LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The Accommodation of Cognitive Style in the Design of the Human Computer Interface

Adrian Brendan Kenneth Parkinson

Thesis submitted in accordance with the requirements for the degree of
Doctor of Philosophy

University of Dublin,
Trinity College,
Department of Computer Science

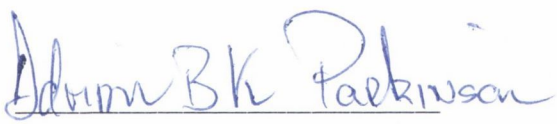
October 2006

TRINITY COLLEGE
27 MAY 2008
LIBRARY DUBLIN
THESIS
8448

Declaration

This thesis has not been submitted as an exercise for a degree at any other university.

Except where stated, the work described herein was carried out by me alone.

Signed: 

Date: 6/11/06.

Permission to lend or copy

I, the undersigned agree that the Trinity College Library may lend or or copy this thesis upon request.

Signed:



Adrian Parkinson

Date: 25th October 2006

Summary

How the design of the human computer interface could be varied to accommodate specific cognitive styles is addressed.

The area of adaptive versus non-adaptive systems was discussed in Chapter 2. Generally non-adaptive systems were favoured due to the difficulties pertaining to adaptable and adaptive systems.

Chapter 3 reviews the Cognive Styles' literature. Generally there have been concerns about the reliability and validity of cognitive styles. The definition and measurement of cognitive styles is still problematic.

The accommodation of cognitive style in interface design is discussed in Chapter 4. This discussion involves i. a review of the literature, ii. two empirical studies examining the accommodation of individual differences (cognitive styles) in the human computer interface (HCI), and iii. a consideration of the matching/mismatching hypothesis which underlies their accommodation. These two studies are evolutionary and lead on to the direct accommodation of cognitive style in the HCI by means of a test.

The reliability and validity of Riding's CSA is discussed in Chapter 5. It was established that Riding's CSA was unreliable, in particular in the Verbal - Imagery (V-I) dimension. Hence the validity of the many studies based on the CSA is questionable. These findings necessitated the design and implementation of a new test of the V-I dimension.

Chapter 6 describes the design and implementation of the new test of cognitive style, VItrap. The reliability of the VItrap in two studies was estimated to be about 0.32, which was much better than Riding's (which fluctuated from -0.21 to about -0.19) over similar test - retest intervals.

The validity of the VItrap was examined in Chapter 7. The definitions of the many types of validity are discussed. The VItrap was shown to be independent of measures of Intelligence and Personality.

The neurophysiological basis of the VItrap test is examined in Chapter 8 using EEGs. The Chapter presents an overview of the EEG and its measurement and a review of the literature on Hemisphericity. Hemisphericity was an important underlying concept in how Riding attempted to establish the neurophysiological basis of his CSA. A neurophysiological study of the VItrap was done in an attempt to broadly

replicate the findings of Riding and his co-workers in 1997. Riding's results could not be replicated. Hence the neurophysiological validity of the VItrap is still an open question.

Chapter 9 examines some prototype applications that incorporated the new test test of cognitive style in the HCI. Of the three prototypes considered, and of the two implemented, the Route Simulators' application proved the most insightful. The QUIS was used to evaluate the usability of the interface. The data from Route Simulator studies consistently demonstrated that matching favoured the Imager and mismatching the Verbaliser.

Chapter 10 summarises, gives conclusions and suggests future work.

Acknowledgements

I wish to thank the following persons who helped me in different ways: Professor Denise Maikowicz, Professor Robert Sternberg and Dr. Cliff Redmond. I would also like to thank Dr. Micheál O'Maoldomhnaigh, my friend and former lecturer, for obtaining, administering and scoring Kirton's KAI and for support throughout my academic career. I would like to thank my sister, Ms. Suzanne Parkinson for getting me started on this work (or laborious trail!) and for encouragement and support in many ways.

I would like to thank Professor Astrid O'Moore for help in obtaining test materials, conducting and overseeing some of the psychometric testing sessions. I am indebted to Professor Barry Sterman for his expertise, trust and friendship in the training and use of his Neuronavigator. I must also thank Dr. David Kaiser of SKIL for his most helpful advice whenever called on. Ms. Tamara Lorenson gave great advice on the practicalities of getting started with the Neuronavigator and how best to equip for it.

I thank Professor Ben Shneiderman and Professor Jenny Preece for many discussions and emails. I am grateful for Professor Jurec Kirakowski for the use of his test of usability - Wammi. I thank Dr. Cathal Walsh for his statistical advice, friendship, dynamism and continual good humour. I am grateful to Ms. Elizabeth Oldham for introducing me to postgraduate research in TCD.

I would particularly like to thank the Elsin group who have been a source of inspiration, feedback, friendship and support in particular: Dr. Steve Rayner, Professor Eugene Sadler-Smith, Dr. Steve Armstrong, Dr. Martin Graff, Dr. Carol Evans and especially Dr. Elizabeth Peterson.

I would also like to thank the IS5 students over the years who have co-operated with my research efforts.

I thank Mrs. Rosemary Welsh for tirelessly solving computer problems at awkward times usually at short notice. I wish to thank Alan Mullally for his support in many areas and his trust in my ability. I wish to thank all my work colleagues for their efforts when called on especially Mrs. Mary McAuley, Ms. Evelyn Lennon, Mrs. Irene Devitt and Mr. Eamon Burke.

I want to thank my wife, Bríd Moran, for her support, friendship and endless patience.

I want to thank my parents for their love and support and their interest in encouraging the family members to fulfil their own unique potential.

Lastly, but not least, I would like to thank my supervisor, Dr. James A. Redmond, for his support and friendship in arduous times.

I wish to dedicate this work to my parents, Eileen and Ken.

Table of Contents

Declaration ii
Permission to Lend or Copy iii
Summary iv
Acknowledgements vi
Dedication vii

Chapter 1 - Introduction 1

Chapter 2 - A Review of the Flexibility of Systems 7
2.1 Introduction 7
2.2 Non–adaptable (Static) systems 7
2.3 Adaptable Systems 8
2.3.1 The Adaptable System: Direct Manipulation Systems 9
2.3.2 Adaptive Systems 12
2.3.3 Models of Adaptive Systems 15
2.3.4 Evaluation of the Push System 20
2.4 Evaluation of Adaptive Systems 21
2.5 Problems of Adaptive Systems Design 23
2.5.1 Usability 23
2.5.2 User Modelling 24
2.5.3 Predictability & Transparency 25
2.5.4 Controllability 26
2.5.5 Unobtrusiveness 25
2.5.6 Privacy 27
2.5.7 Breadth of Experience 27
2.6 Progress to Date in Adaptive Systems 28
2.7 Conclusions 29

Chapter 3 A Review of the Cognitive Styles Domain 31
3.1 Distinctions between Cognitive, Learning & Thinking Styles 31
3.2 The Importance of Cognitive Style 33
3.3 Definitions of Cognitive Style 34
3.4 Cognitive Style and Ability 37
3.5 Cognitive Style, Learning Strategy & Learning Skill 38
3.6 An Overview of the Style Field 40
3.7 The Coffield Report (2004) 42
3.7.1 Existing Difficulties in the Style Area 44
3.7.2 Three Style Traditions after Sternberg & Zhang 45
3.7.3 Curry's Style Model 46
3.8 Riding and Cheema's Model 47

