side (the intact visual field).

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Number of Components Ratio</b>	<b>0.916</b> 0.569	<b>1.75</b> 0.987	<b>1.144</b> 0.512	<b>1.020</b> 0.412	<b>1.052</b> 0.239	<b>1.025</b> 0.092	1v2=0.038, 2v4=0.006 2v5=0.005, 2v6=0.01

*Table 6.49 : Significant grade-based grouped overlay 5/overlay 6 (house) figure completion feature results ratios*

## 6.8 Figure Copying Results

This section presents the results from the four figure copying tasks. As with the figure completion task, images that do not have the requisite number of components are removed from the analysis. The criteria for inclusion is presented with each analysis.

No effects of gender and age were found from features extracted from this task.

### 6.8.1 Square BIT based patient groupings

The first overlay requires the test subject to copy a simple square shape. The inclusion criteria for the square is that the drawn image has four sides (Table 6.50).

<b>Group</b>	<b>% of Group Removed</b>
<b>Neglect</b>	36.6
<b>Stroke Control</b>	14.1
<b>Age Matched Control</b>	0

*Table 6.50 : Percentage of group excluded from analysis, figure copying, square*

Feature-based performance of the included responses is uniform across all three test groupings. Only the movement to drawing time ratio provides a significant difference between SC and neglect groupings - indeed, neglect subjects on average spend longer moving the pen above the tablet surface than drawing on the overlay (Table 6.51).

The increased movement times between components indicates that the neglect subjects adopt a component-based drawing strategy - planning and then drawing individual sides of the shape - which is also indicated by the increased number of pen lifts from the tablet surface. The control groups tend to draw the shape in a more continuous motion.

<b>Feature</b>	<b>Neglect Mean (SD)</b>	<b>SC Mean (SD)</b>	<b>AMC Mean (SD)</b>	<b>AMC vs Neglect</b>	<b>AMC vs SC</b>	<b>SC vs Neglect</b>
<b>Movement to Drawing Time Ratio</b>	<b>1.544</b> 2.298	<b>0.542</b> 0.427	<b>0.859</b> 0.438			0.017
<b>Mean Pen Velocity (mm/sec)</b>	<b>2.879</b> 1.669	<b>2.896</b> 1.890	<b>4.816</b> 3.495	0.046	0.018	

*Table 6.51 : Significant BIT grouped square figure copying feature results*

## 6.8.2 Cross BIT based patient groupings

The second of the figure copying tasks requires the copying of a constructional cross shape. Drawings are excluded if the 5 sub-boxes that form the cross are not present in a response. This shape is more sensitive to neglect as can be seen from the increased number of exclusions detailed in Table 6.52. The number of exclusions from the two control groups is similar to the square copying task.

Group	% of Group Removed
Neglect	53.3
Stroke Control	21.1
Age Matched Control	0

Table 6.52 : Percentage of group excluded from analysis, figure copying, cross

A range of dynamic features provide a separation between the SC and neglect groupings which again show the component based drawing strategy of the neglect group (pen lifts and pen movement time) (Table 6.53). Related to this strategy for the neglect subjects is an increased mean total execution time for the drawing which encompasses slower drawing phases and an increased time spent moving between components.

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neglect	AMC vs SC	SC vs Neglect
<b>Pen Lifts</b>	<b>11.214</b> 8.477	<b>5.866</b> 4.148	<b>7.583</b> 2.906			0.003
<b>Pen Movement Time (sec)</b>	<b>16.297</b> 16.071	<b>7.047</b> 5.875	<b>5.918</b> 2.943	0.009		0.002
<b>Pen Drawing Time (sec)</b>	<b>14.605</b> 5.596	<b>11.960</b> 7.065	<b>7.472</b> 2.965	0.016		
<b>Total Execution Time (sec)</b>	<b>30.905</b> 17.343	<b>19.068</b> 10.113	<b>13.64</b> 4.987	0.001		0.003
<b>Number of Components</b>	<b>3.433</b> 1.851	<b>4.245</b> 1.561	<b>5</b> 0	0.013		

Table 6.53 : Significant BIT grouped cross figure copying feature results

### 6.8.3 Star BIT based patient groupings

The star copying task inclusion criterion is that the drawn shape contains five vertices (defined in Section 5.3.6.2) . The increased complexity of the shape is demonstrated by the number of responses excluded from the analysis in all test subject groups (Table 6.54). An examination of shape complexity vs exclusion rates is presented in Section 6.10.

Group	% of Group Removed
Neglect	60
Stroke Control	52.6
Age Matched Control	30.7

Table 6.54 : Percentage of group excluded from analysis, figure copying, star

Unlike the complex shapes (shapes with a large number of sides and interconnecting components, e.g. the house) used in the figure completion task, which produced uniform performance results from all test subjects who drew the required number of components, the star shape produces dynamic feature based separations between neglect and control groups (Table 6.55). The increased number of pen lifts, movement and execution time indicate that the neglect group, whilst producing an accurate spatial copy of the drawing, perform more slowly during construction. This highlights that while they are able to observe all components of the star, they are unable to visualise the complete construction strategy needed to reproduce the shape. The star, therefore, is the shape most sensitive to the detection of neglect *when the drawing contains the correct number of components*.

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neglect	AMC vs SC	SC vs Neglect
<b>Pen Lifts</b>	<b>16.833</b> 14.440	<b>5.851</b> 4.045	<b>7.444</b> 2.920	0.029		0.001
<b>Pen Movement Time (sec)</b>	<b>43.850</b> 42.089	<b>16.815</b> 18.248	<b>10.785</b> 10.104	0.015		0.011
<b>Total Execution Time (sec)</b>	<b>59.915</b> 45.355	<b>29.220</b> 23.169	<b>20.215</b> 13.577	0.010		0.012
<b>Mean Pressure (range -128 to 128)</b>	<b>-20.506</b> 43.350	<b>25.738</b> 54.144	<b>12.406</b> 52.435			0.038
<b>Image Height (mm)</b>	<b>41.841</b> 21.376	<b>31.624</b> 8.411	<b>27.835</b> 2.750	0.042		
<b>Image Height Error (mm)</b>	<b>15.793</b> 20.618	<b>6.157</b> 7.320	<b>2.074</b> 1.872	0.031		

Table 6.55 : Significant BIT grouped star figure copying feature results

### 6.8.4 Cube BIT based patient groupings

The final shape used in the copying task is the 3-d cube. Table 6.56 details the number of subjects failing to meet the inclusion criterion of drawing a shape containing six vertices - four forming the square and two others forming the top left and bottom right of the three-dimensional section of the drawing. The cube is the most complex shape used as shown by the highest exclusion rate of the figure copying tasks.

<b>Group</b>	<b>% of Group Removed</b>
<b>Neglect</b>	70
<b>Stroke Control</b>	56.1
<b>Age Matched Control</b>	30.7

*Table 6.56 : Percentage of group excluded from analysis, figure copying, cube*

The cube has a similar exclusion rate to the star, however test subjects accurately constructing the shape do not produce the variety of significant dynamic feature responses as found in the star.

A single feature, number of pen lifts, separates the SC and neglect groupings (Table 6.57). This again shows that neglect subjects approach the construction of the shape on a component basis, removing the pen before entering the planning phase of the next component to be drawn. Unlike the star copying task, the cube causes uniform performance from all subject groupings. This indicates the fine balance required in shape selection which must not make the shape too easy or too difficult to cause uniform results within and between test subject groupings (Section 6.10)

<b>Feature</b>	<b>Neglect Mean (SD)</b>	<b>SC Mean (SD)</b>	<b>AMC Mean (SD)</b>	<b>AMC vs Neglect</b>	<b>AMC vs SC</b>	<b>SC vs Neglect</b>
<b>Pen Lifts</b>	<b>14.333</b> 9.151	<b>9.16</b> 3.236	<b>8.666</b> 3.316			0.034
<b>Number of Components</b>	<b>3.9</b> 2.202	<b>4.701</b> 1.614	<b>5.416</b> 1.240	0.043		

*Table 6.57 : Significant BIT grouped cube figure copying feature results*

### 6.8.5 Square grade-based patient groupings

As with the BIT defined groupings, the shape simplicity leads to few differences between the grade-based groupings. The feature shown in Table 6.58 gives the only significant difference obtained between the grade-based groups. The static feature of the measured angle at the lower right hand corner of the square is significantly different between the severe neglect subjects (Group 1) and mild neglect subjects (Group 3) and therefore sensitive to grades of neglect performance.

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Angle at Lower Right Corner (Degrees)</b>	<b>82.354</b> 5.068	<b>85.159</b> 3.386	<b>92.803</b> 5.568	<b>89.084</b> 5.558	<b>87.773</b> 5.258	<b>88.229</b> 4.746	1v3=0.008

Table 6.58 : Significant grade-based grouped square figure copying feature results

### 6.8.6 Cross grade-based patient groupings

The significant grade-based grouping results from the cross copying task are shown in Table 6.59. The results show no significant results within the neglect groups or the SC groups. However, they do show the BIT grouped differences described in Section 6.8.2.

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Pen Lifts</b>	<b>8</b> 4.358	<b>11</b> 8.660	<b>12.5</b> 10.014	<b>6.823</b> 4.965	<b>5.222</b> 3.587	<b>7.583</b> 2.906	3v5=0.012
<b>Pen Movement Time (sec)</b>	<b>12.783</b> 13.036	<b>12.236</b> 9.23	<b>19.137</b> 19.627	<b>8.555</b> 7.074	<b>6.207</b> 5.006	<b>5.918</b> 2.848	3v5=0.006, 3v6=0.014
<b>Total Execution Time (sec)</b>	<b>23.906</b> 16.481	<b>28.72</b> 10.203	<b>34.348</b> 20.363	<b>21.941</b> 10.105	<b>17.713</b> 9.856	<b>13.64</b> 4.987	3v5=0.007
<b>Number of Component</b>	<b>2.625</b> 2.199	<b>2.909</b> 1.921	<b>4.545</b> 0.820	<b>4.227</b> 1.601	<b>4.235</b> 1.577	<b>5</b> 0	1v6=0.011, 2v6=0.017
<b>Image Height Error (mm)</b>	<b>16.054</b> 13.090	<b>2.986</b> 2.796	<b>9.155</b> 10.153	<b>6.057</b> 5.084	<b>4.044</b> 4.005	<b>7.101</b> 4.731	1vs5=0.017

Table 6.59 : Significant grade-based grouped cross figure copying feature results

### 6.8.7 Star grade-based patient groupings

None of the results from the star copying task are significant within the neglect or SC based grade-based groupings (Table 6.60)

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Pen Lifts</b>	<b>19</b> 22.627	<b>19.5</b> 17.935	<b>14.333</b> 12.176	<b>8.166</b> 4.628	<b>4</b> 2.385	<b>7.444</b> 2.92	2v5=0.018
<b>Mean Pressure (range -128 to 128)</b>	<b>30.103</b> 20.77	<b>-7.247</b> 42.828	<b>-46.21</b> 31.453	<b>0.157</b> 56.970	<b>44.384</b> 44.455	<b>12.406</b> 52.435	3v5=0.006

*Table 6.60 : Significant grade-based grouped star figure copying feature results*

### 6.8.8 Cube grade-based patient groupings

As with the cross and the star, no results were extracted that showed significant differences within the neglect or SC based grade-based groupings (Table 6.61)

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
<b>Pen Lifts</b>	<b>5</b> -	<b>15</b> 1.414	<b>15.666</b> 10.670	<b>10.25</b> 2.667	<b>8.166</b> 3.639	<b>8.666</b> 3.316	3v5=0.027

\*ANOVA Significance not calculated on Group 1 due to single group membership.

*Table 6.61 : Significant grade-based grouped cube figure copying feature results*

The failure of the figure copying task to produce significant differences within both neglect and SC grade-based groups can be attributed to the fact that within the simpler tasks, performance is uniform (or when different, not attaining the required level of significance) whereas in the more complex tasks, subjects who would be likely to generate differences find the task too hard and therefore are excluded (in the case of the cube, only one member of the severe neglect group managed to accurately draw the shape, thus preventing the use of

ANOVA). Analysis of the mean results from the features, however, does indicate that performance deterioration correlates with the severity of neglect.

## 6.9 Drawing from Memory Results

Requiring the construction of two shapes used in the figure copying task (the square and the cube), the drawing from memory analysis uses identical inclusion criteria for these geometric figures as defined in the previous task. As with the figure copying results, no significant effects due to gender or age were found in any of the features.

### 6.9.1 Square BIT based patient groupings

Using the same inclusion criteria, the square drawing from memory task produces an almost identical number of exclusions as for the square figure copying task (Table 6.62).

No significant differences were found to separate the SC and neglect subject groupings which indicates that if a subject is able to draw the square, then the drawing does not present a challenge and the performance is uniform. Differences are evident, however, between the AMC and the two stroke groups; three dynamic features are able to differentiate between these subject groupings (Table 6.63)

<b>Group</b>	<b>% of Group Removed</b>
<b>Neglect</b>	40
<b>Stroke Control</b>	17.5
<b>Age Matched Control</b>	7.6

*Table 6.62 : Percentage of group excluded from analysis, drawing from memory, square*

<b>Feature</b>	<b>Neglect Mean (SD)</b>	<b>SC Mean (SD)</b>	<b>AMC Mean (SD)</b>	<b>AMC vs Neglect</b>	<b>AMC vs SC</b>	<b>SC vs Neglect</b>
<b>Pen Drawing Time (sec)</b>	<b>4.629</b> 2.293	<b>4.950</b> 3.243	<b>2.586</b> 1.257		0.035	
<b>Mean Pen Velocity (mm/sec)</b>	<b>4.705</b> 2.899	<b>4.098</b> 3.021	<b>8.502</b> 6.544	0.024	0.001	
<b>Mean Pen Acceleration (mm/sec/sec)</b>	<b>0.214</b> 0.174	<b>0.155</b> 0.130	<b>0.028</b> 0.391	0.046		

*Table 6.63 : Significant BIT grouped square drawing from memory feature results*

## 6.9.2 Cube BIT based patient groupings

The final shape drawing task requires test subjects to draw a cube from memory. Again, the inclusion criteria are identical to the cube figure copying task and a similar number of subjects are excluded (Table 6.64) from the analysis.

<b>Group</b>	<b>% of Group Removed</b>
<b>Neglect</b>	83.3
<b>Stroke Control</b>	57.9
<b>Age Matched Control</b>	23.1

*Table 6.64 : Percentage of group excluded from analysis, drawing from memory, cube*

No dynamic feature produces a significant separation between the BIT defined test groups, but the number of components in the drawn image separates the groups, with the neglect subjects drawing fewer than the two control groups (Table 6.65). As with the cube figure copying task, this analysis shows that the shape used is too difficult for the majority of the neglect subjects and those that can accurately draw the cube perform as the control groups.

Feature	Neglect Mean (SD)	SC Mean (SD)	AMC Mean (SD)	AMC vs Neglect	AMC vs SC	SC vs Neglect
Number of Components	3.466 2.012	4.771 1.464	5.75 0.6216	<0.001		0.001

Table 6.65 : Significant BIT grouped cube drawing from memory feature results

### 6.9.3 Square grade-based patient groupings

The results obtained by grouping the test population by grade-based criteria (as defined in Section 6.3) are detailed in Table 6.66. A significant difference exist between the number of components drawn by the three neglect groups (groups 1, 2 and 3). No other features provided significant results.

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
Pen Drawing Time (sec)	4.524 2.733	3.407 0.761	5.231 2.462	5.99 3.639	4.305 2.849	2.586 1.257	4v6=0.024
Movement to Drawing Time Ratio	0.530 0.207	0.407 0.3	0.910 0.876	0.616 0.448	0.449 0.334	1.024 0.538	5v6=0.024
Image Height (mm)	45.856 17.360	31.258 18.190	44.364 13.865	47.48 25.180	31.654 10.548	35.419 8.901	4v5=0.025
Number of Components	2.875 1.807	3.181 2.136	4.818 2.401	3.826 0.936	4.147 0.359	3.923 0.277	1v3=0.02, 2v3=0.047
Mean Pen Velocity (mm/sec)	5.768 3.549	3.21 1.143	4.778 3.040	4.398 3.688	3.913 2.577	8.502 6.544	5v6=0.011

Table 6.66 : Significant grade-based grouped square drawing from memory feature results

### 6.9.4 Cube grade-based patient groupings

As in the case of the square drawing from memory task, the only grade-based grouped feature result which produces significant differences within the neglect or SC groups is the number of components drawn (Table 6.67).