3.9 Coffield's Continuum of Style Families 49
3.10 Review of some of the main cognitive styles 50
3.10.1 Field dependence – Field independence 50
3.10.2 Leveling and Sharpening 52
3.10.3 Convergent – Divergent Style 53
3.10.4 Concrete – Abstract 54
3.10.5 Intuition – Analytic Style 54
3.10.6 Impulsivity – Reflectivity 55
3.10.7 Holists and Serialists 55
3.10.8 Adaptor – Innovator 56
3.10.9 Models NOT Reviewed 57
3.10.10 Riding's CSA 57
3.10.11 Peterson, Deary and Austin's Work 58
3.10.12 Individual Differences Questionnaire 59
3.10.13 Verbaliser – Visualiser Questionnaire 59
3.11 Conclusions 60

Chapter 4 Cognitive Styles & Human Computer Interface Design 62
4.0 Introduction 62
4.1 Introduction to Study 4.1 63
4.1.2. Study 4.1 Design 64
4.1.3. Study Implementation 65
4.1.4 Treatments 65
4.1.5 Treatment 1 – Text via Floppy Disk 66
4.1.6 Treatment 2 – Multimedia Cd-Rom 66
4.1.7 Treatment 3 – Internet Web-site 66
4.1.8 Learning Performance – Evaluation 67
4.1.9. Psychometric Tests 68
4.1.10. Aim of Study 4.1 68
4.1.11. Results of Study 4.1 68
4.1.12. Discussion of Study 4.1 68
4.1.13 Design Recommendations 69
4.1.14. Conclusions of Study 4.1 69
4.2 The Matching / Mismatching Hypothesis 70
4.2.1 Evolution of Matching / Mismatching Hypothesis 70
4.2.2 Early Discussion of the Matching Hypothesis 71
4.2.3 Riding and the Matching Hypothesis 71
4.2.4 Recent Discussion of the Matching Hypothesis 73
4.2.5 Cognitive Style and HCI design 76
4.3 Study 4.2 Support to give equal performance in a HCI? 80
4.3.1 Support for the Fd Learner - Web-based 81

4.3.2 Motivation for Study 4.2 83
4.3.3 Method for Study 4.2 83
4.3.4 Study 4.2 Design 83
4.3.5 Study 4.2a 84
4.3.6 Study 4.2b 84
4.3.7 Measures 84
4.3.8 Psychometric Test 85
4.3.9 Apparatus 85
4.3.10 Stimulus Materials for Study 4.2 85
4.3.11 Procedure for Study 4.2 85
4.3.12 Learning Performance - Evaluation 85
4.3.13 Website Design 86
4.3.14 Results of Study 4.2 87
4.3.15 Discussion of Studies 4.2a and 4.2b 88
4.3.16 Conclusions of Studies 4.2a and 4.2.b. 90
4.4 Introduction to Study 4.3 91
4.4.1 Method for Study 4.3 93
4.4.2 Apparatus - Description of Website 93
4.4.3 Results of Study 4.3 95
4.4.4 Discussion of Study 4.3 96
4.5 Results of the Three Pilot Studies 4.1, 4.2 and 4.4 97
4.6 Conclusions 97

Chapter 5 The Validity & Reliability of Riding's CSA 100
5.1 Introduction 100
5.2 Validity 101
5.2.1 Validity of the CSA 102
5.2.2 Origins of the CSA 106
5.2.3 The Basis of the CSA - Inspection Time 106
5.2.4 Description of the CSA computerized test 107
5.3 Study 5.1 of Riding's CSA - Ambiguity of Test Items 108
5.3.1 Method 109
5.3.2 Results of Study 5.1 - Ambiguity of Test Items 109
5.3.3 Conclusion of Study 5.1 116
5.4 Study 5.2 of Riding's CSA - Test - Retest Reliability 116
5.4.1 Introduction 116
5.4.2 Methodology 117
5.4.3 Results of Study 5.2 119
5.5 Discussion of Studies 5.2.1, 5.2.1 and 5.2.3 120
5.6 Conclusions 125

Chapter 6 The Design of a New Test of the V-I Dimension - VItrap 127
6.0 Introduction 127
6.1 Test Item Design 128
6.2 Implementation 132
6.3 Studies to Estimate VItrap Reliability 132
6.3.1 Study 6.1 To test the reliability of VItrap 133
6.3.2 Results of Study 6.1 134
6.4 Study 6.2 Second test retest of the reliability of VItrap 135
6.4.1 Results of Study 6.2 135
6.5 Study 6.3 Effect of Incorrectly Answering Test Items 135
6.5.1 Session 6.3.A (Test results) 136
6.5.2 Session 6.3.B (Retest results) 136
6.5.3 Results of Study 6.3 137
6.5.4 Conclusions of Study 6.3 138
6.6 Correlation between times for Test & Retest 138
6.6.1 Reliability when non-English speakers removed 138
6.6.2 Reliability when 6 questions incorrect removed 138
6.7 Summary on the assessment of cognitive style 139
6.8 Discussion 139
6.9 Improving The Test 144
6.10 Conclusions 144

Chapter 7 Validation of the VItrap 145
7.0 Introduction 145
7.1 What is meant by the validation of a psychometric test? 145
7.2 Cognitive Style and Intelligence 148
7.2.1 Study 7.1 - VItrap vs. Intelligence (Raven's, Mill Hill) 149
7.2.2 Results of Study 7.1 150
7.2.3 Conclusions of Study 7.1 150
7.3 Study 7.2 Is VItrap vs. Personality (EPI, STAI) 151
7.3.1 Study 7.2.1 Is VItrap independent of the EPI? 151
7.3.2 Results of Study 7.2.1 151
7.3.3 Study 7.2.2 Is VItrap independent of personality (STAI)? .. 152
7.3.4 Results & Conclusions of Studies 7.2.1 & 7.2.2 154
7.3.5 Overall Conclusions on Personality & the VItrap 154
7.4 Study 7.3 VItrap vs. IPQ 156
7.4.1 Introduction to Study 7.3 156
7.4.2 Results of Study 7.3 160
7.4.3 Internal Reliability of IPQ 160
7.4.4 Conclusions of Study 7.3 161
7.5 Study 7.4 VItrap vs. ILS 162
7.5.1 Introduction to Study 7.4 162

7.5.2 Description of the ILS 162
7.5.3 Results & Conclusions of Study 7.4 on ILS 164
7.6 Study 7.5 Investigating the Convergent Validity of VItrap 164
7.6.1 Introduction to the VVIQ 165
7.6.2 Methods 166
7.6.3 The Test-Retest Reliability of the VVIQ 166
7.6.4 The Internal Reliability of the VVIQ 167
7.6.5 The Test-Retest Correlation between Vitrap and VVIQ 167
7.6.6 Factor Analysis 167
7.6.7 Conclusions on Study 7.5 on VVIQ 165
7.7 Comment on Reliability and Validity 170
7.8 Overall Chapter Conclusions 170

Chapter 8. A Neurophysiological Approach to Validating VItrap 172
8.0 Introduction 172
8.1 Brief overview of the EEG 172
8.1.1 Hemisphericity of the Brain 175
8.1.2 Hemisphericity and the lower animal kingdom 176
8.1.3 Language hemispheric specialisation 176
8.1.4 Visuo-spatial processing, visual imagery & hemisphericity.. 178
8.2 The Work of Riding and his Co-workers using EEG 181
8.2.1 Basis for the Riding and Glass (1997) study 184
8.2.2 Method of Riding and Glass 186
8.2.3 Results of Riding and Glass 186
8.2.4 Conclusions of Riding and Glass 187
8.3 Study 8.1 Study of the Neurophysiological Basis for the VItrap 187
8.3.1 Introduction to Study 8.1 187
8.3.2 Method 188
8.3.3 Filter Settings used in the Neuronavigator 188
8.3.4 Description of method 188
8.3.5 Preparation of Participants for EEG recording 188
8.3.6 Fitting the Electrode cap 190
8.3.7 Impedances 191
8.3.8 Earlobes 191
8.3.9 Sample 192
8.3.10 Task One: Eyes Closed - Ec 193
8.3.11 Task Two 193
8.3.12 Task Three 194
8.3.13 Statistical Analysis of Data 194
8.3.14 Calculation of an Approximate Cognitive Style Ratio 194
8.3.15 Data Processing & Hardware Calibration 195
8.3.16 Interactions of interest 195

8.3.17 Statistical Analysis 196
8.4 Results of Study 8.1 196
8.4.1 Midline cluster (Fz, Pz, Cz) 196
8.4.2 Discussion of Study 8.1 197
8.5 Conclusions 200

Chapter 9. Design and Implementation of Prototypes using VItrap 201
9.0 Introduction 201
9.1: Design of Route Simulator in Authorware (Cd-Rom) 202
9.1.1 Design & Description of the CD based route simulator 203
9.1.2 The Four Experiments 204
9.1.3 Authorware 205
9.1.4 Procedure 207
9.1.5 Route Simulator Study 9.1: Analysis of data 208
9.1.6 Psychometric Tools 208
9.1.7 VItrap 209
9.1.8 Study 9.1 Results (n = 168) 209
9.1.9 Study 9.1 Experiment 2 211
9.1.10 Study 9.1 Experiment 3 212
9.1.11 Study 9.1 Experiment 4 214
9.1.12 Combined Plot for all Four Experiments and Groups 215
9.1.13. The Internal Reliabilty of a Truncated Quis 216
9.2 Study 9.2 (IS5 Group Route Simulator (n = 40)) 216
9.2.1 Introduction - Study 9.2 CD-rom Study with IS5 216
9.2.2 Method 217
9.2.3 Results of Study 9.2 Experiment 1 217
9.2.4 Study 9.2 Experiment 1 Analysis 218
9.2.5 Study 9.2 Experiment 2 218
9.2.6 Study 9.2 Experiment 3 220
9.2.7 Study 9.2 Experiment 4 221
9.3 Study 9.3 Web-based Route Simulator (n = 20) 223
9.3.1 Introduction to Study 9.3 223
9.3.2 Design of the Web-based Route Simulator 224
9.3.3 Method 224
9.3.4 Design of the Application 224
9.3.5 Screenshots of the application 225
9.3.6 Introduction - Study 9.3 Web-based Route Simulator 227
9.3.7 Results of Study 9.3 Web version of Route Simulator 228
9.3.8 Study 9.3 Web - Experiment 2 229
9.3.9 Study 9.3 Web - Experiment 3 230
9.4 Study 9.4 and Study 9.5 233
9.5 Discussion Pertaining to Studies 9.1, 9.2 and 9.3 233

9.6 Quis Analysis - Usability Results - Study 9.1 CD route simulator .. 235
9.6.1 (a) General reaction to the system 235
9.6.2 (b) Ease of interface interpretation 236
9.6.3 (c) Ease of use 236
9.6.4 (d) Evaluation of the tutorial 236
9.6.5 (e) Quality of photographs 236
9.6.6 Discussion of the Quis as an Evaluatory Tool 237
9.7 Discussion of Matching & V-I after Riding & Co-Workers 238
9.8 Limitations of these Studies 239
9.9 Overall Conclusions 240

Chapter 10 Conclusions 242
10.1 Purpose 242
10.2 Limitations of this Thesis 246
10.3 Contributions to Knowledge 246
10.3.1 Major Contributions to Knowledge 246
10.3.2 Minor Contributions to Knowledge 247
10.4 Future Directions 247

References 248

Appendices
Appendix 1-2.1.A Rothrock's Generic Adaptive Model -Table A 281
Appendix 2-2.2.B Rothrock's Review of Eighteen Adaptive Systems 282
Appendix 3-4.1.A Pilot Study (Study 4.1) 286
Appendix 4-4.2.B Statistics for Geft Scores - Studies 4.1, 4.2a & 4.2b 287
Appendix 5-4.3.C Statistics for Exam. Scores - Studies 4.1, 4.2a & 4.2b 288
Appendix 6-5.1.A Additional Statistics - Riding's CSA 289
Appendix 7-5.2.B Study 5.2.2 Analysis W-A & V-I vs. APM2 & Mill Hill ... 290
Appendix 8-6.1.A Spearman Browne Prophecy Formula 291
Appendix 9-6.2.B Inspection Times - Verbal Forms vs. Image Forms 292
Appendix 10-6.3.C VItrap: Total Correct - Test & Retest 293
Appendix 11-6.4.D VItrap: Additional Statistics - Test & Retest 295
Appendix 12-6.5.E Survey of Potential Test Item clarity 299
Appendix 13-6.6.F Summary Statistics for the VItrap Test 303
Appendix 14-6.7.G Post Hoc Test Results - Test & Retest 304
Appendix 15-7.1.A Vividness of Visual Imagery Questionnaire 307
Appendix 16-7.2.B Summary Statistics For Felder's Scale 309
Appendix 17-8.1. A Sample EEG Scans via SKIL Analysis Software 311
Appendix 18-9.1 A Usability Histograms pertaining to Study 9.1 313
Appendix 19-9.2.B Golf Application 1 (Ultan) 332
Appendix 20-9.3.C Golf Application 2 (Piotr) 335

Appendix 21-9.4.D Web screenshots of the Route Navigator 338
Appendix 22-9.5.E Quis 344
Appendix 23-9.6.F: Statistical Analyses for Studies 9.1, 9.2 and 9.3 349
Appendix 24-9.6.F CD Study 9.2 (n = 40) 353
Appendix 25-9.6.F Study 9.2 Experiment 2 Results 354
Appendix 26-9.6.F Study 9.2 Experiment 3 Analysis 355
Appendix 27-9.6.F Study 9.2 Experiment 4 Analysis 356
Appendix 28-9.6.F Study 9.3 Web version of Route Simulator (n = 30) 357
Appendix 29-9.7.G Quis: Statistical Data for each of the Questions 360

(**N.B. Key to Appendix Nomenclature:** For example, in the title Appendix 15-7.2.B, 15 means Appendix number 15, 7.2.B stands for second Appendix in Chapter 7, while B (also means second but is used as an index appended to Appendix Figure and Table labels to indicate they are from an Appendix rather than the main text).

Chapter 1:
Introduction

This thesis examines the facilitation of individual differences in the design of the human computer interface. Generally it is important to cater for individual differences in how systems are designed for everyday use.

The human computer interface has become of more and more importance with the continued evolution of computer power. As computers become more powerful, much more power becomes available for matching their interfaces to individual human needs. Because of the variety of individual needs, there is no one way for designing effective interfaces. The fact remains that in structuring information within an information system, the systems designer is designing the way in which the user must interact with the system in order to access the information (broadly his/her cognitive style) and is ipso facto 'designing' the way in which that person does his/her thinking.

This thesis considers the question whether the accommodation of individual differences such as cognitive style in the human computer interface proves advantageous for user performance. When accommodating individual differences such as this, one might realistically assume better user performance and satisfaction would result.

There may also be other advantages that could accrue such as safer use of computer systems resulting in less catastrophic accidents, greater utilisation and more effective use of the interface by those who use them. Given the increasing reliance on human computer interfaces in society whether browsing the Net, or engaging in such activities as eLearning, eBusiness and eGovernment, this study is timely.