Feature	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)	Group 4 Mean (SD)	Group 5 Mean (SD)	AMC Mean (SD)	Significance
Number of Components	1.75 1.982	3.090 1.640	5.090 0.943	4.727 1.386	4.764 1.538	5.769 0.599	1v3<0.001, 1v4<0.001 1v5<0.001, 1v6<0.001 2v3=0.02, 2v4=0.035 2v5=0.014, 2v6<0.001

*Table 6.67 : Significant grade-based grouped cube drawing from memory feature results*

As discussed in Section 6.9.3, the performance on the cube task, if drawn successfully, is similar to the control groups and does not, therefore, produce any differences between the groupings (apart from the number of components drawn). The complexity of the shape drawn prohibits its use in assessing a graded neglect performance.

## 6.10 Shape Complexity

It is evident that the shape complexity and the ability to draw the correct number of components is the primary diagnostic feature in neglect drawing assessment. In many of the tasks, the shape is too difficult and therefore performance of those subjects able to produce an accurate response is uniform both statically and dynamically.

Applying a set of inclusion criteria to the shape drawing tasks, the percentage of exclusions indicates the difficulty in construction of the individual figures for the particular test groupings. To assess the effect shape complexity has on exclusion rate, a complexity metric was devised, based on the number of straight line edges within each model drawing. For example, the square is assigned a complexity index of 4, whereas the 'man' completion shape has an index of 8 (individual arm and leg - 3 components each - arm, body and head). Figures 6.3 to 6.5 separately show the graphs for the three subject groups.

The neglect group results shown in Figure 6.3 show the correlation between complexity and exclusion. The man and the cube shapes produce a higher exclusion rate with respect to their complexity index in comparison with the other shapes. The cube, as well as having a high complexity index, also involves perspective-based drawing which proves difficult for the neglect population. It can also be noted that the neglect patients produce less accurate results when drawing in to the left hand visual field (L) with the diamond and man completion tasks.

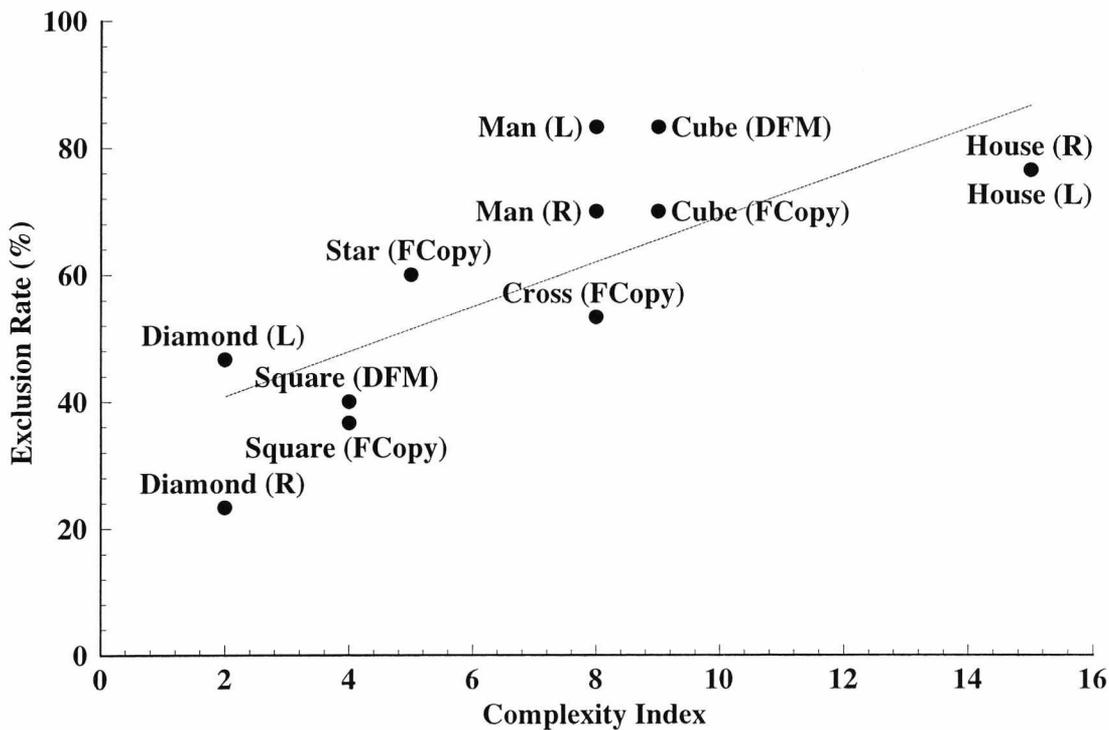


Figure 6.3 : Neglect group shape complexity vs exclusion %

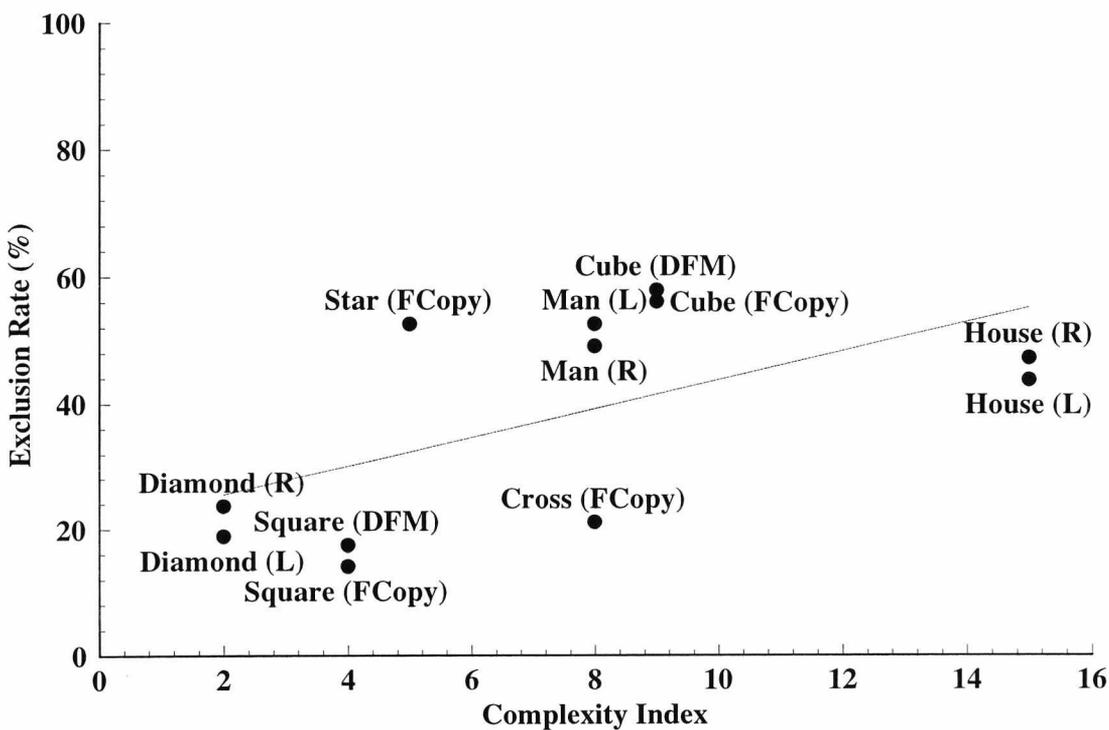


Figure 6.4 : Stroke control group shape complexity vs exclusion %

The stroke control groups exclusions are shown in Figure 6.4. It is obvious from the graph that the exclusion are less than for the neglect group, however, a correlation still exists between shape complexity and exclusions. While the man and the cube shape still cause a high exclusion rate, the star has a similar rate to the neglect group. This indicates that the static performance on the star is stable across all CVA subjects. As shown, dynamic features are able to detect significant differences between test populations.

The age matched control group (Figure 6.5) have a lower exclusion rate for all shapes, indicating the more accurate performance of this group. Again the star and the cube task prove the most difficult for the group. Interestingly, these shapes both contain acute angles indicating that the angular composition has an effect on accuracy in copying and drawing for all test groupings.

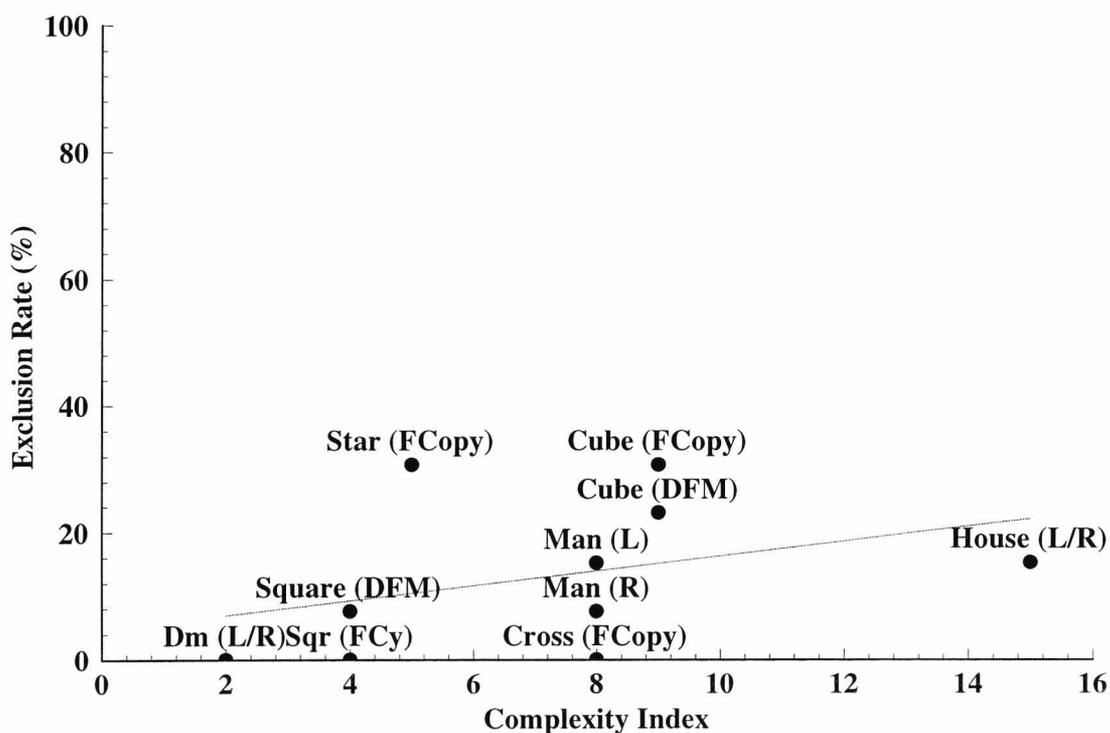


Figure 6.5 : Age matched control group shape complexity vs exclusion %

## 6.11 Kinematic profiling results

This section presents the results from a single neglect subject case study performing the kinematic profile overlays as described in Section 5.2.6. Intra-subject analysis is performed on this single patient comparing pen movement times, drawing times and velocity profiles within and between individual sides of the visual field.

The test subject in the trial is female, aged 82 and scored a conventional BIT score of 110. Tables 6.68 to 6.72 detail the timing based measurements from the test overlays. No left to right movements were made on overlay 1 - the right to left timings are presented in Table 6.68 as a comparison with the other 'zig-zag' drawing tasks. Timing data from these tests are presented in Tables 6.69 (Overlay 2) and 6.70 (Overlay 3). The pause and movement time is the time the pen is stationary or off the tablet *before* a movement in the specified direction was made.

<b>Direction</b>	<b>Mean Drawing Time (sec) (SD)</b>	<b>Mean Pause and Movement Time (sec) (SD)</b>
<b>Right to Left</b>	<b>3.046</b> 0.681	<b>1.257</b> 0.188

*Table 6.68 : Overlay 1 timing based measurements*

<b>Direction</b>	<b>Mean Drawing Time (sec) (SD)</b>	<b>Mean Pause and Movement Time (sec) (SD)</b>
<b>Right to Left</b>	<b>4.659</b> 1.336	<b>1.857</b> 0.875
<b>Left to Right</b>	<b>2.256</b> 0.536	<b>0.632</b> 0.681

*Table 6.69 : Overlay 2 timing based measurements*

<b>Direction</b>	<b>Mean Drawing Time (sec) (SD)</b>	<b>Mean Pause and Movement Time (sec) (SD)</b>
<b>Right to Left</b>	<b>2.540</b> 0.995	<b>1.187</b> 0.186
<b>Left to Right</b>	<b>1.915</b> 0.589	<b>0.380</b> 0.134

*Table 6.70 : Overlay 3 timing based measurements*

The data shows that the drawing time is greater when the pen is moved across the tablet surface into the inattentive visual field (right to left). A greater time is also taken prior to drawing (pause and movement time) when drawing in this direction. This indicates that the subject has problems both initially locating the target in the inattentive field and while executing the drawing. Overlays 4 and 5 were not analysed as the test subject only drew two out of the four lines required in drawing the square on both overlays.

Overlays 6 and 7 also required the drawing of a square, this time with all four corner points located in one half of the vertical visual field. The results shown in Table 6.71 indicate both greater drawing time per component and also a greater pause and movement time between drawing components for the square drawn on the left hand side of the visual field (Overlay 7). As all the drawing targets were located in the inattentive field, it appears that the position of the drawing destination affects the length of the planning phase and drawing time. This effect is also seen in Overlay 8, the clock face task (Table 6.72). Dividing the targets into left and right visual field positions (ignoring the vertically central 12 and 6 o'clock positions), drawing and movement times within the left visual field are longer than those on the right.

<b>Side</b>	<b>Mean Drawing Time (sec) (SD)</b>	<b>Mean Pause and Movement Time (sec) (SD)</b>
<b>Left Side Drawing (Overlay 7)</b>	<b>1.799</b> 0.585	<b>1.348</b> 0.739
<b>Right Side Drawing (Overlay 6)</b>	<b>1.534</b> 0.320	<b>0.289</b> 0.203

*Table 6.71 : Overlays 6 and 7 timing based measurements*

Side	Mean Drawing Time (sec) (SD)	Mean Pause and Movement Time (sec) (SD)
<b>Left Side Drawing</b>	<b>2.126</b> 0.921	<b>12.157</b> 10.992
<b>Right Side Drawing</b>	<b>1.355</b> 0.454	<b>2.631</b> 1.077

*Table 6.72 : Overlay 8 timing based measurements*

Four velocity based measures were used to assess individual drawing components. Tables 6.73 to 6.75 detail the results from the first three overlays. No performance trends are found on these tasks other than a longer time to reach the peak velocity when drawing into the inattentive field (right to left).

Direction	Mean Velocity (cm/sec) (SD)	Mean Peak Velocity (cm/sec) (SD)	Mean Time to Peak Velocity (Sec) (SD)	Mean Velocity Profile Skew
<b>Right to Left</b>	<b>0.503</b> 0.343	<b>1.730</b> 1.118	<b>18.878</b> 8.532	<b>0.653</b> 0.128

*Table 6.73 : Overlay 1 velocity based measurements*

Direction	Mean Velocity (cm/sec) (SD)	Mean Peak Velocity (cm/sec) (SD)	Mean Time to Peak Velocity (Sec) (SD)	Mean Velocity Profile Skew
<b>Right to Left</b>	<b>0.483</b> 0.064	<b>2.083</b> 0.137	<b>20.553</b> 5.429	<b>0.467</b> 0.203
<b>Left to Right</b>	<b>0.940</b> 0.123	<b>2.567</b> 0.619	<b>9.707</b> 2.933	<b>0.430</b> 0.164

*Table 6.74 : Overlay 2 velocity based measurements*

Direction	Mean Velocity (cm/sec) (SD)	Mean Peak Velocity (cm/sec) (SD)	Mean Time to Peak Velocity (Sec) (SD)	Mean Velocity Profile Skew
<b>Right to Left</b>	<b>0.903</b> 0.382	<b>3.407</b> 1.095	<b>8.013</b> 3.085	<b>0.327</b> 0.145
<b>Left to Right</b>	<b>0.917</b> 0.456	<b>2.007</b> 1.354	<b>6.097</b> 4.478	<b>0.413</b> 0.332

*Table 6.75 : Overlay 3 velocity based measurements*

The left and right visual field results from overlays 6, 7 and 8 reveal a performance pattern: drawings made on the left hand side of the overlay tend to be slower (also shown in the increased drawing times) and with a lower peak velocity. This peak velocity also occurs later in the profile indicating a greater acceleration phase when drawing in the inattentive field. Tables 6.76 and 6.77 show these results.