Employment opportunities are some times constrained by such differences as: ability, motivation, skills, personality types and physical characteristics. Indeed even seemingly trivial applications such as computer dating agencies use such characteristics in matching prospective partners. Individual Differences also include cognitive styles that can be described as being preferred ways of habitually representing and processing

information. The definitions of cognitive style are more closely addressed in Chapter 3.

Individual Differences are defined by the Penguin Dictionary of Psychology (Reber & Reber (2001)) as "a label used for an approach to psychological phenomena that focuses on characteristics or traits in respect of which individual organisms may be shown to differ". Differential psychology is the area of psychology that concentrates on the study of individual differences. Cooper (2002) describes the psychology of individual differences as seeking "to understand how we should best describe the ways in which people vary and to understand how and why such variation comes about". Theories that test how people are hypothesized to behave, and vary, are dependent on the field of psychometrics. This can be thought of as the accurate measurement of Individual Differences, and has been defined as "pertaining to the measurement of that which is psychological". Hence the definition of Psychometrics can be extended to: "Pertaining to mental testing in any of its facets, including assessment of personality, evaluation of intelligence and determining of aptitudes" (Reber & Reber (2001)).

The properties of any psychometric test, most importantly its reliability and validity, require evaluation before the test should be selected for further use. There are different ways of assessing reliability and different types of validity. For example the temporal stability, or the test - retest reliability, examines the degree of association between the same test given to the same cohort on two separate occasions. It might be expected that there should be a high degree of similarity between an individual's scores on both occasions if the test under consideration was stable and reliable and that there was no change in the trait being measured. The validity of a psychometric test establishes whether the test measures what it is supposed to measure. There are different types of validity (discussed later): content validity, construct validity, convergent and divergent validity. A psychometric test with low reliability has little if any validity whereas a test with high reliability does not automatically imply test validity. For example, a weighing scales may measure weight accurately, and while there is a broad correlation between weight and height, the validity of using the scales to estimate height may be very poor and may give unreliable estimates of

height. For instance estimating the heights of a group of Sumo wrestlers versus the heights of a group of basketball players from their weights is unlikely to be accurate. Reliability and validity are addressed in greater depth in Chapters 5, 6 and 7.

Some researchers in the field of Human Computer Interaction have emphasised the importance of individual differences such as cognitive style as a basis for designing the human computer interface. Cognitive style is investigated as a basis for accommodating individual differences in the design of the human computer interface in this thesis. This chapter presents an introduction to each of the subsequent chapters in this thesis, highlighting the applicable issues in each chapter as they arise.

The broad research question is whether Cognitive Style should be accommodated in the design of a human computer interface. A computer-based test is uded to determine the cognitive style of the user and then the HCI system is matched to the user's style. The research design adopted in all studies is experimental performed in computer laboratories in TCD. The target population were largely engineering and computer science students who volunteered. In some experiments the students were paid a small amount. The sample sizes varied from about twelve to about one hundred and eighty five. In each instant the sample were representative of the larger populations from which they were taken. Psychometric tests were purchased. These tests were administered by a qualified educational psychologist where necessary. The data were analysed using a range of appropriate computer statistical packages. Further details are given in more depth where necessary. In many minor ways, This thesis evolved over a period of some years. Some studies were done relatively early with two being done in parallel. The studies are presented in the thesis in chronological order.

In Chapter 2, the main types of flexible system design are discussed. The spectrum of computer systems is discussed ranging from non-adaptive to adaptive systems.

In Chapter 3 an overview of the cognitive style area is presented. Issues such as the definition of cognitive style, characteristics of cognitive style, a description of some of the various style models

and finally a brief account of the most pertinent of the popular style labels are considered.

Chapter 4 reviews the literature that has incorporated cognitive style as a constituent part of interface design. Some of the studies described have scrutinised the usefulness of cognitive style in the various designs. This chapter also addresses the matching/ mismatching style hypothesis. Many of the studies that have accommodated cognitive style have been predicated on the assumption that matching is actually favourable for the user of the system, but this assumption has been questioned by some other researchers.

Riding's test of cognitive style (Riding (1991)) is investigated as to its psychometric properties in Chapter 5. This was necessary before deciding to use it as a basis for designing the human computer interface.

Riding's test was picked for a number of reasons: i. it was easy to use, ii. self-scoring for students, iii. it was relatively inexpensive for a one year license, iv. it did not require an extensive training programme to administer and use it (unlike, for example, Kirton's KAI which required a two day training program, v. it did not require the presence of a qualified wducational psychologist to adminsiter it or interpret the scores (unlike Witkin's Group Embedded Figures Test), vi. the test was readilyavailable on an inexpensive 3.5 inch floppy disk and the results were written to the disk, vii. this test was the first of its kind to assess cognitive style in a computerised format, viii. the test took only about twenty minutes to administer and did not result in user fatigue and finally, ix. being computerised the test had the potential to be directly embedded in the human computer interface. Riding's test met the requirements for test construction and it generally well accepted as evidenced by the large literature based around it.

The validity of Riding's CSA seemed very well validated in studies which examined its construct validity. The temporal stability and the internal reliability of the CSA had not been reported upon previously in the literature before this research. This chapter addresses the topic of the temporal stability and the

internal consistency of the CSA and reviews other characteristics of this computerized test of cognitive style.

Chapter 6 describes the design of a new test of cognitive style, which assesses the Verbaliser - Imager dimension. The test retest reliability and the internal reliability of this new test VItrap is discussed. The new test of the Verbaliser- Imager dimension is discussed in the context of Riding's CSA and Peterson's paper which describes her own new test of this dimension - VICS (Peterson, Deary & Austin (2005a)).

The validity of the new test of cognitive style VItrap is examined in Chapter 7. The approach taken in this chapter is similar to that advocated by Riding (1998). The test is correlated with a number of tests of Intelligence: Raven's Matrices and the Mill Hill Inventory (Court & Raven (1975)). This new test VItrap is also investigated as to independence of measures of personality as assessed by Eysenck's EPI (Eysenk & Eysenck (1975)) and Speilberger's STAI (Speilberger (1977)). Curry's style model (Curry (1983)) is used to categorise the new test of cognitive style (placing it in the innermost core) and its relationship to measures pertaining to the instructional preference layer investigated.

Chapter 8 examines the neurophysiological evidence for cognitive style as assessed by the VItrap. Riding and Glass's studies (Riding, Glass and Douglas (1993), Riding, Glass, Butler, and Pleydell-Pearce (1997)) have reported some physiological evidence for cognitive style and this discussed in detail. This Chapter describes whether some of Riding's findings and general conclusions were replicated in the studies described in this chapter.

In Chapter 9, some prototype applications that incorporated the new test of the Verbaliser - Imager dimension are evaluated. These applications provide insight into the practicality of designing and developing such prototypes and are an opportunity of investigating the matching / mismatching hypothesis. The inadequacies of the Quis (Quis (2006)) and Wammi (Wammi (2006)) are mentioned in assessing the usability and user satisfaction of the prototypes examined in the context of the experimental design.

The thesis concludes by weaving together the main threads of the arguments, giving a discussion of the main conclusions and suggests areas for further work.

Chapter 2:
A Review of the Flexibility of Systems

2.1 Introduction

This chapter provides an overview of the different types of system that can be customised to cater for individual differences.

Systems can be divided up into three types (a) non-adaptable (static), (b) adaptable and (c) adaptive (Opperman (1994)). These systems are discussed and the relative merits of each of the design approaches are outlined.

This review of the different systems has been influenced in part by, and is adapted mainly from, the reviews of adaptive systems undertaken by Jameson (2000) and Rothrock et al. (2000).

Cognitive Style will introduced in Chapter 3 as a dimension along which a system can be adapted.