Side	Mean Velocity (cm/sec) (SD)	Mean Peak Velocity (cm/sec) (SD)	Mean Time to Peak Velocity (Sec) (SD)	Mean Velocity Profile Skew
<b>Left Side Drawing (Overlay 7)</b>	<b>0.498</b> 0.152	<b>1.160</b> 0.193	<b>7.335</b> 3.674	<b>0.440</b> 0.113
<b>Right Side Drawing (Overlay 6)</b>	<b>0.675</b> 0.209	<b>1.633</b> 0.389	<b>5.265</b> 3.955	<b>0.383</b> 0.148

*Table 6.76 : Overlays 6 and 7 velocity based measurements*

Side	Mean Velocity (cm/sec) (SD)	Mean Peak Velocity (cm/sec) (SD)	Mean Time to Peak Velocity (Sec) (SD)	Mean Velocity Profile Skew
<b>Left Side Drawing</b>	<b>0.420</b> 0.113	<b>1.247</b> 0.242	<b>11.797</b> 8.064	<b>0.493</b> 0.165
<b>Right Side Drawing</b>	<b>0.634</b> 0.139	<b>1.576</b> 0.367	<b>5.940</b> 5.057	<b>0.378</b> 0.236

*Table 6.77 : Overlay 8 velocity based measurements*

These results demonstrate the ability of velocity profiling to identify performance characteristics within a neglect subject. In particular, this occurs when drawing in or into the left hand side of the overlay, causing slower drawing times and an increased planning phase between components.

## 6.12 Feature Space Analysis and Classification

This section utilises the results detailed earlier in this chapter to assess the automated classification ability of a number of feature space clustering methodologies. The analysis represents a preliminary study on the basis of which potentially fruitful further areas for

investigation can be identified. Pattern classification presents a complex and diverse area for research and as such this study only attempts a very basic analysis using a set of standardised and established classification architectures. Optimisation of the classifiers is not a primary goal of the study, rather the identification that it is feasible to classify the data obtained from the visuo-spatial analysis system.

The aim of the feasibility study is to identify classification performance, firstly in identifying neglect performance from the age matched and stroke control groupings and secondly between levels of neglect based performance. Successful classification using the first ten principal components will indicate techniques for further investigation using a wider range of features, features from an individual sub-task of the computer battery and the possibility for the automatic classification of a finer resolution in neglect performance.

Identifying the principal features extracted from the computer based test responses, five separate categorisation techniques (Bayesian statistical, K-means clustering, Euclidean nearest neighbour clustering, the Kohonen self-organising map and Adaptive Resonance Theory) are assessed. The basic parameters within each of these methodologies are examined to assess performance variation.

### 6.12.1 Classification Aims and Methodology

The classification techniques used in this study have been described in Section 2.5. The performance of each classifier is assessed on two criteria:

- Using the BIT based groupings defined in Section 6.2, investigate the percentage of neglect responses that can be successfully identified from the control groupings.
- Within the neglect subject responses establish if the three performance based groupings (severe neglects, moderate neglects and mild neglects) as defined in Section 6.2 can be detected.

Selection of features presented to each classifiers was performed by a two stage process: All features producing significant differences (as detailed in Chapter 6) between neglects and stroke controls were identified, along with any additional features that indicated differences *within* the neglect population (e.g. a significant performance difference between severe

neglects and moderate neglects). A total of 48 features were identified. For the second selection stage, the 10 principal features were found by a Principal Component Analysis. These features comprised the vector presented to a classifier for each test subject. All features results within the vector were normalised within the range 0.0 to 1.0.

Where a particular classifier required training (such as the Kohonen network), the data set was divided into two, resulting in a training and test set each of 50 test subjects. Subject group membership proportions were maintained within these sets by arranging subjects by BIT score and selecting alternate subjects for inclusion into the training and test data.

### 6.12.2 Principal Features

Table 6.78 shows the principal components (PC's) identified by a Principal Component Analysis (PCA). PCA was used to identify which features cause the greatest performance variance across the test population. As the difference between subjects groupings are significant (as identified by an ANOVA) then a large variation indicates a feature which maximises separation. Analysis of the PC's shows that the majority of the features are static and are extracted from the cancellation based tasks. Three of the features are extracted from the drawing based tasks using the component based performance measures defined in Chapter 5. This demonstrates the performance detection ability of strictly defined assessment criteria. A single dynamic feature, the number of matches between the drawn sequence and a series of predefined sequences from the first OX cancellation test, was also identified. Normalisation of each feature was performed by dividing by the maximum obtainable value as detailed in Table 6.78.

### 6.12.3 Bayes' Statistical Classification

The results of presenting the entire data set to a Bayes' classifier are shown in Figure 6.6. Both the rate of detection of neglect and identification of the three graded performance sub-groups of neglect improve as the number of features included in the feature vector presented to the classifier are increased. The classifier was trained using 50 data sets of training data. This data was selected from alternate entries in the total data set of 100 test subjects, thus ensuring that the group membership proportionalities were maintained.

Principal Component	Normalisation Divisor	Feature
1	3	OX 1, number of cancellations, top left quadrant
2	3	OX 1, number of cancellations, bottom left quadrant
3	5	House figure completion, drawing to right hand side, number of components drawn
4	5	House figure completion, drawing to left hand side, number of components drawn
5	12	OX 1, number of cancellations, total
6	11	Albert's cancellation, number of cancellations, bottom left quadrant
7	12	OX 1, number of sequence matches
8	100	Bisection error %, 140mm line, top left quadrant
9	6	Cube drawn from memory, number of components drawn
10	36	Albert's cancellation, number of cancellations, total

*Table 6.78 : Principal features from computer based test battery*

As the results showed that the classification ability of the Bayes' improved as the number of features increased, the features ranked 11 to 13 by PCA were also included in the analysis. This identifies the number of features within the input set where performance deterioration occurs.

#### 6.12.4 Cluster Analysis

Two separate approaches have been used to assess the classification of the data set by cluster analysis. The results of the first approach, K-mean classification are shown in Figure 6.7 indicating a fluctuation in classification performance as the number of features included in the input vector are increased. The K-means algorithm was sequentially presented with each of the 100 feature vectors and asked to fit the data to 2 cluster centres for the detection of neglect subjects from the control groups. The 30 neglect features vectors were fitted to 3 cluster centres for graded neglect performance assessment.

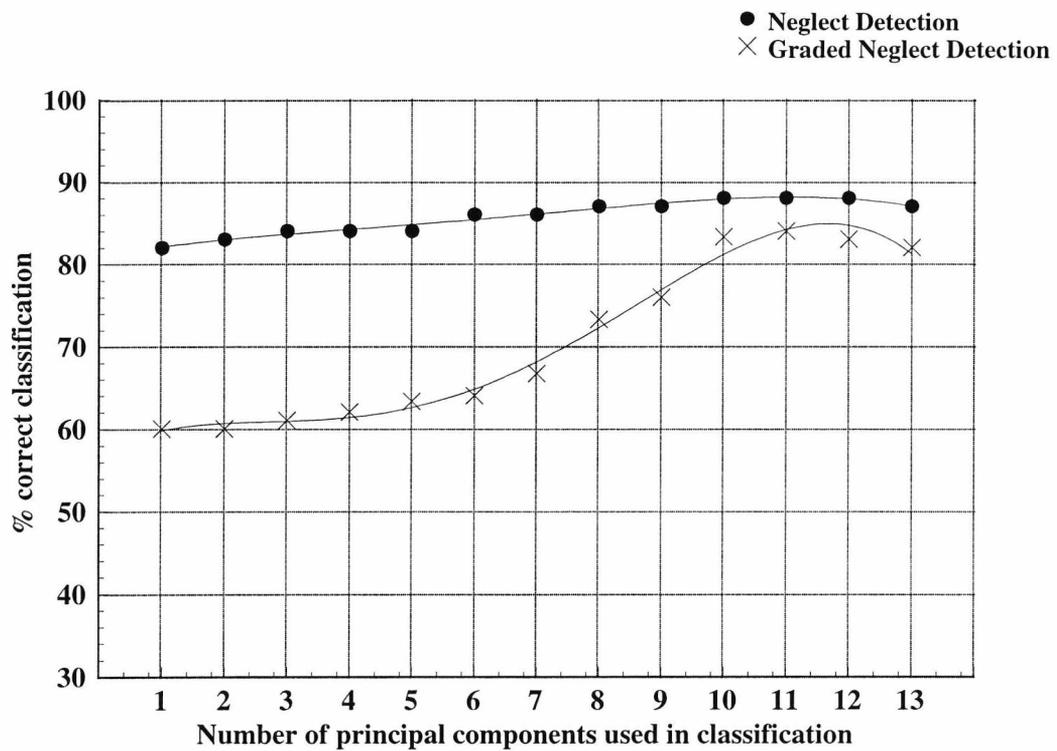


Figure 6.6 : Bayesian classifier performance

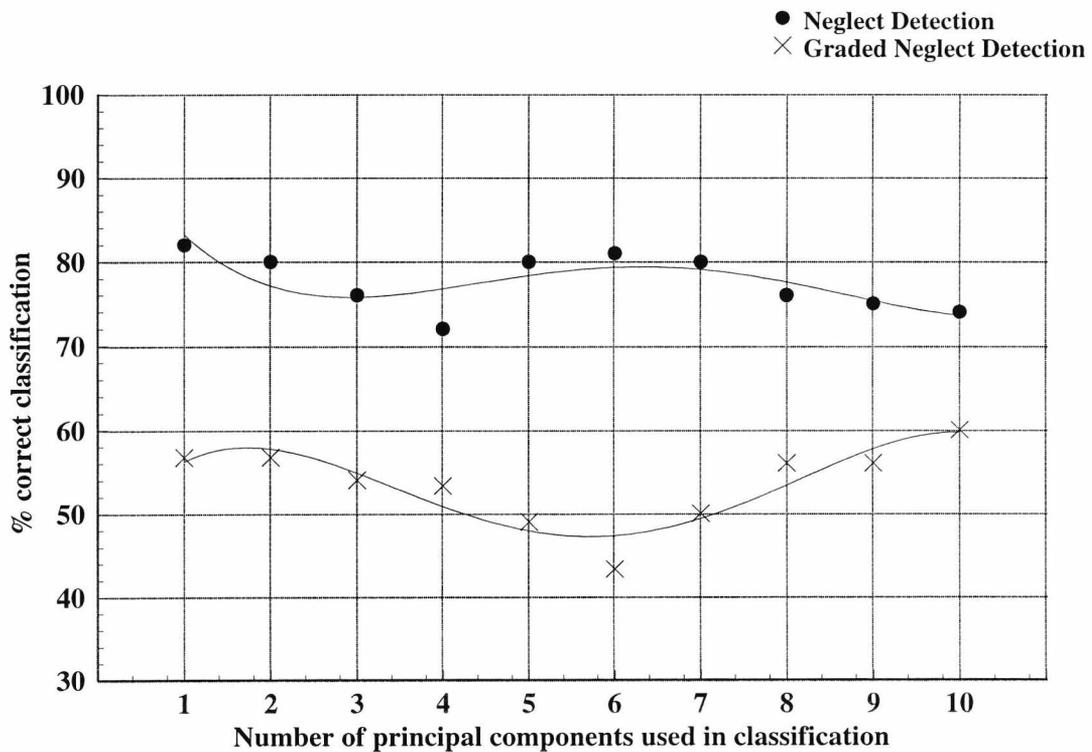


Figure 6.7 : K-means classifier performance

Although the detection of neglect from the control groups is on average 80 % correct, the detection of graded neglect performance is lower. The standard K-means clustering algorithm is sensitive to the initial cluster centres. As the data was presented ‘blind’ to the classifier, these centres were chosen at random for the trial. Implementation of algorithms such as the adaptive K-means with dynamic initialisation may solve this problem and improve the rates of classification and stability of the system [189].

The Euclidean distance cluster analysis results are shown in Figure 6.8. The classifier was initialised to form 2 clusters. Classification was performed by passing the feature vectors through the classifier twice. The first pass modified two cluster centres according to the specified group membership specified with each vector. The second pass classified each case according to the nearest cluster centre to the vector’s position in n-dimensional space.

Success in detection of neglect is approximately the same as the K-mean algorithm for the smaller size feature vectors, the performance deteriorates as more features are used for classification. Neglect performance grade classification is similar to the K-means algorithm across all sizes of feature vector included in the trial. This is possibly caused by the use of the same distance metric (Euclidean) within both of the clustering methods.

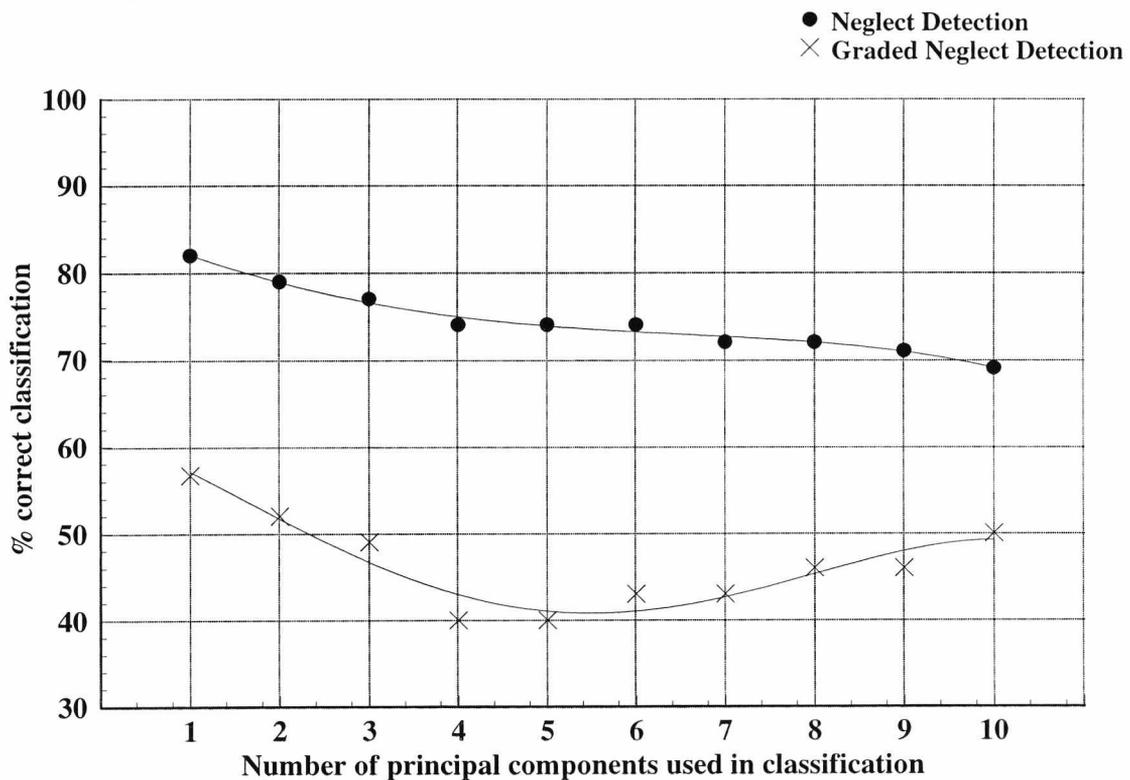


Figure 6.8 : Euclidean cluster classifier performance

While the Euclidean distance is probably the most widely used distance metric, the accuracy of classifier using this measure relies on a circular distribution of data within each cluster. Other measures, such as the Mahalanobis distance [190] result in elliptical probability contours which may provide a more accurate model for the description of the data belonging to each group. A range of distance metrics and the related modelling contours within a feature space are described in Fairhurst [126] and Everitt [191].

### 6.12.5 Kohonen Self Organising Map

The results of classification experiments using four sizes of network surface are detailed in Table 6.79. A set of 50 test subjects was used to train the network and a further 50 used to test the classification performance. The quantisation results are calculated by inputting the training data again to the trained network and finding the difference between the feature vector activation and the winning node activation. A smaller quantisation error indicated that the network is able to represent all input vectors without generalisation. The results shown in Table 6.79 indicate that a larger map results in a smaller quantisation error.

The network activation maps using the training set are shown in Figure 6.9 (AMC and Stroke Controls) and Figure 6.10 (Neglect). The areas shown in yellow and green through to blue and purple indicate the accumulative winning totals for each node. It can be noted that the control group nodes are located in the top left corner of the topological map, whereas the neglect groups are more scattered, with a greater bias towards the bottom right hand corner. Figure 6.11 shows the mean cluster activation centres for the two groups, confirming these map positions.

The neglect detection classification ability of the map was obtained by finding the nearest (Euclidean) activation centre to the winning node for each of the test vectors input to the network. These results are detailed in Table 6.79.

The deterioration in performance of the classifier as the size of the map increases indicates the optimal size for the map for maximum classification ability. Results from the classification of graded performance within the neglect groups are also shown in Table 6.79. These results show a similar deterioration of performance as the size of map increases. The cluster centres for the three neglect groups is shown in Figure 6.12. Again a Euclidean distance measure was used to calculate the nearest cluster centre and hence form a classification for each input

vector. The results from both classification criteria show similar results to the Euclidean and K-means clustering techniques. However, the SOM has the disadvantage of requiring a training phase within the classification process.