2.2 Non–adaptable (Static) systems

A non-adaptable (static) system is where the structural template and interface variables have been predefined. In this type of system the designer, consciously or unconsciously, has identified some characteristics of the user and modified the interface to suit those characteristics in the design phase. The designer has a good model of the user and designs the system accordingly.

In some cases depending on the type of application, the general design of a non-adaptable system can exclude specific groups of users, for example a sports car is designed to cater for a relatively fit user and may present difficulties for the obese person. Much of general design is aimed at the "average" user.

Benyon, Crerar and Wilkinson (2001) suggest that interaction between human and computer can be enhanced in a non-adaptable system by means of an 'adaptor': "An adaptor is another way of affording interaction ... that mediates between two incompatible systems". They discuss the role of the computer keyboard in this context. They argue that the keyboard is necessary for data entry as the computer is poor at processing the signals for which humans "would normally use speech or

handwriting". Some modern speech and gesture recognition programs could possibly function as adaptors in this instance.

An in-built test of personality or cognitive style could be used as an adaptor, in a similar fashion to that discussed by Benyon, Crerar and Wilkinson (2001): e.g. a score obtained from a test could be used to select a more suitable interface for the individual user. They state that: "all non-adaptable systems impose all the load of change onto the user."

While there is a trend towards more adaptability in systems, some researchers (Somberg (1987) and Wickens (1992)) suggest that non-adaptive systems "promote superior performance than adaptive ones" (Rothrock, Koubeck, Fuchs, Haas, Salvendy (2002)).

2.3 Adaptable Systems

An adaptable interface is where the user can modify some aspect of the system as he/she is using it. Benyon, Crerar and Wilkinson (2001) categorise this as a customisable system. They look at customising the interface for a car in terms of seat position, steering wheel height and mirror settings. This approach to system design is challenging, as the need for the user to adapt the interface to his style can be distracting to his primary task. This may misdirect the user's locus of attention resulting in an inconsistent performance and lack of satisfaction with the interface. Benyon, Crerar and Wilkinson (2001) assert that "The problem with customisable systems is that they rely on the users to configure them, thus forcing users to do something tangential to their main purpose". For example in MS Word Style sheets are available for use to help customise the document but most users do not use them directly. They cite writing macros as another example of a problem with customisable systems: for instance in the case of Excel, for users to fully exploit the package functionality, they not only need to know the package basics but also should be able to write macros. Basically going off in a sidetrack from the main purpose does not interest many users. Benyon, Crerar and Wilkinson (2001) also maintain that "in a customisable system, the user may be willing to adapt, but cannot understand the customizing system."

Direct Manipulation systems, a subset of Adaptable Systems, have received considerable attention recently in the literature (Shneiderman (1998)).

2.3.1 The Adaptable System: Direct Manipulation Systems

These applications allow the presentation of data visually, and direct manipulation of specific data points. For example specific sets of data can be distinguished from each other by use of colour or icons within the display. When data are presented visually, trends can be identified almost immediately thereby presenting a wholistic view of the analysis. Using this type of application also provides the user in some cases with a greater insight as to the effect of different variables in the statistical calculations.

Shneiderman (1995) gives an example of a direct manipulation interface and discusses it in depth: i.e. a film finder that provided the user with the opportunity to select a film title. In this application the user was able to control the display by zooming in on recent popular films. The y-axis denoted the popularity of the film, the x-axis the year of the film and colour the type of film. Shneiderman (1997) states that "users have greater control over the display and as they select items, the details appear in windows on the sides".

Shneiderman (1997) describes such a system as having "rapid incremental and reversible actions, selection by pointing and immediate feedback". He states that "direct manipulation depends on visual representation of the objects and actions of interest, physical actions or pointing instead of complex syntax, and rapid incremental reversible operations whose effect on the objects of interest is immediately visible (Shneiderman (1997)). He envisages that this direct manipulation interface is an alternative to the agent-based scenarios. He maintains enthusiastic users have reported "mastery of the interface, competence in performing tasks, ease in learning, enjoyment in using the system and a desire to explore more powerful aspects of the system" (Shneiderman & Maes (1997)).

Other applications giving the user a holistic visual overview of the search space are discussed in the article. The user can then filter out what information he/she requires and ultimately

obtains specific details about the query. Shneiderman & Maes (1997) comment "my claim is that this gives the users the feeling of being in control and therefore they can be responsible for the decisions they make".

The disadvantages of direct manipulation systems include applications that may not be suitable because the viewing area (the screen size) may be too small (Shneiderman (1998)). An example is the small size of PDA (Personal Digital Assistant) screens that are not suitable for graphic displays. "Direct manipulation designs may consume valuable screen space and thus force valuable information off screen, requiring scrolling or multiple actions" (Shneiderman (1998)). He also comments, that "users must learn the meaning of components of the visual representation". In other words, how meaningful are icons, "a graphic icon may be meaningful to the designer but may require as much or more learning time than a word". Also visual representation may be misleading; the incorrect representation of an icon may lead to false conclusions about permissible actions. Some other disadvantages include difficulties in programming the interface, and perhaps the possibility that visualisation might not suit expert users (e.g. an expert might prefer to use a Unix command line rather than an icon).

Shneiderman's (1998) direct manipulation interface approximates the adaptable system. He favours the direct manipulation interface as he claims it empowers the user. It provides transparency, a greater sense of predictability and gives the users a sense of control. The features of direct manipulation systems are: rapidity, easy reversibility, incremental actions; high visibility and direct manipulation of the object of interest.

On the other hand some advantages of direct manipulation interfaces include: they are easy to learn, they provide opportunities to explore the data, concepts can be presented visually, the user is in control and the system may be transparent and predictable.

In direct manipulation interface applications, information visualisation is at their very core. The user is provided with the opportunity to interact directly with the interface by

manipulation of the data. However it is also possible that colour-blind users would find these types of applications rather difficult to use.

What is surprising is the fact that Shneiderman assumes that everyone prefers to represent and interrogate data in this way. These applications may favour Imagers i.e. those who tend to represent information graphically or using images. No cognizance is taken of those who might prefer to represent information in other ways. This is a weakness in the advocacy of the direct manipulation interface. Could enabling audio input provide better support for direct manipulation of the object of interest? This interaction style might suit Verbalisers who are not catered for in current direct manipulation interfaces.

It is worth noting that the psychometric tool (Quis (2006)), which Shneiderman designed (albeit some time ago) to assess software usability and user satisfaction, does not categorise on the basis of preferred mode of information representation i.e. it does not query the users as to their own preferences.

An example of a direct manipulation interfaces would be programmes like AutoCad (Autocad (2006)) where the user can manipulate the drawing in multidimensional formats. This package also provides the user with opportunities to amend drawings without too much difficulty. Using such packages encourages exploration and fosters creativity and in this specific instance provides instantaneous feedback. Another application based on the direct manipulation interface is the Data Desk statistical package (Data Desk V6, 1980-2006).

The Data Desk package, for example, is much more visual and more directly manipulable than SPSS, which does not facilitate direct manipulation of the data. The SPSS interaction style is based on menu selection. This type of interaction limits exploration of the data sets, and may lead to information overload, if the calculations being performed are complicated and require the use of numerous menus.

2.3.2 Adaptive Systems

While Adaptable Systems provide the user with the opportunity to adjust part of the interface; users of Adaptive Systems have no

choice at all. In Adaptive Interfaces the system is responsible for any modification. Wickens (1992) defines adaptive systems as "those in which some characteristic of the system changes or adapts usually in response to measured or inferred characteristics of the human user".

Langley (1998) defines an adaptive interface as being "a software artefact that improves its ability to interact with a user by constructing a user model based on partial experience with that user".

Others have defined an adaptive interface as one that: "Autonomously adapts its displays and available actions to current goals and abilities of the user by monitoring user status, the system task, and the current situation" (Rothrock, Koubeck, Fuchs, Haas & Salvendy (2002)).

Adaptive Systems are generally dependent on software programs known as "agents". Jennings and Woolridge (1995) define a software agent as being "a computer system situated in some environment ... that is capable of autonomous action in this environment in order to meet its design objectives".

Jennings and Woolridge (1995) discuss domains where agents have been deployed: industrial applications (air traffic control, manufacturing, process control), commercial applications (information management), eCommerce, medical applications and entertainment including computer games.

Not everybody is enthusiastic about agents. Some contrary views have been expressed by Sheiderman (1995) and Lanier (1996).

Autoformat within the MS Word package is an example of an adaptive (system-controlled) system. It well illustrates the advantages and disadvantages of adaptive systems. Some users (perhaps a majority) find this feature disconcerting, as it is difficult to reverse decisions made by the system. Sometimes the system assumes that the user requires a particular change made to the text when in actual fact this may not be the case at all. For instance the inclusion of bullet points or font changes to the entire preceding paragraph. What these users find intensely

irritating is the complete absence of a button to toggle Autoformatting off completely. This exemplifies the lack of control users find when using adaptive systems generally and also perhaps illustrates the attitude to control of the parent company of this product.

Shneiderman and Maes (1997) distinguish software agents from other types of software in the following manner: "A software agent knows the individuals user's habits, preferences and interests "and "a software agent is proactive" (Maes (1997)). Maes argues that the user can delegate tasks to software agents thereby reducing information overload: "We need to be able to delegate to what could be thought of as extra eyes or extra ears".

Maes contends that the "World Wide Web and browser is becoming the one and only interface" so therefore agents are beneficial in this type of open unstructured environment.

Parasuraman, Mouloua and Hillburn (1998) argue that there are three advantages to the implementation of Adaptive Systems – regulated workload, better performance and less dependency on static automation.

In the Adaptive System, the system learns something from the user and modifies the interface accordingly. These types of interfaces have been termed user-adaptive systems (Jameson, 2000). An adaptable system can be defined as one in which "the user can explicitly tailor to his/her own preferences (e.g. by choosing options that determine the appearance of the user interface). In contrast, an adaptive system can be defined as being "an interactive system that adapts its behaviour to individual users on the basis of processes of user model acquisition and application that involves some form of learning, inference or decision making" (Jameson (2000)). Keeble and Macredie (2000) define the adaptive interface as follows:
"one where the appearance, function or content of the interface (or underlying application) adjusts itself in response to the user's interaction with it".
Jameson (2000) lists the possible uses of adaptive systems and illustrates them by reference to existing applications. These applications support: system use, adapting the interface, giving advice about system use and controlling user dialogue.

Adaptive systems can also support information acquisition: helping users find out information using agents; browsing; searching; support for query based searching; recommendation of products; provision of information; collaboration and learning. Jameson (2000) in this review provides examples of different applications that serve to illustrate system use and information acquisition such as:

- Elm–Art: a system that supports students learning a programming language called Lisp,
- Phelps: a system that provides virtual peer support to inexperienced workers, and
- FindMe: a product recommender which provides advice about restaurants in a given city.

Findlater and McGrenere (2004) compared the efficiency of different approaches to menu design. This study investigated static, adaptable and adaptive menu design. Their results can be summarised as follows: the adaptable menu was preferred to the adaptive menu, and the adaptive menu was preferred to the static menu. Their findings also report, that "the adaptive menu was slower than the static menu. The adaptive menu was slower than the adaptable menu, except when subjects used the adaptable menu first". Or put more simply, the adaptive menu was slower than the adaptable menu, which was slower than the static menu. The sample size was small in their study (n = 27) but the study has interesting implications. The students were paid to participate. A potential weakness of this study is the subjective rating of the participants' own levels of their own computing expertise. The students were university psychology students so this sample may have been accustomed to these types of experiments. A more naïve, randomised sample type might have been more useful.

The users' perception of their own efficiency in using the menus and actual performance differed quite significantly. The authors also report, that "the ability of the users to customise their own menus to achieve a result that was not found to be significantly different from the optimal efficiency of the static split menu is a strong result". In other words, the users' customisation was not particularly advantageous to them compared to what would have been achieved using a fixed (non-adaptable) interface in less time. That most users want a personalised interface

however was another finding from the study – but the question still remains what is the best way to provide this? The authors suggest that combining adaptable and adaptive "in a mixed initiative design may be the best way to satisfy a wide range of users". Another question about this study is how generalizable are these findings to other populations and applications?

2.3.3 Models of Adaptive Systems

There have been some adaptive system theoretical models developed to guide development in this area. Examples of some of these models are Wickens' (1992) Closed Loop Adaptive Systems model (CLAS), Benyon and Murray's model (1993) and Jameson's users' adaptive system model (Virvou (1999), Jameson (2000)).

Egan and Gomez (1985) have examined the facilitation of individual differences. They have presented a three-stage model. They advocate assessing the individual difference and then separating it from confounding factors. They suggest that once this has been accomplished, it is then necessary to accommodate it in the design.

Their approach may be used in the following manner, using cognitive style as the example. There is more than one cognitive style; which one should be chosen and why? This requires reliable, valid, cost effective, and simple to use instrumentation. Perhaps a confounding factor may be intelligence, so it therefore becomes necessary to control for this. Finally, a decision has to be made, as to how this cognitive style is going to be accommodated within the system design.

Wicken's Closed Loop Adaptive System model CLAS (1992) comprises five main components: the environment, the automated performer, the CLAS manager containing the decision rules, the human operator and the task. The change to system automation is dictated by a set of decision rules located within the CLAS manager. This CLAS manager receives inputs from the environment, the human performer and the user model – the automated performer.

The term adaptive interface is approximately synonymous with the term Intelligent User Interface. Hook (1999) for instance uses this term, as do Benyon and Murray (1993).

Benyon and Murray (1993) present a more sophisticated model of the adaptive system, calling it an architecture for intelligent systems. It comprises a user model, a domain model and a third component, called an interaction model. Both user and domain model provide and receive inputs and outputs to the interaction model. The user model provides a profile of the user and this may include such data as user ability, user experience (novice, expert), psychological data (personality characteristics, cognitive style, and general cognitive abilities).

The domain model defines the scope of the system. This model consists of three other sub-modules. These levels are: intentional, conceptual and physical.

The main component of this model consists of the interaction model. Again it comprises four sub-components: a dialogue record, an evaluation mechanism, an adaptation mechanism and an inference mechanism (Benyon and Murray (1993)).

Jameson's model (2000) models the user and then relies on this user-generated model to provide a basis for adaptation.

Virvou's (1999) model consists of four parts: the Domain Knowledge, the Advice Generation component, the User Modelling component and lastly the User Interface.

In these systems, there are a few principal components and issues: the user model, user model acquisition, information about user, user model application, predictions and decisions about the user (Jameson, 2000 p306).

Rothrock, Koubek, Fuchs, Haas and Salvendy (2002) (Appendix 1-2.1.A) summarized these models in their review of the area and then presented their own model. Their model encompasses some characteristics of the previously conceptualised models already mentioned. It includes the following list of common features: inputs, user variables, identification inference mechanism, interaction model, decision inference mechanism,

systems or environmental variables. It could be thought of as being a generic model encapsulating some of the features of the specific models already described. It primarily provides a universal model of adaptive interfaces and is useful for those researchers wishing to design systems based on this paradigm.