Size	Quantisation Error	Neglect Detection Rate	Neglect Severity Detection Rate
12 x 8	0.179	74 %	50 %
16 x 12	0.107	76 %	66 %
20 x 16	0.056	78 %	62 %
24 x 20	0.038	74 %	44 %

*Table 6.79 : Kohonen network size effects on classification rate*

### 6.12.6 Adaptive Resonance Theory

The classification results after presenting all 100 feature vectors to an Adaptive Resonance Theory 2 (ART) system are detailed in Table 6.80. The ART2 system, initiated with the vigilance parameter ( $\rho$ ) set at 0.9 (the comparison ratio threshold between input and comparison vectors within the F1 layer above which a new output pattern is formed), formed 13 pattern classifications from the input vectors. As the ART does not require an explicit training phase, classifications are formed by a single data pass through the architecture. The structure of the ART enables the network to learn as feature vectors are presented therefore requiring no formal training phase of operation. Analysis of these output patterns in relation to the test subject populations enabled the assignment to a particular patient group which are also detailed in Table 6.80. The results show a higher classification rate than the other methodologies and enable the detection of both BIT defined populations and grades of neglect performance.

Figure 6.13 shows the mapping between the BIT score for the stroke subjects (controls and neglects) and the output pattern identified by the ART indicating a high correlation between BIT score and output pattern assignment. Figure 6.14 shows the mean feature vector profile (a graph of the feature vector normalised value plotted against the principal component identifier) for each of output classification patterns. It is evident that the type of profile identified with the SC and AMC groups contain values closer to 1.0. The dip in the profile represents the bisection deviation measurement.

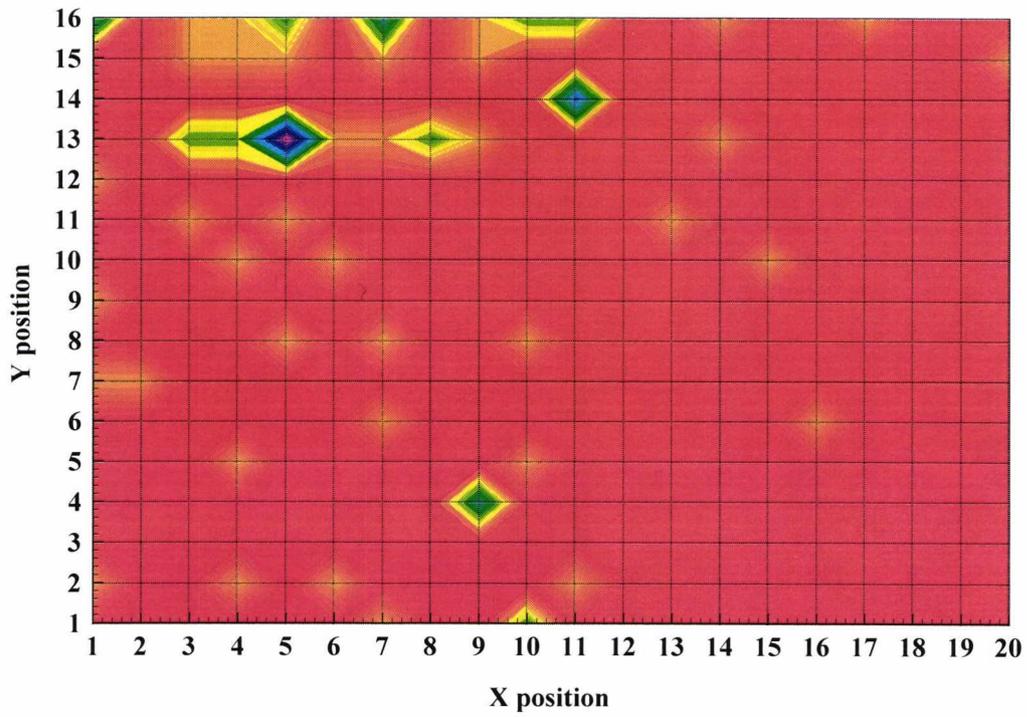


Figure 6.9 : 20 x 16 Kohonen activation map, control groups

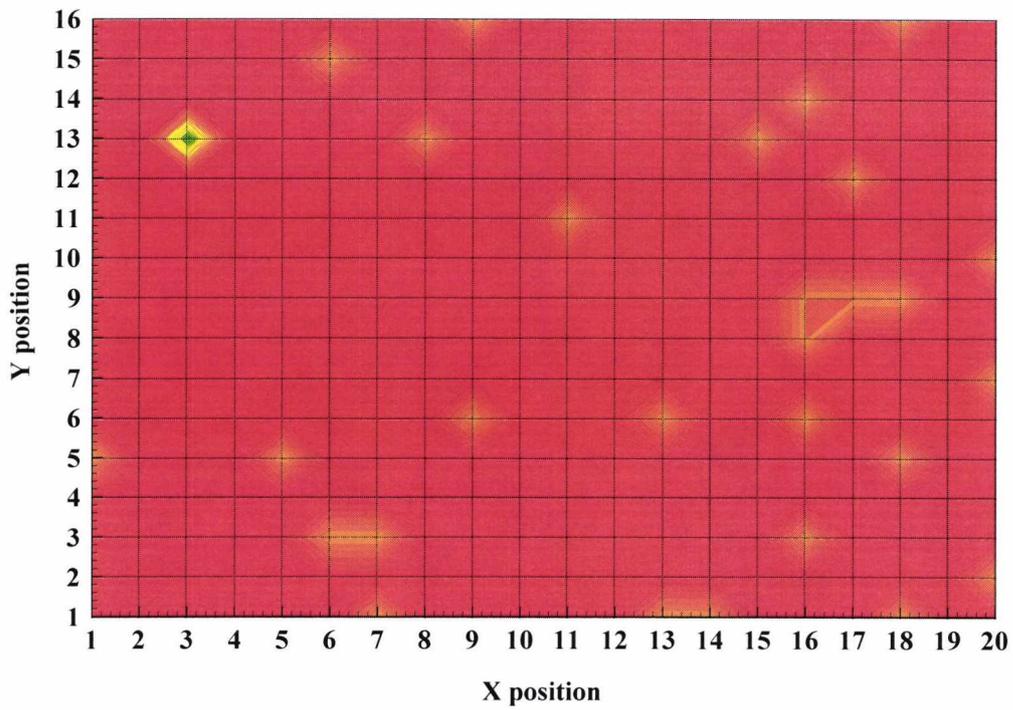


Figure 6.10 : 20 x 16 Kohonen activation map, neglect group

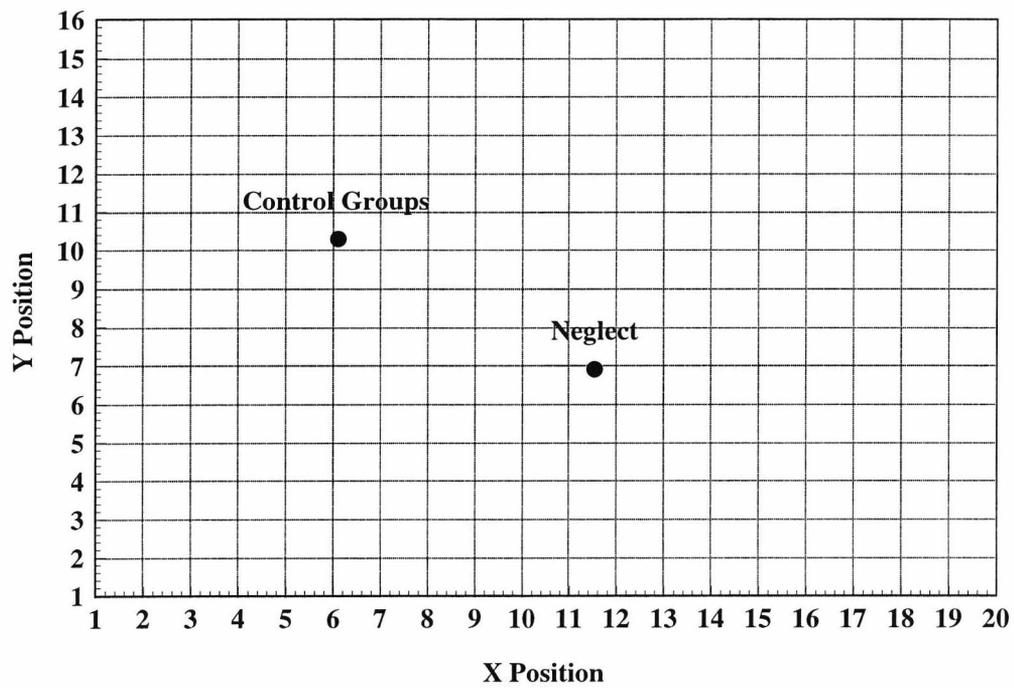


Figure 6.11 : Kohonen activation centres

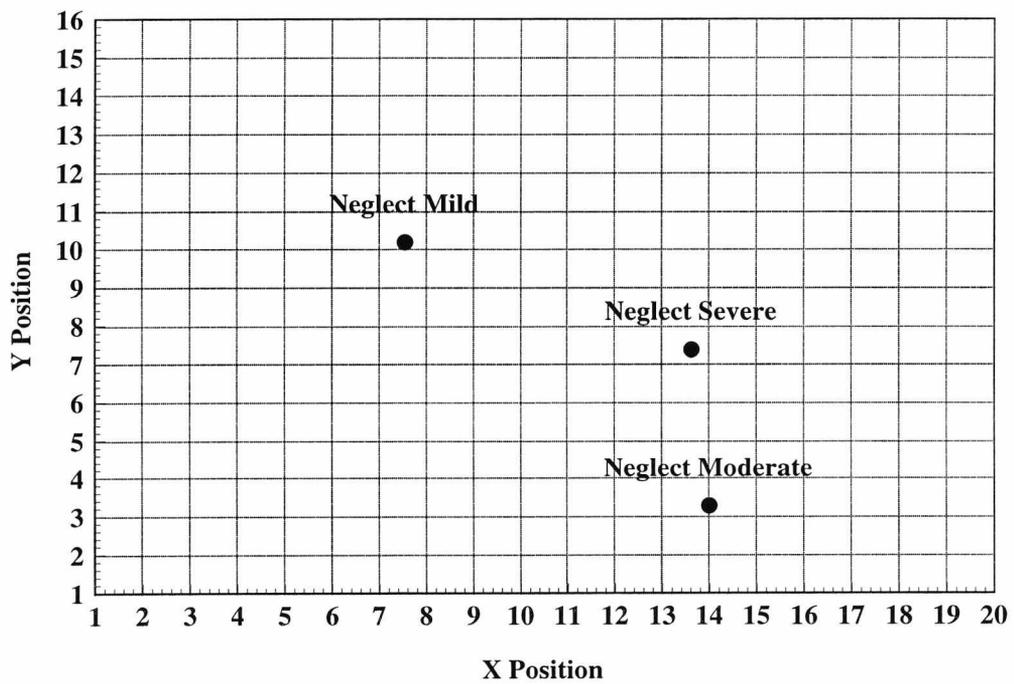


Figure 6.12 : Kohonen activation centres, neglect sub-groups

Group	Detection %	ART pattern assigned to group
Neglect	94	1,2,3,4,5,6,8,9
SC	85	7,10,11
AMC	93	12,13
Severe Neglect	91	1,2,3
Moderate Neglect	92	4,5,6
Mild Neglect	93	8,9

Table 6.80 : ART classification patterns

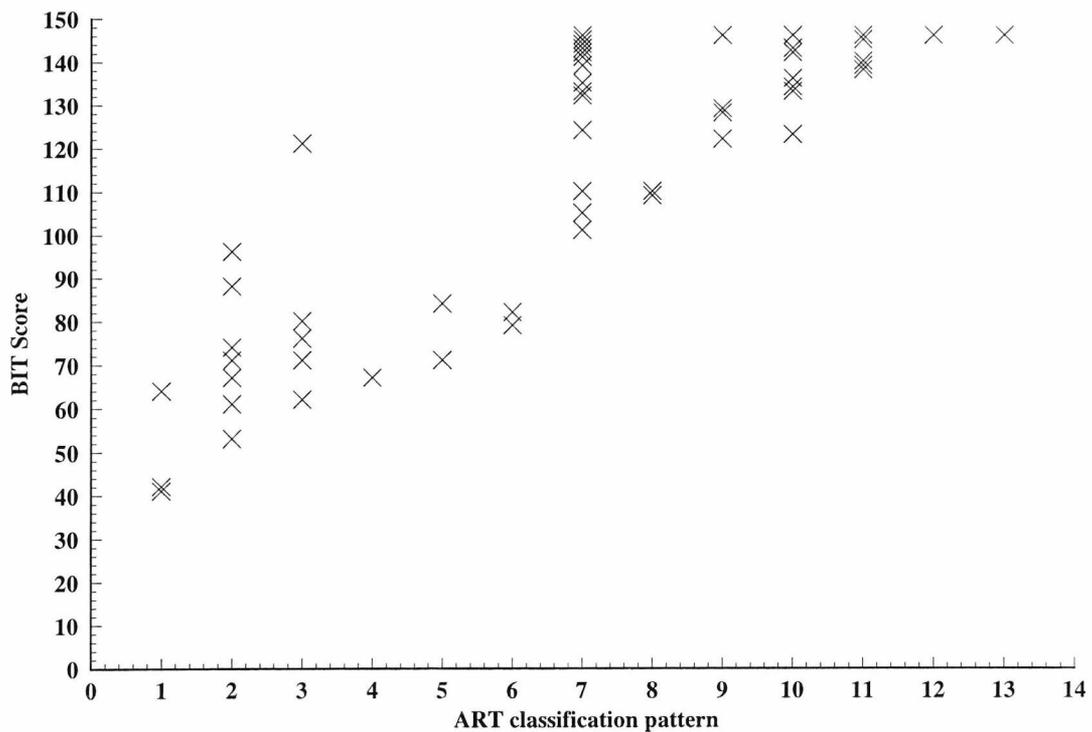


Figure 6.13 : BIT score vs ART classification pattern

### 6.12.7 Classifier Summary

This study has identified and evaluated classification methodologies for the detection of neglect and the classification of graded performance within a neglect population. Examination of the five classification techniques shows that using the first ten principal components, the ART system produces the highest correct classification for both the detection of neglect and

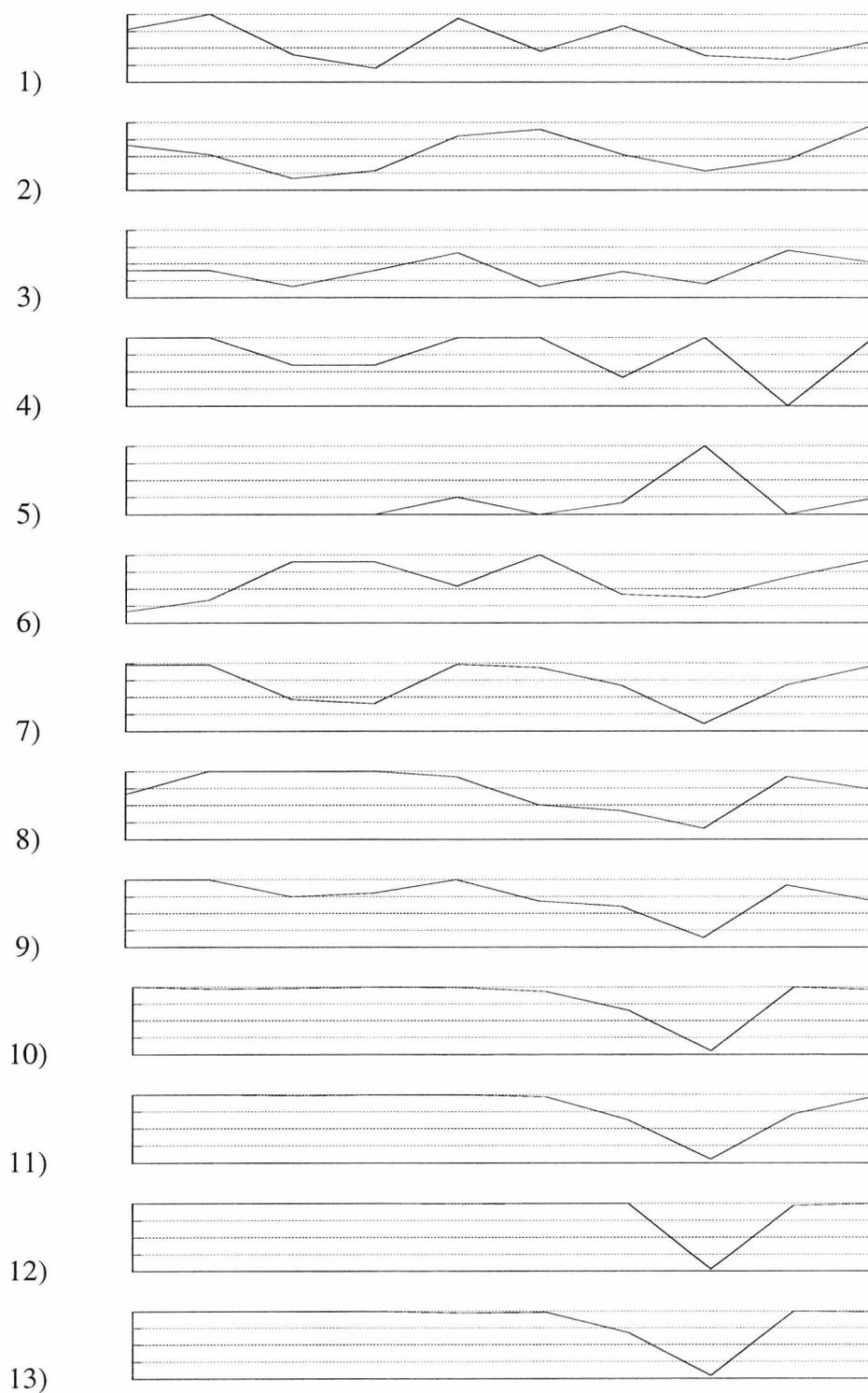


Figure 6.14 : Mean feature vector output patterns identified by the ART

graded performance within the neglect population. The ART is also able to accurately differentiate between age matched control and stroke control performance. Modification of both the network vigilance parameter (a reduction in this parameter causes the network to form a greater number of output patterns) and the number of features used in the input vector may increase the classification ability of the system.

The results from the Bayesian statistical classifier indicate a performance improvement proportional to the size of input vector used. The K-means, nearest neighbour Euclidean distance and the Kohonen SOM clustering techniques all return similar performance results, the K-means algorithm begin sensitive to the starting positions of the clusters. Increased accuracy in classification by this method would be obtained through the use of a variant of the existing algorithm, such as the adaptive K-means clustering with dynamic initialisation.

In this trial, classifiers have been used independently. Arrangement of multiple classifiers into a multi-expert configurations will improve classification as will the determination of classifier combinations and the method of result fusion [192] [193].

### 6.13 Summary

This chapter has presented the results from the tasks contained within the computer based test battery. Four assessment groupings have been used to analyse the results based on Conventional BIT battery result, graded-based performance within BIT categorisation, age and gender. None of the implemented tasks showed any significant performance effects when grouped by age or gender indicating the suitability of use across the range of geriatric patients.

The most sensitive tests to neglect detection are the bisection and cancellation based tasks. The longer lines within the bisection task were more sensitive, especially when placed to the left of the overlay. Bisection was also able to identify differences in neglect subjects grouped by grade-based criteria and hence the extent of neglect with a subject.

The three cancellation tasks all showed significant differences in the number of cancellations made when subjects were both grouped by BIT and grade-based performance. Again the most significant effects were to the left of the overlays. A number of dynamic features also produce significant differences with both the BIT and grade-based schemes. These features include the

division of intercancellation timings into drawing, movement and premovement phases, again with performance differences particularly prevalent on the left hand side of the overlay. Other sequence based results such as starting coordinate and quadrant and similarity to predefined sequences highlight constructional differences between the subject groups.

The drawing tasks did not produce the clearly defined significant differences in detecting neglect based performance across all task responses that were found in the bisection and cancellation tasks. This is consistent with the findings of Marshall and Halligan [86]. The ability of the drawing tasks to separate subject groups is dependent on the complexity of the drawing. With the simpler shapes, such as the square and the diamond completion task, a large proportion of all test groups are able to draw the required number of components and perform with similarity across a range of features. The most significant feature extracted from all drawing tasks was the number of components drawn, particularly for the more complex shapes. Standardisation of the component based assessment criteria has enabled the accurate and consistent component based assessment. Some dynamic features, such as mean pen velocity and movement timings, are also able to separate the groupings. This demonstrates the diagnostic ability of these novel features to classify a response which previously would have been considered identical to control subjects. The number of test subjects producing accurate drawings with the correct number of components is related to the complexity of the shape required to be drawn. When the shape contains many components (in the case of the cube and house) the subjects that produce accurate response all do so with similar static and dynamic performance. With these shapes, either the test subject is able to do the task, and will do so well, or not at all. Other drawing based findings have concluded that neglect subjects produce a more accurate response when copying from their inattentive (left hand) visual field than when drawing into it. However, they also tend to compress the drawing made on the right. Neglect patients tend to be significantly slower in movement and drawing phases when copying from the right hand side of the page (drawing in the left hand field). Test subjects tended to draw the square and cube with little variation in comparing pairs of responses drawn by individuals from the figure copying and drawing from memory tasks.

A number of dynamic features enable differences within both the BIT based and grade-based groupings to be obtained using the figure completion task, most prominently on the simpler diamond drawings. These graded differences were not found on the figure copying and drawing from memory tasks. In all drawings, the number of pen lifts and movement times between drawing components is greater for the neglect subjects, indicating the drawing

strategy used. Instead of treating the shape as a global image, the drawing is approached on a component basis.

The neglect test case results for the kinematic analysis have demonstrated the ability of the tasks to identify performance differences, most noticeably that neglect patients have difficulty locating targets on the left hand side of the overlay and are slower drawing in and moving towards this side of the visual field.

The results from the classification feasibility trial show that it is possible to use automated classification on features extracted from computer based test. No single classifier is able to produce a totally accurate performance which indicates the need for investigation into individual classifier optimisation and multiple classifier configurations.

## Chapter 7

### Summary and Conclusions

#### 7.1 Introduction

This chapter reviews the research undertaken and assesses the extent to which the aims and design requirements have been met. Highlighted are the major findings documented in the thesis along with suggestions for further research related to the work undertaken.

#### 7.2 Summary of Research Programme

The design objectives for the study were defined in Chapter 1:

- a) The implementation of accurate and consistent assessment of standard neuropsychological tests for visuo-spatial neglect using proven computer-based extraction methods.
- b) An investigation of the diagnostic ability of dynamic time-based features extracted from hand-drawn data.
- c) Classification of a test subject's performance by feature space analysis of the test responses.

These objectives were driven by the need for increased accuracy within the fields of stroke assessment and rehabilitation and secondly by the findings of recent investigations identifying the efficient use of computer techniques to increase the measurement resolution of pencil and paper based tests. The system development required the standardisation and algorithmic definition of features extracted from hand-drawn responses and also the implementation of a computer-based measurement system which would maintain the test infrastructure of proven neuropsychological assessments while enabling the capture of a series of constructional-based features.

---

A review of the relevant areas within the fields of neuropsychological assessment of neglect, computer-based feature extraction from hand-drawn data and feature space classification techniques were presented in Chapter 2. Important issues identified were the task specific spatial performance effects that are produced by neglect subjects and the need for objective assessment criteria, particularly within the drawing tasks. Most relevant to the research programme, a number of studies [2][48] have recently been undertaken to establish the kinematic movement profiling of neglect patients which reveal a correlation between spatial and constructional performance of neglect subjects when completing pencil and paper based tasks. The chapter introduced the concept of static and dynamic based features extractable from a computer-based capture system. These feature types were discussed along with a series of pattern recognition techniques which are able to provide performance classification from the assessment of results. A set of ten design objectives were defined for the research programme.

The subjectivity of current assessment techniques was demonstrated by the study reported in Chapter 3. The existing standard for neglect based assessment, the Rivermead Behavioral Inattention test (BIT) was used to examine 20 stroke patients (10 with neglect and 10 without the condition). The test responses were marked independently by eleven trained Occupational Therapists. While the inter-rater agreement on the overall assessment from the test was satisfactory, analysis of individual sub-test scores revealed disagreement caused by ambiguous marking criteria (particularly in the drawing tasks) and human error through difficulty of assessment (for example in the star cancellation task where many targets are required to be cancelled and assessed). In some cases, the variation in assessment caused an overall classification disagreement between raters from the same set of test responses. Areas in which a computer-based test system could aid these deficiencies were identified including the algorithmic application of marking schemes and a reduction in administration time from the two hours typically spent testing and marking a patient with the BIT battery. This time reduction has been realised in the 20 to 25 minutes required to collect and analyse test responses from the computer based implementation.

Chapter 4 described some of the practical considerations concerned with pen-based capture including sampling rates, data handling, pre-processing techniques (such as sample rate interpolation and low-pass filtering for noise reduction) and considerations for use within a hospital based environment. The data transfer protocol from the selected Wacom digitisation tablet was described as were the storage requirements of the data response files containing

pen position and status data. An example of feature extraction was provided, demonstrating the process of converting the raw coordinate data into a performance-based measurement.

Chapter 5 defined the computer-based test battery and detailed the range of features extracted from each sub-task. The sub-tasks are implementations or modifications from standardised neuropsychological tests which enables task verification against expected test performances. Extracted from the sub-tasks were a series of static-based features which assess both conventional performance characteristics (such as number of cancellations made) and new measurements, increasing the accuracy and sensitivity to spatial deficits (for example, performance on a quadrant basis and algorithmically assessing the presence of components within a drawing). New dynamic features, which assess the constructional and time-based properties of the test response enable a novel understanding of the task execution.

The significant results between three test subjects groupings (stroke subjects with neglect, stroke subjects without neglect and a set of age matched control subjects) defined in relation to their total scores from the Conventional battery of the BIT and collected in a clinically-based study of 100 test subjects, were presented in Chapter 6. A range of features were identified that detected significant performance differences between the stroke controls and neglect groups, thereby identifying performance characteristics sensitive to the presence of neglect. These features included the conventional assessments of the neglect identified in Chapter 2, thus confirming that a normal test response is not modified by the computer-based implementation. Also significant were a range of dynamic features, the majority of which mapped the spatial differences noted with static features; that effects were obtained when drawing or cancelling to the left hand side of the overlay (the neglected side of the visual field for right CVA subjects). It was noted that the bisection and cancellation tasks were the most sensitive to neglect. However, a range of dynamic features extracted from the drawing tasks were also significant, indicating several constructional aspects of neglect performance which were previously unobserved. In particular, the component-based drawing strategy used by neglect subjects, shown by increased movement and pre-movement times and the number of pen lifts made within the drawing process. Selection of drawing task proved critical to task sensitivity; many of the shapes used in the tasks were too complex, resulting in a large proportion of the all test groups producing unrecognisable responses. A component-based analysis of drawings provided standardisation and objective assessment criteria to this range of tasks.

An additional analysis identified features which were sensitive to graded performance within the neglect and stroke control population. Again the conventional static assessment features (including the standardised component-based drawing analysis) identified significant differences between these graded groupings.

The significant features identified from the analysis were used as the basis of a feasibility study into automated test response classification. Five classification methodologies were assessed using a set of features which were identified by Principal Component Analysis as producing the largest performance differential between the test groups. An analogue Adaptive Resonance Theory system produced the best classification rates for detection of both neglect and performance grades within the neglect group. The results from this study indicate that classifier performance is dependent on configuration variables such as classifier size, training set size, learning rates and number of features used to form the input vector.

In summary, the research has identified several new aspects of neglect performance, mainly from the automated dynamic and constructional measurements of the test response. These provide a deeper understanding of the mechanisms of neglect and serve as an aid to clinical diagnosis of patients. The study has identified the need for the clear definition of assessment rules, particularly for drawing-based tasks. The developed computer-based test battery has been used within a hospital environment (Figure 7.1) over a period of four years to collect over 150 sets of patient responses (see Section 7.3) and reduces the time and resources needed to collect data and produce an accurate and consistent assessment of a test subject.

The developed system has the ability to archive test responses and thus is able to monitor test performance and rehabilitation rates over time. The set of tools developed for data collection and feature extraction can be used for clinical investigation by replaying test response construction in real time and assessing performance differences between and within individual patients.

Most importantly, dynamic features which were previously unmeasured by conventional assessments of neglect have been shown to indicate significant performance differences between neglect and control populations. While principal differences are contained within the static spatial deficit measurements, the dynamic features show that constructional aspects of performance are related to the spatial errors made in the neglect visual field across a range of cancellation and drawing tasks. In some instances, where static performance appears normal, dynamic features identify a neglect based response, thus increasing the sensitivity of the task.

Feasibility investigations into classifier performance show an accurate detection of neglect using principal components. Classification techniques such as Adaptive Resonance Theory and Bayesian classification are able to identify levels of performance within the neglect group. Further work is required to identify the optimum classification architecture to increase the detection of graded performance levels across the range of test subjects.



*Figure 7.1 : Data capture within a hospital environment*

### 7.3 Suggestions for Further Research

While the research reported in this thesis indicates the ability of the devised system to collect, extract and assess data to determine the severity of neglect, several techniques, which may lead to increased accuracy and performance from the computer-based test environment, have not been fully explored. This section suggests some of the areas for continuing research in this field and indicates the type of clinically-based trial within which a validated computer test system could be utilised.

1. Investigation into user interface implementation will enable efficient use of the system within both the clinical and research based environments. Whereas for research based assessment, manual manipulation of raw feature data is acceptable, for clinical use, the display and computer interaction must be intuitive and require simple *automated* procedures. The information must be concise, unambiguous and tailored to what is required specifically by a therapist or other hospital staff [102]. The system output may not be required to list individual feature results, although these may be displayed if necessary to investigate a particular aspect of a patient's performance. Instead, an overall test performance metric may provide an adequate system result for everyday clinical use in much the same way as the total result of the conventional BIT test is currently used. While this masks individual performance characteristics, the therapist has the ability to interrogate any area of performance from any of the devised features.
2. The use of classification architectures has concentrated on using single implementations of the most widely used and understood methods [194]. Investigations into multi-expert combinations of classifiers arranged in different configurations (Figure 7.2) and optimised to produce a more accurate, reliable solution will increase the usability of the system in a clinical setting. Identifying the optimum operational parameters for each classifier (size, learning rates etc.) will also have an effect on the overall performance.

The consistency and objectivity with which features are extracted from data responses enables clinically-based trials using the system to obtain a greater understanding of the nature of visuo-spatial neglect in the fields of Neuropsychology and Occupational Therapy.

1. Alongside the Conventional BIT results obtained for all patients within the trial, the Behavioural battery of tasks was also administered (Section 3.2). An assessment of these results against the performance metric obtained from the computer-based features may identify correlations between everyday activity ability patterns and the severity of neglect exhibited by a patient.
2. A body of data comprising BIT results and computer based responses has been collected from stroke patients with a left side CVA as well as the right CVA group used in this study. While it is clinically recognised that neglect is less prevalent in Left CVA subjects, the sensitivity of some of the new dynamic features may improve the accuracy in detecting neglect.

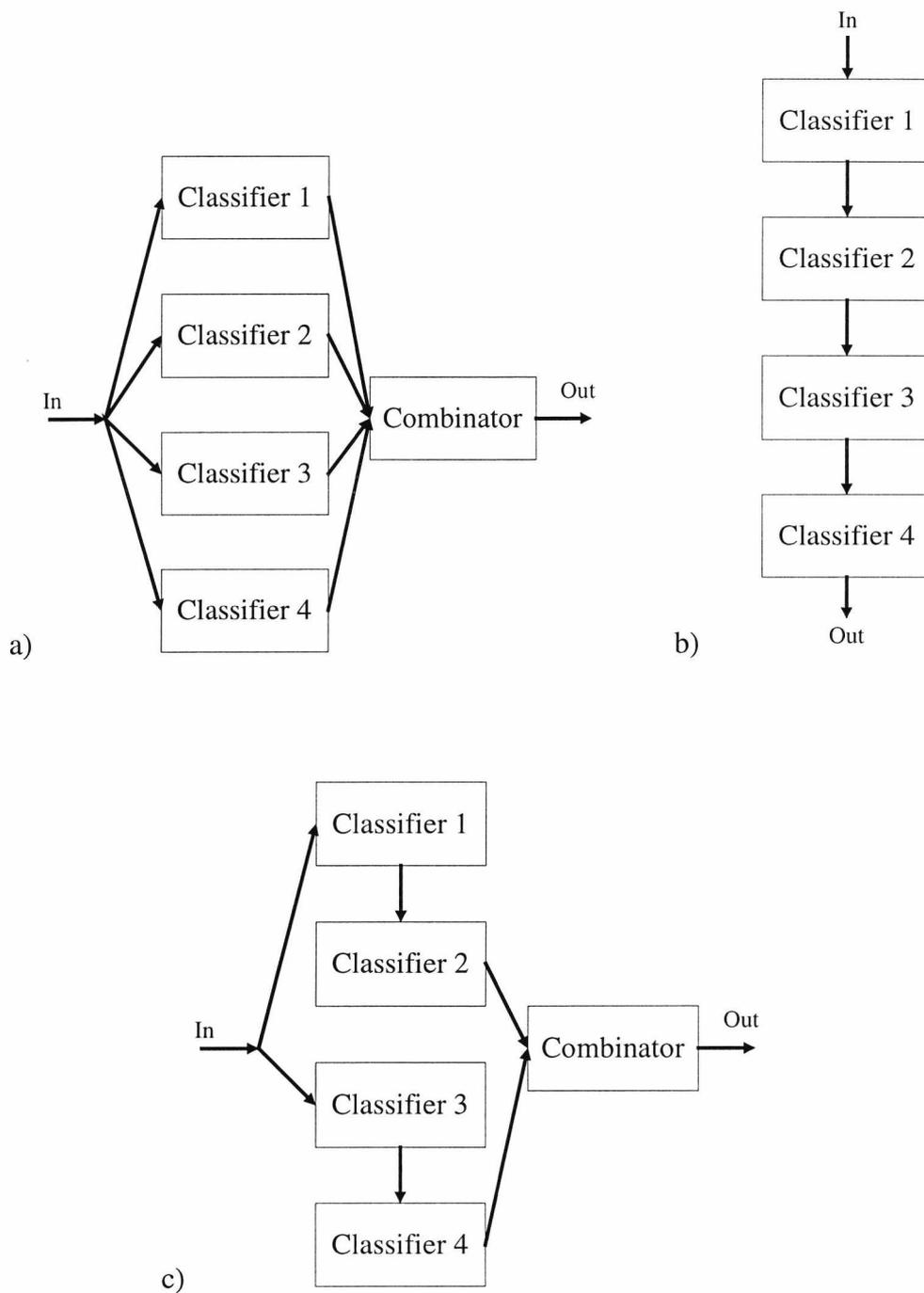


Figure 7.2 : Classifier architecture configurations: a) parallel b) pipeline c) hierarchical

3. The neurological location of the lesion causing the stroke is also recorded for each patient within the trial. Again, an assessment against the computer-based performance metric may indicate an associated between the stroke location and neglect severity.