This model is based on previous models of adaptive systems and may be regarded as a generic model comprising a comprehensive set of possible features of any adaptive system. This model categorises the main types of input variables which can be adapted, including: user performance, user goals, user workload, user situational awareness, user knowledge, groups of users, user personality and cognitive style, situational awareness, task variables. Some of these variables are static while others are dynamic.

Changes in any of these variables result in adaptation of the interface. Broadly the authors argue that adaptation consists of two main types – content and dialogue.

Rothrock, Koubeck, Fuchs, Haas, Salvendy (2002) classify interface adaptation into four categories: content adaptation, human system dialogue, task allocation and real time adaptation.

Brusilovsky (1996) and De Bra, Brusilovsky and Houben (1999) have studied dialogue adaptation in web-based application. This type of dialogue adaptation involves reducing the number of steps to complete a task. Ultimately, the amount of time spent on a task is reduced, while the overall aim is accuracy and therefore improved performance.

Another example is dialogue annotation of previously used links in a hypermedia program or through use of the Web. Other methods include placing frequently used applications or menu choices at the top of menu lists.

Content adaptation can consist of the selection of information, quantity of information, augmentation of information, layout of information and finally modality of information.

Dialogue adaptation applies mainly to hypermedia /hypertext environments. This involves the adaptation and annotation of links and menus, to suit the user model.

The next component of their model concerns task allocation and level of automation. The important factors considered are: "what is the best level of automation" and "who or what decides to switch on /off the automation?" The pertinent issue is to what extent the task should be completed by the system, and to what extent by the user. Relevant also is "who or what decides when the control of the entire system should become automated or switched to manual control?"

This is an important consideration when adaptive systems are designed where system inputs are dynamic. Examples of these are in domain areas such as aviation and the control of nuclear plants. In relatively early work, Sheridan and Verplank (1978) outline ten levels of varying degrees of automation ranging from the human completing the full task with no computer intervention to the complete opposite with the computer doing the work.

This model does not indicate examples of tasks commensurate with each of the ten levels of automation described in this model. It is probable that task type, context, characteristics of the user could be determining factors in deciding the different appropriate levels of automation in the design of adaptive systems. A relationship between possible types of adaptive system design and varying levels of automation as outlined by these authors would have been instructive and beneficial to those interested in adaptive system design.

The delegation of control to an adaptive system is a contentious issue as it requires user trust. Jennings and Woolridge (1995) comment "for individuals to be comfortable with the idea of delegating tasks to agents they must first trust them" and they further state that "the agent must strike a balance between continually seeking guidance (and needlessly distracting the user) and never seeking guidance".

As with any model of an adaptive system there should be a consideration of the inference mechanism. This essentially

classifies the input variables and determines the effect of these in the context of system change.

In other early work, Rosen (1985) in Rothrock, Koubek, Fuchs, Haas, Salvendy (2002) defines an anticipatory agent "as a system containing a predictive model of itself and its environment, to allow it to change its state at an instant in accord with the model's prediction pertaining to a latter instant". In advanced systems this component as discussed by these authors, has also the capacity to evaluate the effect of the proposed change on the system before implementation. Fundamentally in these instances, the system displays a capacity for self-evaluation. Underlying computer systems involved in the implementation of this may include neural networks, fuzzy logic systems, decision trees and Bayesian networks. This model provides designers of adaptive systems with a plethora of possibilities. However the efficacy of adaptive systems lies in their ability to perform in real time environments.

Rosen's (1985) comments that the requirements of the adaptive interface must be "sensitive along all three dimensions of abilities, tasks, and situations in order to be effectively adapted to the user in real time applications". Rothrock, Koubeck, Fuchs, Haas, Salvendy (2002) in their review of adaptive interfaces state that "no application was sensitive to changes along all three". They argue that successful adaptive interfaces must adjust to complex situations while maintaining adequate communications with the user. Perhaps at this point in time their expectations of adaptive systems are utopian although their model is stimulating and quite complete.

Modelling and predicting human behaviour remain the most enduring difficulties in the area of adaptive system design.

The problems of user control, system transparency and predictability have been cited as some of the disadvantages of adaptive systems. Hook (2000) argues that by virtue of designing adaptive systems some of these usability principles designed for direct manipulation systems are debased. In attempting to address these issues she designed an application (PUSH) that provides different levels of transparency to the user of the system. This important system is now discussed.

2.3.4 Evaluation of the Push System.

Hook (1997) discusses the evaluation of the PUSH system already referred to above. The PUSH system was designed to provide users with help, adaptively, to find information in an online manual. Hook states "that the PUSH system decides what to show and what to hide in a page ... thus attempting to avoid information overload". The goals of the PUSH system were: "to reduce the information overflow problem" and "to aid users in getting at the most relevant information with respect to their information seeking task". The implemented adaptation specifically affects how information was presented on the page, not how the users would have had navigated through the pages.

Hook (1998b) applied the "metaphor of a black box in a glass box" in the design of the PUSH system. Hook quotes "in the PUSH system we hide the complex inferencing of user goals (in the black box) and show a quite simplified view on what is going (in the glass box) to the user". Essentially the issue becomes what should be transparent and what should be opaque to the user in the design of adaptive systems. User characteristics will impinge on how this type of decision will be made, for instance user's goals, knowledge, cognitive abilities, personality, task and purpose of interaction are important variables relevant to this type of decision.

The application was compared with a non-adaptive version and the same individuals used both systems. The sample size was very small consisting of three females and six males. These participants were all computer literate and had experience of using the Web.

The participants were assessed as to how successful they were at locating the most relevant information when presented with various tasks. The author reported that adaptivity "reduced the amount of actions needed", and that "subjects clearly preferred the adaptive system". The conclusion that could be drawn is that the adaptive system reduced the information overload when compared with the non-adaptive system. An interesting issue is that one might have expected perhaps a significant difference between task time completions for each individual when the two systems were compared; this was not really explored to any great depth in her paper. The results would suggest that this

adaptive system is preferable when compared with its equivalent non-adaptive system.

This study was interesting in that it sought to compare an adaptive system with a non-adaptive system using many of the same tasks. The experimental design sought to compare the performance of the experimental sample using both systems. However this study is limited given its small sample size and more than likely is not powerful enough to test many of the hypotheses presented in the paper. The author presents bar charts and draws many of her conclusions based on these, which, while being informative, could not be taken to be persuasive because it lacks statistical validity. What the author could have done was to isolate one variable and investigate the participants' performance on both systems, and establish differences using a paired t-test.

The results reported in this study have to be interpreted in light of these limitations. The participants were also videotaped and this may have affected their performance. Would they have performed similarly in a real working environment?

2.4 Evaluation of Adaptive Systems

Weibelzahi (2005) asserts: "Empirical evaluations help to estimate the effectiveness, the efficiency and the usability of the system" and it provides opportunities to test the assumptions made about the user's model. These assumptions include users' skills, preferences, needs, cognitive profile, previous knowledge and expected style of interaction with the system. Weibelzahi (2005) also comments: "publications on user modelling systems and adaptive hypermedia rarely contain empirical studies" and states that researchers "have been lamenting about this frequently". He also argues that evaluation methodologies are often neglected in the design of computer curricula at third level and that they should be included as part of an undergraduate syllabi.

Hook has also referred to the difficulty in assessing adaptive systems and this is also commented upon by Weibelzahi (2005): "the evaluation of adaptive systems includes some inherent systems and pitfalls that can easily corrupt the quality of the

results and make further conclusions impossible". These problems are summarized in Table 2.1 below.