4. The consistent application of an assessment scheme enables the monitoring of patient performance over time, indicating recovery rates and responses to rehabilitation schemes. Indeed, by applying a number of different schemes to separate patients, the effectiveness of a method applied to a particular category of patient (classified by neglect severity, lesion location etc.) can be established.
5. Using the defined features for analysis of hand-drawn responses which have effectively detected neglect performance characteristics, these can perhaps be applied to other areas of neuropsychological testing and handwritten data analysis.

The research documented in this thesis has shown that additional diagnostic indicators can be extracted using a computer-based assessment of neglect which can be used alongside conventional assessment methods to aid diagnosis of neglect in stroke patients. It is hoped that the increased sensitivity of detection can be utilised further within a clinical setting to provide a more accurate aid to recovery.

## References

- [1] *Research and Development - Towards an evidence-base for health services, public health and social care*, U.K. Department of Health, **1998**.
- [2] Mattingley, J.B., Phillips, J.G., Bradshaw, J.L., Impairments of movement execution in unilateral neglect: A kinematic analysis of directional bradykinesia. *Neuropsychologia*, **32**, pp. 1111-1134, **1994**.
- [3] Fairhurst, M.C., Higson, N., Interactive two dimensional image analysis in the assessment of visual-motor integration through figure copying. In: *Handwriting and Drawing Research: Basic and Applied Issues* (eds. Simner, M.L., Leedham, C.G., Thomassen, A.J.W.M.), IOS Press, Netherlands, **1996**.
- [4] Halligan, P.W., Marshall, J.C., Wade, D.T., Visuospatial neglect - underlying factors and test sensitivity. *Lancet*, Vol. II, pp. 908-911, **1989**.
- [5] Stone, S.P., Wilson, B., Wroot, A., Halligan, P.W., Lange, L.S., Marshall, J.S., Greenwood, R.J., The assessment of visuo-spatial neglect after acute stroke. *Journal of Neurology, Neurosurgery and Psychiatry*, **54**, pp. 345-350, **1991**.
- [6] Azouvi, P., Marchal, F., Samuel, C., et al., Functional consequences and awareness of unilateral neglect: study of an evaluation scale. *Neuropsychological Rehabilitation*, **6**, pp. 133-150, **1996**.
- [7] Plamondon, R., Yu, L.D., Stelmach, G.E., Clement, B., On the automatic extraction of biomechanical information from handwriting signals. *IEEE Transactions On Systems Man And Cybernetics*, **21**, pp. 90-101, **1991**.
- [8] Pao, Y.H., *Adaptive Pattern Recognition and Neural Networks*, Addison Wesley, **1989**.
- [9] Edmans, J.A., Lincoln, N.B., The frequency of perceptual deficits after stroke. *British Journal of Occupational Therapy*, **52**, pp. 266-270, **1989**.
- [10] Friedland, R.P., Weinstein, E.A., Hemi-inattention and hemisphere specialization: introduction and historical review. *Advances in Neurology*, **18**, pp. 1-31, **1977**.
- [11] Heilman, K.M., Watson, R.T., Valenstein, E., Neglect and related disorders. In: *Clinical Neuropsychology* (eds. Heilman, K.M., Valenstein, E), Oxford University Press, Oxford, UK. pp. 279-336, **1993**.
- [12] Vallar, G., The anatomical basis of spatial hemi-neglect in humans. In: *Unilateral Neglect: Clinical and Experimental Studies* (eds. Robertson, I.H., Marshall, J.C.), Lawrence Erlbaum Associates, Hove, UK. pp. 27-59, **1993**.
- [13] Halligan, P.W., Marshall, J.C., Left visuo-spatial neglect: A meaningless entity? *Cortex*, **28**, pp. 525-535, **1992**.
- [14] Halligan P.W., Robertson I.H., The assessment of unilateral neglect. In: *A Handbook of Neuropsychological Assessment* (eds. Crawford, J.R., Parker, D.M., McKinlay, W.W.), Lawrence Erlbaum Associates, Hove, UK. pp. 151-175, **1992**.

- [15] Behrmann, M., Moscovitch, M., Black, S., Mozer, M., Perceptual and conceptual mechanisms in neglect dyslexia. *Brain*, **113**, pp. 1163-1183, **1990**.
- [16] Vallar, G., Perani, D., The anatomy Of unilateral neglect after right-hemisphere stroke lesions - a clinical ct-scan correlation study in man. *Neuropsychologia*, **24** (5), pp. 609-622, **1986**.
- [17] Skilbeck C., Neuropsychological assessment in stroke. In: *A Handbook of Neuropsychological Assessment* (eds. Crawford, J.R., Parker, D.M., McKinlay, W.W), Lawrence Erlbaum Associates, Hove, UK. pp. 339-361, **1992**.
- [18] Critchley, M., *The Parietal Lobes*. Hafner Publishing Company, New York. **1953**.
- [19] Edmans, J.A., Lincoln, N.B., The relation between perceptual deficits after stroke and independence in activities of daily living. *British Journal of Occupational Therapy*, **53**, pp. 139-42. **1990**.
- [20] Ladavas, E., Menghini, G., Umilta, C., On the rehabilitation of hemispatial neglect. In: *Cognitive Neuropsychology and Cognitive Rehabilitation* (eds. Riddoch, M.J., Humphreys, G.W.), Lawrence Erlbaum Associates, Hove, UK. pp. 151-172, **1994**.
- [21] Robertson, I.H., The rehabilitation of attentional and hemi-attentional disorders. In: *Cognitive Neuropsychology and Cognitive Rehabilitation* (eds. Riddoch, M.J., Humphreys, G.W.), Lawrence Erlbaum Associates, Hove, UK. pp. 173-186, **1994**.
- [22] Lennon, S., Task specific effects in the rehabilitation of unilateral neglect. In: *Cognitive Neuropsychology and Cognitive Rehabilitation* (eds. Riddoch, M.J., Humphreys, G.W.), Lawrence Erlbaum Associates, Hove, UK. pp. 187-204, **1994**.
- [23] Colombo, A., De Renzi, E., Gentilini, M., The time course for visual hemi-attention. *Archives fur Psychiatrie und Nervenkrankheiten*, **29**, pp. 644-653, **1992**.
- [24] Gainotti, G., Les manifestations de negligence et d'inattention par l'hemispace. *Cortex*, **4**, pp. 64-91, **1968**.
- [25] Ericsson, K., Forrsell, L.G., Holmen, K., Viitanen, M., Copying and handwriting ability in the screening of cognitive dysfunction in old age. *Archives of Gerontology and Geriatrics*, **22**, pp. 103-121, **1996**.
- [26] Ishiai, S., Sugishita, M., Ichikawa, T., Gono, S., Clock-drawing test and unilateral spatial neglect. *Neurology*, **43**, pp. 106-110, **1993**.
- [27] Halligan, P.W., Marshall, J.C., Figural perception and parsing in visuo-spatial neglect. *Neuroreport*, **5**, pp. 537-539, **1994**.
- [28] Oxbury, J.M., Campbell, D.C., Oxbury, S.M., Unilateral spatial neglect and impairment of spatial analysis and visual perception. *Brain*, **97**, pp. 551-564, **1974**.
- [29] Halligan, P.W., Marshall, J.C., Completion in visuo-spatial neglect: A case study. *Cortex*, **30**, pp. 685-694, **1994**.
- [30] Stangalino, C., Semenza, C., Mondini, S., Generating visual mental images: Deficit after brain damage. *Neuropsychologia*, **33**, pp. 1473-1483, **1995**.

- [31] Doricchi, F., Incoccia, C., Galati, G., Influence of figure-ground contrast on the implicit and explicit processing of line drawings in patients with left unilateral neglect. *Cognitive Neuropsychology*, **14**, pp. 573-94, **1997**.
- [32] Andrews, K., Brocklehurst, J.C., Richards, B., Laycock, P.J., The prognostic value of picture drawings by stroke patients. *Rheumatology and Rehabilitation*, **19**, pp. 180-188, **1980**.
- [33] Kirk, A., Kertesz, A., Hemispheric contributions to drawing. *Neuropsychologia*, **27**, pp. 881-886, **1989**.
- [34] Swindell, C. S., Holland, A. L., Fromm, D., Greenhouse, J. B., Characteristics of recovery of drawing ability in left and right brain-damaged patients. *Brain and Cognition*, **7**, pp. 16-30, **1988**.
- [35] Thurmond, J.B., Hancock, J.B., Effects of figural complexity on the identification of different solid and outlined shapes. *Psychonomic Science*, pp. 315-317, **1969**.
- [36] Peru, A., Moro, V., Avesani, R., Aglioti, S., Overt and covert processing of left-side information in unilateral neglect investigated with chimeric drawings. *Journal of Clinical and Experimental Neuropsychology*, **18**, pp. 621-630, **1996**.
- [37] Cahn, D.A., Salmon, D.P., Monsch, A.U., Butters, N., Wiederholt, W.C., Coreyblom, J., Barrettconnor, E., Screening for dementia of the Alzheimer-type in the community - the utility of the clock drawing test. *Archives Of Clinical Neuropsychology*, **11**, pp. 529-539, **1996**.
- [38] DiPellegrino, G., Clock-drawing in a case of left visuo-spatial neglect: A deficit of disengagement. *Neuropsychologia*, **33**, pp. 353-8, **1995**.
- [39] Kirk, A., Kertesz, A., On drawing impairment in Alzheimer's disease. *Archives of Neurology*, **48**, pp. 73-77, **1991**.
- [40] Walton, J., Handwriting changes due to aging and Parkinson's syndrome. *Forensic Science International*, **88**, pp. 197-214, **1997**.
- [41] Oliveira, R.M., Gurd, J., Nixon, P., Marshall, J., Passingham, R., Micrographia in Parkinson's Disease: the Effect of Providing External Cues. *Journal of Neurology Neurosurgery and Psychiatry*, **63**, pp. 429-33, **1997**.
- [42] Fairhurst, M.C., Smith, S.L., Application of image-analysis to neurological screening through figure-copying tasks. *International Journal Of Bio-Medical Computing*, **28**, pp. 269-87, **1991**.
- [43] Smith, S.L., Fairhurst, M.C., Bucher, A., Image analysis tools for the monitoring of patients on long term haemodialysis programmes. *Proceedings of IPA'97, Dublin, Ireland*, pp. 571-575, **1997**.
- [44] Clar, C., *Evaluation of a computer based method for the analysis of a geometric figure copying test in learning disabled children*. M.Sc. Thesis, University of Kent, Canterbury, UK, **1994**.
- [45] Higson, N., *An innovative approach to visuo-perceptual testing using image analysis and pattern recognition techniques*. Ph.D. Thesis, University of Kent, Canterbury, UK, **1997**.

- [46] Beery, K., *The developmental test of visual-motor integration*, Modern Curriculum Press, Cleveland, **1967**.
- [47] Plamondon, R., On the origin of asymmetric bell-shaped velocity profiles in rapid-aimed movements. In *Tutorials in Motor Neuroscience* (eds. Requin, J., Stelmach, G.E.), Kluwer Academic Publishers, Netherlands, pp. 283-295, **1991**.
- [48] Konczak, J., Karnath, H., Kinematics of goal-directed arm movements in neglect: control of hand velocity, *Brain and Cognition*, **37**, pp. 387-403, **1998**.
- [49] Albert, M. L. A simple test of visual neglect. *Neurology*, **23**, pp. 658-664, **1973**.
- [50] Wilson, B., Cockburn, J., Halligan, P., Development of a behavioral test of visuospatial neglect. *Archives Of Physical Medicine And Rehabilitation*, **68**, pp. 98-102, **1987**.
- [51] Heilman, K. M.; Watson, R. T.; Valenstein, E., Localization of lesions in neglect and related disorders. In: *Localization and neuroimaging in neuropsychology* (ed. Kertesz, A.), pp. 495-523, Academic Press, San Diego, CA, **1994**.
- [52] Chatterjee, A.; Mennemeier, M.; Heilman, K. M., A stimulus response relationship in unilateral neglect: The power function. *Neuropsychologia*, **30**, pp. 1101-1108, **1992**.
- [53] Treisman, A., Gormican, S., Feature analysis in early vision - evidence from search asymmetries, *Psychological Review*, **95**, pp.15-48, **1988**.
- [54] Kaplan, R.F., Verfaellie, M., Meadows, M.E., Caplan, L.R., Pessin M.S., DeWitt. L.D., Changing attentional demands in left hemispatial neglect. *Archives of Neurology*, **48**, pp. 1263-1266, **1991**.
- [55] Henderson, L., Effects of letter-names on visual search. *Cognitive-Psychology*, pp. 90-96, **1973**.
- [56] Geldmacher, D. S., Stimulus characteristics determine processing approach on random array letter-cancellation tasks. *Brain and Cognition*, **36**, pp. 346-354, **1998**.
- [57] Geldmacher, D. S., Effects of stimulus number and target-to-distractor ratio on the performance of random array letter cancellation tasks. *Brain and Cognition*, **32**, pp. 405-415, **1996**.
- [58] Vanier, M., Gauthier, L., Lambert, J., Pepin, E. P., Evaluation of left visuospatial neglect: Norms and discrimination power of two tests. *Neuropsychology*, **4**, pp. 87-96, **1990**.
- [59] Halligan, P.W., Marshall, J.C., Is neglect (only) lateral? A quadrant analysis of line cancellation. *Journal Of Clinical And Experimental Neuropsychology*, **11**, pp. 793-798, **1989**.
- [60] Mark, V. W., Heilman, K. M., Diagonal spatial neglect. *Journal of Neurology, Neurosurgery and Psychiatry*, **65**, pp. 348-352, **1998**.
- [61] Chatterjee, A., Mennemeier, M., Heilman, K.M., Search patterns and neglect: A case study. *Neuropsychologia*, **30**, pp. 657-672, **1992**.
- [62] Geldmacher, D.S., Doty, L., Heilman, K.M., Spatial performance bias in normal elderly subjects on a letter cancellation task. *Neuropsychiatry, Neuropsychology and Behavioral Neurology*, **7**, pp. 275-280, **1994**.

- [63] Ruff, R.M., Evans, R.W., Light, H., Automatic detection vs. controlled search: a pencil and paper approach. *Perceptual and Motor Skills*, **62**, pp. 407-416, **1986**.
- [64] Heilman, K.M., Valenstein, E., Mechanisms underlying hemispatial neglect. *Annals Of Neurology*, **5**, pp. 166-170, **1979**.
- [65] Schenkenberg, T., Bradford, D.C., Ajax, E.T., Line bisection and unilateral visual neglect in patients with neurological impairments. *Neurology*, **30**, pp. 509-517, **1980**.
- [66] Nichelli, P.; Rinaldi, M., Cubelli, R., Selective spatial attention and length representation in normal subjects and in patients with unilateral spatial neglect. *Brain and Cognition*, **9**, pp. 57-70, **1989**.
- [67] Halligan, P.W., Marshall, J.C., Two techniques for the assessment of line bisection in visuo-spatial neglect: A single case study. *Journal of Neurology, Neurosurgery and Psychiatry*, **52**, pp. 1300-1302, **1989**.
- [68] Dellatolas, G., Vanluchene, J., Coutin, T., Visual and motor components in simple line bisection: An investigation in normal adults. *Cognitive Brain Research*, **4**, pp. 49-56, **1996**.
- [69] Manning, L., Halligan, P.W., Marshall, J.C., Individual variation in line bisection: A study of normal subjects with application to the interpretation of visual neglect. *Neuropsychologia*, **28**, pp. 647-655, **1990**.
- [70] Milner, A.D., Brechmann, M., Pagliarini, L., To halve and to halve not: An analysis of line bisection judgements in normal subjects. *Neuropsychologia*, **30**, pp. 515-526, **1992**.
- [71] Roig, M., Cicero, F., Hemisphericity style, sex, and performance on a line-bisection task: An exploratory study. *Perceptual and Motor Skills*, **78**, pp. 115-120, **1994**.
- [72] Chokron, S., DeAgostini, M., Line bisection in French and Israeli subjects: A developmental approach. *Brain and Cognition*, **32**, pp. 108-111, **1996**.
- [73] Scarisbrick, D.J., Tweedy, J.R., Kuslansky, G., Hand preference and performance effects on line bisection. *Neuropsychologia*, **25**, pp. 695-699, **1987**.
- [74] Butter, C.M., Mark, V.W., Heilman, K.M., An experimental analysis of factors underlying neglect in line bisection. *Journal of Neurology, Neurosurgery and Psychiatry*, **51**, pp. 1581-1583, **1988**.
- [75] Harvey, M., Milner, A.D., Roberts, R.C., Differential effects on line length on bisection judgements in hemispatial neglect. *Cortex*, **31**, pp. 711-722, **1995**.
- [76] Mozer, M.C., Halligan, P.W., Marshall, J.C., The end of the line for a brain-damaged model of unilateral neglect. *Journal of Cognitive Neuroscience*, **9**, pp. 171-190, **1997**.
- [77] Halligan, P.W., Marshall, J.C., Line bisection in visuo-spatial neglect: disproof of a conjecture. *Cortex*, **25**, pp. 517-521, **1989**.
- [78] Fujii, T., Fukatsu, R., Yamadori, A., Kimura, I., Effect of age on the line bisection test. *Journal of Clinical and Experimental Neuropsychology*, **17**, pp. 941-946, **1995**.

- [79] Hjaltason, H., Tegner, R., Line bisection with a head-mounted pointing device: An investigation of the role of right hand use and bilateral hemisphere activation in left neglect. *Neuropsychologia*, **35**, pp. 1175-1179, **1997**.
- [80] Wilson, B., Cockburn, J., Halligan, P.W., *The Behavioural Inattention Test*, Thames Valley Test Company, Fareham, Hampshire, **1987**.
- [81] Bisiach, E., Ricci, R., Lualdi, M., Colombo, M., Perceptual and response bias in unilateral neglect: Two modified versions of the Milner Landmark task. *Brain and Cognition*, **37**, pp. 369-386, **1998**.
- [82] Bucks, R.S., Willison, J.R., Development and validation of the location learning test (LLT): A test of visuo-spatial learning designed for use with older adults and in dementia. *Clinical Neuropsychologist*, **11**, pp. 273-286, **1997**.
- [83] Caplan, B., Assessment of unilateral neglect: A new reading test. *Journal of Clinical and Experimental Neuropsychology*, **9**, pp. 359-364, **1987**.
- [84] Marshall, R.S., Lazar, R.M., Binder, J.R., Desmond, D.W., Intrahemispheric localization of drawing dysfunction. *Neuropsychologia*, **32**, pp. 493-501, **1994**.
- [85] Warrington, E. K.; James, M., Kinsbourne, M., Drawing disability in relation to laterality of cerebral lesion. *Brain*, **89**, pp. 53-82, **1966**.
- [86] Marshall, J.C., Halligan, P.W., The yin and the yang of visuo-spatial neglect: A case study. *Neuropsychologia*, **32**, pp. 1037-1057, **1994**.
- [87] Casagrande, M., Violani, C., Curcio, G., Bertini, M., Assessing vigilance through a brief pencil and paper letter cancellation task (LCT): Effects of one night of sleep deprivation and of the time of day. *Ergonomics*, **40**, pp. 613-630, **1997**.
- [88] Elithorn, A., Mornington, S., Stavrou, A., Automated psychological testing: Some principles and practice. *International Journal of Man Machine Studies*, **17**, pp. 247-263, **1982**.
- [89] Allen, C.C., Ellinwood, E.H., Logue, P.E., Construct validity of a new computer-assisted cognitive neuromotor assessment battery in normal and inpatient psychiatric samples. *Journal of Clinical Psychology*, **49**, pp. 874-882, **1993**.
- [90] Kerkhoff, G., Marquardt, C., Standardised analysis of visual-spatial perception after brain damage. *Neuropsychological Rehabilitation*, **8**, pp. 171-189, **1998**.
- [91] Paul, S., Effects of computer assisted visual scanning training in the treatment of visual neglect: Three case studies. *Physical and Occupational Therapy in Geriatrics*, **14**, pp. 33-44, **1996**.
- [92] Lincoln, N. B., The assessment and treatment of disorders of visual perception. *Reviews in Clinical Gerontology*, **5**, pp. 77-82, **1995**.
- [93] Lynch, W.J., Ecological validity of cognitive rehabilitation software. *Journal of Head Trauma Rehabilitation*, **7**, pp. 36-45, **1992**.

- [94] Bergego, C., Azouvi, P., Deloche, G., Samuel, C., Louis-Dreyfus, A., Kaschel, R., Willmes, K., Rehabilitation of unilateral neglect: A controlled multiple-baseline-across-subjects trial using computerised training procedures. *Neuropsychological Rehabilitation*, **7**, pp. 279-293, **1997**.
- [95] McGuire, B.E., Computer-assisted cognitive rehabilitation. *Irish Journal of Psychology*, **11**, pp. 299-308, **1990**.
- [96] Carr, A.C., Automated testing of geriatric patients using a microcomputer-based system. *International Journal of Man Machine Studies*, **17**, pp. 297-300, **1982**.
- [97] Tseng, H., Macleod, H.A., Wright, P., Mood assessment and cognitive performance: are they influenced by computer anxiety? *Proceedings of Computers in Psychology '94*, York, UK. **1994**.
- [98] Ogden, M., Kellett, J.M., Merryfield, P., Millard, P.H., Practical aspects of automated testing of the elderly. *Bulletin of the British Psychological Society*, **37**, pp. 148-149, **1984**.
- [99] Roberts, G.I., Samuels, M.T., Handwriting remediation: a comparison of computer-based and traditional approaches. *Journal of Educational Research*, **87**, pp. 118-125, **1993**.
- [100] Collinson, P.O., Jones, R.G., Howes, M., Nicholls, J., Sheehy, N., Boran, G.R., Cramp, D.G., Of mice and men: data capture in the clinical environment. *International Journal of Clinical Monitoring and Computing*, **6**, pp. 217-222, **1989**.
- [101] Overton, R.C., Taylor, L.R., Zickar, M.J., Harms, H.J., The pen-based computer as an alternative platform for test administration. *Personnel Psychology*, **49**, pp. 455-464, **1996**.
- [102] Acuff, R.D., Fagan, L.M., Rindfleisch, T.C., Carlson, R.W., Tuttle, M.S., Sherertz, D.D., Integration of pen-based computer technology in clinical settings. *Journal Of The American Medical Informatics Association*, **55**, pp. 1042, **1994**.
- [103] Lee, L.L., Berger, T., Aviczer, E., Reliable online human signature verification systems. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, **18**, pp. 643-647, **1996**.
- [104] Wilfong, G., Sinden, F., Ruedisueli, L, On-line recognition of handwritten symbols, *IEEE Transactions On Pattern Analysis And Machine Intelligence*, **18**, pp. 935-940, **1996**.
- [105] Found, B., Rogers, D., Schmittat, R., A computer-program designed to compare the spatial elements of handwriting. *Forensic Science International*, **68**, pp. 195-203, **1994**.
- [106] Schomaker, L.B., Van-Galen, G.P., Computer models of handwriting. In : *Computational Psycholinguistics* (eds. Dijkstra, T., Smedt, K.), Taylor and Francis, London, UK, pp. 386-420, **1996**.
- [107] Doermann, D.S., Rosenfeld, A., Recovery of temporal information from static images of handwriting. *International Journal Of Computer Vision*, pp 143-164, **1995**.
- [108] Smith, S.L., *A new approach to neurological monitoring using image analysis*. Ph.D. Thesis, University of Kent, Canterbury, UK, **1990**.
- [109] Marquardt, C., Mai, N., A computational-procedure for movement analysis in handwriting. *Journal Of Neuroscience Methods*, **52**, pp. 39-45, **1994**.

- [110] MacKenzie, C.L., Marteniuk, R.G., Dugas, C., Liske, D., Three-dimensional movement trajectories in Fitts' task: implications for control. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **39**, pp. 629-647, **1987**.
- [111] Teulings, H.L., Stelmach, G.E., Control of stroke size, peak acceleration and stroke duration in Parkinsonian handwriting. *Human Movement Science*, **10**, pp. 315-334, **1991**.
- [112] Desbiez, D., Vinter, A., Meulenbroek, R.J., Biomechanical and perceptual determinants of drawing angles. *Acta-Psychologica*, **94**, pp. 253-271, **1996**.
- [113] Slavin, M.J., Phillips, J.G., Bradshaw, J.L., Visual cues and the handwriting of older adults: A kinematic analysis. *Psychology and Aging*, **11**, pp. 521-526, **1996**.
- [114] Morgan, M., Bradshaw, J.L., Philips, J.G., Mattingley, J.B., Iansek, R., Bradshaw, J.A., Effects of hand and age upon abductive and adductive movements - a kinematic analysis. *Brain and Cognition*, **25**, pp. 194-206, **1994**.
- [115] Bellgrove, M.A., Phillips, J.G., Bradshaw, J. L., Gallucci, R.M., Response (re) programming in aging: A kinematic analysis. *Journals of Gerontology: Series A: Biological Sciences and Medical Sciences*, **53A**, pp. 222-227, **1998**.
- [116] Geogiou, N., Bradshaw, J., Phillips, J., Cunnington, R., Rogers, M., Functional asymmetries in the movement kinematics of patients with Tourette's Syndrome. *Journal of Neurology, Neurosurgery and Psychiatry*, **63**, pp. 188-195, **1997**.
- [117] Carnahan, H., Aguilar, O., Malla, A., Norman, R., An investigation into movement planning and execution deficits in individuals with schizophrenia. *Schizophrenia Research*, **23**, pp. 213-221, **1997**.
- [118] Bellgrove, M.A., Phillips, J., Bradshaw, J., Hall, K., Presnell, I., Hecht H., Response programming in dementia of the Alzheimer type: a kinematic analysis. *Neuropsychologia*, **35**, pp. 229-240, **1997**.
- [119] Phillips, J.G., Bradshaw, J., Chui, E., Teasdale, N., Iansek, R., Bradshaw, J., Bradykinesia and movement precision in Huntington's Disease. *Neuropsychologia*, **34**, pp. 1241-1245, **1996**.
- [120] Elble, R.J., Brilliant, M., Leffler, K., Higgins, C., Quantification of essential tremor in writing and drawing. *Movement Disorders*, **11**, pp. 70-78, **1996**.
- [121] Heilman, K.M., Bowers, D., Coslett, H.B., Whelan, H., Watson, R.T., Directional hypokinesia: Prolonged reaction times for leftward movements in patients with right hemisphere lesions and neglect, *Neurology*, **35**, pp. 855-859, **1985**.
- [122] Jolliffe, I.T., *Principal Component Analysis*, Springer, **1986**.
- [123] Dunteman, G.H., *Principal Component Analysis*, Sage, London, **1989**.
- [124] Zhang, B.L., Fu, M.Y., Yan, H., Handwritten digit recognition by a mixture of local principal component analysis, *Neural Processing Letters*, **8**, pp. 241-251, **1998**.
- [125] Jacobs, E.L., O'Brien, G., Investigation of computational vision and principal component analysis with application to target classification, *Optical Engineering*, **37**, pp. 2022-2028, **1998**.

- 
- [126] Fairhurst, M.C., *Computer vision for robotic systems*, Prentice Hall, **1988**.
- [127] Scutt, L.E., Chow, E., Honer, W.G., Hogan, J., Jones, C., Weksberg, R., Bassett, A.S., Identifying subgroups in schizophrenia using minor dysmorphic features and cluster analysis, *Schizophrenia Research* (Abstract), **36**, pp. 95-96, **1999**.
- [128] Schweitzer, L., Renehan, W.E., The use of cluster analysis for cell typing. *Brain Research Protocols*, **1**, pp. 100-108, **1997**.
- [129] Winters, J.C., Groenier, K.H., Sobel, J.S., Arendzen, H.H., Classification of shoulder complaints in general practice by means of cluster analysis. *Archives of Physical Medicine and Rehabilitation*, **78**, pp. 1369-1374, **1997**.
- [130] MacQueen, J., Some methods for classification and analysis of multivariate observations. In: *Proceedings of the 5<sup>th</sup> Berkley Symposium*, pp. 281-297, **1967**.
- [131] Duda, R.O., Hart, P.E., *Pattern Classification and Scene Analysis*, Wiley-Interscience, **1973**.
- [132] Devijver, P.A., Kittler, J., *Pattern Recognition: A Statistical Approach*, Prentice-Hall, New Jersey, **1982**.
- [133] Kohonen, T., *Self-Organisation and Associative Memory*, Springer Series in Information Sciences Vol. 8, Springer-Verlag, **1984**.
- [134] Cleva, C., Cachet, C., Clustering of infrared spectra with Kohonen networks, *Analysis*, **27**, pp. 81-90, **1999**.
- [135] Kirew, D.B., Chretien, J.R., Bernard, P., Ros, F., Application of Kohonen neural networks in classification of biologically active compounds. *SAR and QSAR in Environmental Research*, **8**, pp. 93-107, **1998**.
- [136] Lozano, S., Guerrero, F., Onieva, L., Larraneta, J., Kohonen maps for solving a class of location-allocation problems. *European Journal of Operational Research*, **108**, pp. 106-117, **1998**.
- [137] Rosenblatt, R., *Principals of Neurodynamics*, Spartan Books, Washington D.C., **1962**.
- [138] Matthews, C.P., Warwick, K., Practical application of self-organising feature maps to process modeling. In: *Proc. EANN'95*, Helsinki, Finland, pp. 452-499, **1995**.
- [139] Weldemark, J., An automated procedure for cluster analysis of multivariate satellite data. *International Journal of Neural Systems*, **8**, pp.3-15, **1997**.
- [140] Grossberg, S., Adaptive pattern classification and universal recoding II: feedback, expectation, olfaction and illusions. *Biological Cybernetics*, **23**, pp. 187-202, **1976**.
- [141] Carpenter, G.A, Grossberg, S., ART2: self-organization of stable category recognition codes for analog input patterns. *Applied Optics*, **26**, pp. 4919-4930, **1987**.
- [142] Carpenter, G.A, Grossberg, S., Rosen, D.B., ART 2-A: an adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks*, **4**, pp. 493-504, **1991**.

- [143] Carpenter, G.A, Grossberg, S., Rosen, D.B., Fuzzy ART : fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks*, **4**, pp. 759-771, **1991**.
- [144] Carpenter, G.A, Grossberg, S., ARTMAP : supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, **4**, pp. 543-564, **1991**.
- [145] Beale, R., Jackson, T., *Neural Computing: an introduction*, Institute of Physics Publishing, Bristol, **1994**.
- [146] Freeman, J.A., Skapurs, D.M., *Neural Networks, Algorithms, Applications and Programming Techniques*, Addison Wesley, **1992**.
- [147] Cross, S.S., Harrison, R.F., Drezet, P., Making the distinction between chronic idiopathic inflammatory bowel disease and normality by histopathological examination: A comparison of human performance, logistic regression and adaptive resonance theory mapping neural network. *Journal of Pathology* (Abstract), **186**, pp. A2, **1998**.
- [148] Wang, X.Z., Chen, B.H., Clustering of infrared spectra of lubricating base oils using adaptive resonance theory. *Journal of Chemical Information and Computer Sciences*, **38**, pp 457-462, **1998**.
- [149] Webster, J. S., Roades, L. A., Morrill, B., Assessment of patients with unilateral neglect. In *The neuropsychology handbook Vol. 1: Foundations and assessment (2nd ed.)* (eds. Horton, A.M. Jr., Wedding, D.), Springer Publishing Co, New York, NY, USA, **1997**.
- [150] Hartman-Maeir, A., Katz, N., Validity of the Behavioral Inattention Test (BIT): Relationships with functional tasks. *American Journal of Occupational Therapy*, **49**, pp. 507-516. **1995**.
- [151] Whitting S., Lincoln N. B., An ADL assessment for stroke patients. *British Journal of Occupational Therapy*, **43**, pp. 44-46. **1980**.
- [152] Shiel, A., Wilson, B.A., The relationship between unilateral neglect and dependence in activities of daily living. In *Proc. British Occupational Therapy Conference*, Dublin, **1992**.
- [153] Cermak, S.A., Hausser, J., The behavioural inattention test for unilateral visual neglect: A critical review. *Physical and Occupational Therapy in Geriatrics*, **7**, pp. 43-53, **1989**.
- [154] Bakeman, R., Gottman, J.M., *Observing Interaction*, Cambridge University Press, Cambridge, UK, **1986**.
- [155] Williamson, J.M., Manatunga, A.K., Assessing interrater agreement from dependant data, *Biometrics*, **53**, pp. 707-714, **1997**.
- [156] Cohen, J., A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, pp. 37-46, **1960**.
- [157] Carletta, J., Assessing agreement on classification tasks: the kappa statistic, *Computational Linguistics*, **30**, pp. 408-414, **1995**.
- [158] Fleiss, J.L., Measuring agreement between judges on the presence on absence of a trait, *Biometrics*, **31**, pp. 651-9, **1975**.

- 
- [159] Landis, J.R., Koch, G.C., The measurement of observer agreement for categorical data. *Biometrics*, 33, pp. 1089-91, **1977**.
- [160] Krippendorff, K., *Content analysis: an introduction to its methodology*. Sage Publication. **1980**.
- [161] Cronbach, L., Coefficient alpha and internal structure of tests. *Psychometrika*, 16, pp. 297-334, **1951**.
- [162] Bland, J.M., Altman, D.G., Statistics notes: Cronbach's alpha, *British Medical Journal*, 314, pp. 572, **1997**.
- [163] *SPSS Professional Statistics*, SPSS Inc., **1993**.
- [164] Barkto, J.J., The intraclass correlation coefficient as a measure of reliability, *Psychological Reports*, 19, pp. 3-11, **1966**.
- [165] Donner, A., Eliasziw, M., Sample size requirements for reliability studies. *Statistics in medicine*, 6, pp. 441-8, **1987**.
- [166] Downie, N.M., Heath, R.W., *Basic Statistical Methods*, Harper and Row, **1970**.
- [167] Beckwith, M., Restle, F., Process of enumeration, *Psychological Review*, **73**, pp. 437-444, **1966**.
- [168] Teulings H.L., Maarse, F.J., Digital recording and processing of handwriting movements, *Human Movement Science*, 3, pp. 193-217, **1984**.
- [169] Maamari, F., Plamondon, R., Extraction of the analog pen tip position, velocity and acceleration signals from a digitizer, in *Ergonomics: Contemporary Research in Handwriting* (eds. Keo, S.R., Van Galen, G.P., Hoosain, R.), Elsevier Science, Netherlands, **1986**.
- [170] Whitefield, A., Human factors aspects of pointing as an input technique in interactive computer systems, *Applied Ergonomics*, 17, pp 97-104, **1986**.
- [171] Hollerbach, J.M., An oscillation theory of handwriting, *Biological Cybernetics*, 39, pp. 139-156, **1981**.
- [172] Plamondon, R., A kinematic theory of rapid human movements, *Biological Cybernetics*, 72, pp. 295-307, **1995**.
- [173] Dooijes, E.H., Analysis of handwriting movements, *Acta Psychologica*, 54, pp. 99-114, **1983**.
- [174] Schomaker, L.R.B., Plamondon, R., The relation between pen force and pen-point kinematics in handwriting, *Biological Cybernetics*, 63, pp. 277-289, **1990**.
- [175] Wann, J., Nimmo-Smith, I., The control of pen pressure in handwriting: a subtle point, *Human Movement Science*, 10, pp. 223-246. **1990**.
- [176] Kline, Paul, *The handbook of psychological testing*, Routledge, London, **1995**.
- [177] Davies A. D., Davies D. R., The effects of noise and time of day upon age differences in performance at two checking tasks. *Ergonomics*, **18**, pp. 321-336. **1975**.

- 
- [178] Mennemeier M., Chatterjee A., Heilman K. M. A comparison of the influences of body and environment centered reference frames on neglect. *Brain*, **117**, pp. 1013-1021. **1994**.
- [179] *Wacom software interface reference manual WTC-1002*, Wacom Company Ltd., Vancouver, **1996**.
- [180] Willams, C.S., *Designing Digital Filters*, Prentice Hall, New Jersey, **1986**.
- [181] Avro, J., *Graphic Gems II*, Academic Press, **1991**.
- [182] Press, W.H., Flannery, B.H., Teukolsky, S.A., Vetterling, W.T, *Numerical Recipes in C*, Cambridge University Press, **1988**.
- [183] Girden, E. R., *ANOVA Repeated Measures*, Sage Publications. **1992**.
- [184] Bryman, A., Cramer, D., *Quantitative data analysis with SPSS for Windows*, Routledge. **1999**.
- [185] *SPSS Base 8.0, User's Guide*, SPSS Inc. **1998**.
- [186] Everitt, B.S., *Making Sense of Statistics in Psychology*, Oxford University Press. **1996**.
- [187] Howell, D.C., *Statistical Methods for Psychology*, Duxbury. **1997**.
- [188] Greenwood, P. E., *A Guide to Chi-Squared Testing*, John Wiley & Sons. **1996**.
- [189] McRae, D.J., MIKA, a FORTRAN IV iterative K-means cluster analysis program, *Behavioral Science*, 16, pp. 423-424, **1971**.
- [190] Brizzotti, M.M., Carvalho, A.C.P.L.F., The influence of clustering techniques in the RBF networks generalization, *Proc. IPA '99*, pp. 87-93, Manchester, **1999**.
- [191] Everitt, B., *Cluster Analysis*, Heinemann Educational, **1974**.
- [192] Ho, T.K., Hull, J.J., Srihari, S.N., Decision combination in multiple classifier systems, *IEEE Trans. PAMI*, **16**, pp. 66-75, **1994**.
- [193] Rahman, A.F.R., Fairhurst, M.C., An evaluation of multi-expert configurations for recognition of handwritten numerals. *Pattern Recognition*, **31**, pp. 1255-1273, **1998**
- [194] Waldemark, J., An automated procedure for cluster analysis of multivariate satellite data, *International Journal of Neural Systems*, **8**, pp.3-15, **1997**.

## Appendix A : Stroke Test Subject Details

## Neglect Subjects

Hospital Number	Date of Birth	VST ID Number	Type of Stroke	Stroke Location	BIT Conventional Total Score	Gender
F744836	10/11/06	31	infarct	temporal lobe/ occipital lobe/ basal ganglia	41	Female
F438646	05/10/38	1009	infarct	post parietal	42	Female
F357581	21/09/27	56	infarct	MCA	53	Male
F350938	27/12/21	75	infarct	R occipital	61	Male
F007397	16/06/23	53	infarct	temporoparietal	62	Female
F765727	01/01/35	67	infarct	post parietal	64	Male
F407588	20/01/24	1	haemorrhage	occipital/thalamic	67	Male
F439579	19/06/11	1022	infarct	temporo-parietal	67	Female
F216974	13/07/26	33	infarct	occipital/cerebral	71	Male
F363591	03/01/16	1024	infarct	MCA	71	Male
F377254	11/07/30	1059	infarct	R Temporo-parietal	71	Male
F433666	09/08/24	57	haemorrhage	occipital	74	Male
F884079	15/03/22	41	infarct	parietal	76	Male
F003113	07/10/12	60	infarct	basal ganglia/internal capsule	79	Female
F420317	02/10/12	10	ischemia	preventricular	80	Female
F780394	24/07/26	63	infarct	parietal	82	Female
F047458	09/03/19	1016	infarct	occipital/internal capsule	84	Male
F821694	31/05/22	77	infarct	R parietal	88	Female
F277937	11/06/28	8	infarct	R fronto-parietal / part of occipital	96	Male
F203205	16/10/22	52	infarct	temporoparietal	101	Female
F168984	27/07/25	1008	infarct	parietal	105	Female
F423353	21/05/32	11	infarct	R parietal	109	Male
F879301	14/04/20	65	ischaemia	periventricular ischaemia	110	Female
F406742	19/06/17	72	infarct	lateral ventricle	110	Male
F088451	23/12/17	22	infarct	R periventricular	121	Male
F2005863	09/11/20	1012	infarct	preventricular	122	Female
F811491	07/11/09	55	infarct	parietal	123	Male
F345612	11/02/22	1021	infarct	parietal	124	Male
F380550	02/02/22	1045	infarct	R basal ganglia/internal capsule	128	Male
F109992	15/02/35	1013	infarct	basal ganglia/parietal	129	Female

## Stroke Control Subjects

Hospital Number	Date of Birth	VST ID Number	Type of Stroke	Stroke Location	BIT Conventional Total Score	Gender
F037650	15/10/19	1023	haemorrhage	parieto-occipital	132	Female
F051491	13/08/09	1029	infarct	parietal	133	Female
F326533	13/06/21	1049	infarct	MCA	133	Male
F113083	18/05/97	37	infarct	internal capsule	134	Male
F100800	17/04/19	1033	infarct	R basal ganglia/internal capsule	135	Female
F717102	24/11/24	1035	infarct	R parietal	135	Male
F177345	25/11/24	58	infarct	internal capsule	136	Female
F809349	30/09/20	1036	infarct	R temporo-parietal	138	Male
F306496	22/04/37	1031	infarct	MCA	139	Male
F715839	29/01/29	1037	haemorrhage	MCA	139	Female
F402196	01/07/27	18	no CT found		140	Male
F419247	24/12/18	2	infarct	internal capsule	141	Male
F394883	25/01/22	13	infarct	cerebellar lacunar/intraventricular internal capsule	142	Male
F847844	03/01/19	64	infarct	frontal lobe	142	Male
F324222	14/08/18	12	generalised atrophy	no specific localisation	143	Male
F134065	24/04/37	16	infarct	temporo-parietal	143	Male
F738647	03/09/19	43	infarct	parietal	143	Female
F258470	09/01/12	25	infarct	internal capsule	144	Male
F204606	28/08/28	38	infarct	temporo-parietal	144	Male
F085014	11/06/14	3	infarct	temporal	145	Female
F329770	20/07/14	15	infarct/haemorrhage	subarachnoid/cerebral/parietal	145	Female
F400693	24/03/13	19	ischaemia	no focal lesion	145	Male
F350671	27/05/15	1026	infarct	parieto-occipital	145	Male
F088768	06/07/29	5	infarct	R internal capsule	146	Female
F411537	19/07/30	6	infarct	R fronto-parietal	146	Female
F126212	02/11/10	7	haemorrhage	subarachnoid	146	Male
F091509	15/04/24	14	infarct	intercerebral/R ventricular	146	Female
F192888	20/04/37	17	infarct	parietal	146	Female
F320194	14/03/18	23	infarct	internal capsule	146	Male
F371321	03/01/33	26	no CT found		146	Female

F423085	13/03/22	28	haemorrhage	posterior cerebrum	146	Female
F371301	17/10/19	29	haemorrhage	lateral ventricle	146	Male
F427663	19/06/23	34	infarct	middle cerebral artery	146	Female
F153959	25/06/32	35	no CT found		146	Female
F021895	13/12/22	36	infarct	internal capsule	146	Male
F371351	03/05/09	39	no CT found		146	Female
F204024	16/09/35	40	infarct	MCA and basal ganglia	146	Male
F845107	28/07/19	42	infarct	parietal	146	Male
F854252	31/08/20	46	infarct	cerebellar	146	Male
F751299	31/07/33	48	infarct	parietal	146	Male
F880804	25/03/38	50	infarct	cerebral	146	Male
F751059	26/04/40	69	infarct	parasagittal and parietal	146	Male
F027676	16/05/10	70	infarct	cerebellar	146	Female
F168984	09/11/20	71	infarct	parietal	146	Female
F803069	12/03/14	76	infarct	R parietal	146	Female
F211687	06/04/35	1003	infarct	cerebellar / fronto-parietal	146	Female
F745342	11/08/25	1007	infarct	parietal	146	Male
F797054	08/05/22	1011	infarct	parietal	146	Female
F412939	11/08/12	1017	infarct	R parietal	146	Male
F321410	19/01/27	1038	infarct	Basal ganglia	146	Male
F870528	07/11/22	1040	infarct	occipital/parietal	146	Male
F2007556	11/11/12	1041	infarct	parietal	146	Male
F091879	26/01/30	1043	infarct	R lateral ventricle	146	Female
F300434	04/02/14	1050	infarct	R basal ganglia/internal capsule	146	Female
F444739	20/07/21	1056	haemorrhage	R frontal	146	Female
F450948	03/01/22	1057	haemorrhage	R parietal	146	Female
F017719	12/11/14	1060	atrophy	No focal lesions	146	Male
F878495	22/08/16	1062	infarct	cerebellar	146	Female

## Appendix B : Computer based assessment script used by therapists administering test

*Italic text indicates non-verbal instructions*

### Point Location Task:

- 1 Look at this page. I would like you to press this pen (*Hand client the pen*) down on any black circles that you can see. When you have finished I would like you to put your pen down. Please start when I say start.
- 2 Look at this page. I would like you to press the pen down on any black circles that you can see. When you have finished I would like you to put your pen down. Please start when I say start.
- 3 Look at this page. I would like you to press the pen down on any black circles that you can see. When you have finished I would like you to put your pen down. Please start when I say start.
- 4 Look at this page. I would like you to press the pen down on any black circles that you can see. When you have finished I would like you to put your pen down. Please start when I say start.

### Cancellation Task :

- OX 1 Cancellation This page is made up of the letters X (*Point to midline X*) and O (*Point to midline O*). Look at the page carefully and cross out all of the letter O's that you can see. Try not to miss any and when you have finished please put your pen down.
- OX 2 Cancellation Again this page is made up of the letters X (*Points to midline X*) and O (*Points to midline O*). Look at the page carefully and cross out all of the letter O's that you can see. Try not to miss any and when you have finished please put your pen down.
- Star Cancellation This page contains stars of different sizes. Look at the page carefully – this is a small star. Every time you see a small star, cross it out like this (*Illustrate by crossing out the two small stars immediately above the centralising arrow on the stimulus sheet*). I would like you to go through this page and cross out all the small stars without missing any of them.
- Albert's Test On this page we have many lines pointing in different directions. Follow my pen as I indicate these lines (*Move pen right to left, top to bottom over all the lines on the page*). Now with this pen, I want you to cross out all the lines which you can see on the page, like this (*Illustrate by crossing out two of the four central lines*). Some patients may initially cross out only those lines which appear to correspond to the orientation of the example. In such a case the patient should be instructed to cross out all the lines irrespective of orientation.

## Kinematic Analysis Tasks :

- 1 Look at this page. There are black circles on the right (*Point to column on subjects right*) and left side of this page (*Point to the column on the subjects left*). I want you to join these black circle starting from the right hand side moving to the left then back to the next dot on the right (*Demonstrate by tracing the route with your finger*). Join all of the dots in this way and when you have finished put your pen down.
- 2 On this page are a series of black circles and squares. I would like you to join all of the circles together starting from the right side (*subjects right*) going to the left side and back to the right (*Demonstrate by tracing the route with your finger*). When you have finished put down your pen.
- 3 Again on this page are a series of black circles and squares. I would like you to join all of the circles together starting from the right side (*subjects right*) going to the left side and back to the right (*Demonstrate by tracing the route with your finger*). When you have finished put down your pen.
- 4 On this page are a series of black circles, each has a number. I would like you to join these circles in the numerical order they appear. So, 1 to 2, 2 to 3 and so on. When you have finished please put your pen down.
- 5 On this page are a series of black circles, each has a number. I would like you to join these circles in the numerical order they appear. So, 1 to 2, 2 to 3 and so on. When you have finished please put your pen down.
- 6 Look at this page. Here are some black circles join them together to form a square.
- 7 Look at this page. Here are some black circles join them together to form a square.
- 8 Look at this page. Here is a central black circle (*Point at the circle*). Around the central circle are a series of other circles (*Point to each of the circles in the outer ring*). I want you to join these circles together. Starting from the central circle I want you to draw a line out to the first circle then back to the central circle then back out to the next circle. Join all the circles and when you have finished please put down your pen.

## Line Bisection Tasks :

- 1 Please draw a line where you think the middle of the line should be.
- 2 Please draw a line where you think the middle of the line should be.
- 3 Please draw a line where you think the middle of the line should be.
- 4 Please draw a line where you think the middle of the line should be.
- 5 Please draw a line where you think the middle of the line should be.
- 6 Please draw a line where you think the middle of the line should be.
- 7 Please draw a line where you think the middle of the line should be.
- 8 Please draw a line where you think the middle of the line should be.