Table 2.1 Problems in assessing adaptive systems
(Adapted from Weibeizahi (2005))

Section	Problem	Recommendation
Formative /Summative Evaluations	Occurs at end of Project - often too late to recover failures	Conduct several formative studies across the development cycle of the project
Allocation of resources	Empirical studies require - personnel, organisational and financial resources.	Use expert evaluation Evaluate interpretative models with simulated users and empirical data
Specification of Control conditions	Switching off the adaptivity might result in an incomplete or even useless system	Compare various adaptation decisions conditions that are based on the same user characteristics
Sampling	High variance corrupts statistical analysis	Use repeated measures designs Samples - heterogenous on user modelled characteristics, homogenous on other factors Control for variables that might have an impact on results
Definition of criteria	Difficult to establish a single evaluation criterion	Definition of goals of user adaptivity and use these to derive other criteria
Asking for adaptivity effects	Users might be unable to report adaptivity effects (novice users for example)	Use user feedback in combination with associated measures
Reporting the results	Incomplete reporting of the results corrupts and makes the interpretation of the study difficult	Report on effect size, empirically identified user characteristics, following guidelines in reporting statistical data

Weibeizahi believes that the evaluation of an adaptive framework could comprise the following phases and should be evaluated independently of each other "collection of input data, interpretation of the data, modelling the current state of the world, deciding on the type of adaptation, and applying adaptation". He argues "that groups or even single users might be observed across these layers thus the focus of evaluation using the approach has moved from evaluating the whole sample group at one layer as opposed to a subset of the sample across the different layers. Others believe that the evaluation of real usage data using visualisation techniques "are a useful part of good adaptive multimedia strategies and help form user model attributes" (Sobol and Stones (2005)).

Weibelzahi's (2005) perspective and general recommendations towards the evaluation of adaptive systems provide guidelines as to how this might be achieved effectively in the future. What are evident from this analysis are the difficulties in evaluating adaptive systems in a productive and informative manner. Consequently there is a difficulty in establishing a rich empirical basis to guide the design of such adaptive system applications in the future.

2.5 Problems of Adaptive Systems Design

There are many limitations of adaptive systems including: Usability, User Modelling, Predictability and Transparency, Controllability, Unobtrusiveness, Privacy and Breadth of Experience. Some of the typical properties of user-adaptive systems can lead to usability problems that may outweigh the benefits of adaptation to the individual user (Jameson (2000). Keeble and Macredie (2000) comment that "adaptation may make the interface less usable, rather than enhancing the interaction by supporting the user in his task ". Fig. 2.1 gives a high-level summary of many of the relevant ideas.

2.5.1 Usability

Carroll (2002) comments on the migration of the computer as a tool used by highly skilled and trained personnel to society at large . Today people use computers in their offices and homes in many cases not being trained in their use at all. Carroll states "the usability needs of these new computer users propelled the growth of HCI as a discipline of study and focused its primary goal on making computers easier to use". He also said "Basic usability ... involves an assortment of support for needs such as ease of use, ease of learning, error protection, graceful error recovery and efficiency of performance". These needs, as outlined by Carroll, may seen as being broadly espoused as usability goals by some other researchers e.g. Jameson (2000).

Fig. 2.1 (Jameson (2000)) illustrates many of the main issues connected with the usability of Adaptive Systems.

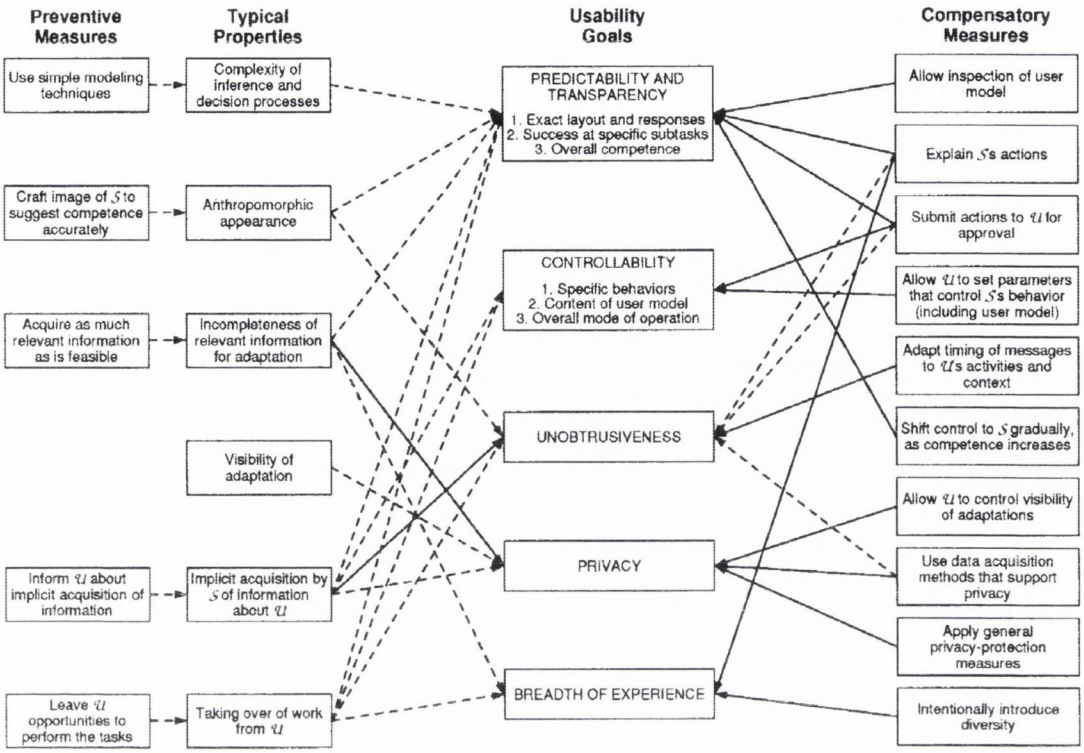


Fig. 2.1 Overview of usability challenges for user-adaptive systems

2.5.2 User Modelling

Some parameters of the user models can be difficult to predefine – for example, user goals, and some of those based on cognitive style and personality. How can one specify user goals in advance? User's goals may change during the course of the interaction with the system. Brusilovsky (1996) states, that "the users' goal is the most changeable user feature".

Hook (2000) refers to the problem of unreliable user characteristics, the problem of diagnosing user characteristics some of which should not be used as a basis for adaptation. She states: "A proper analysis of users and their characteristics will reveal many individual differences that we should not necessarily use as the basis for adaptive behaviour in the system" e.g. gender friendliness, and IQ. It can be difficult to accurately measure cognitive style and personality variables. Much depends on the reliability of the instruments used in this regard.

User performance presents difficulty in that how decision-making occurs can be difficult to predetermine, encapsulate,

assess and predict. User performance can be monitored by means of reaction time, number of keystrokes or number of errors made. A change in these variables can be used as a basis for causing adaptation in the interface.

Adaptive systems depend on how well they can model the human user's activities and upon the efficiency of the algorithmic techniques to do so. User models are always constrained, as it is difficult to design ones which realistically approximate the flexibility and intuitiveness of human nature. Maybe the question might be, if it were possible to design such models, to what extent should they be designed in this manner anyway?

Adaptive systems can be conspicuous; they can interfere with task performance by providing distracting information that may hinder task efficiency and therefore put humans at risk depending on the task.

It is questionable whether the design of the user model actually serves the users of adaptive systems. Fischer (2001) states that "past research has shown that there is often quite a difference between modelling certain aspects of a user's work and behaviour, and applying this knowledge for the benefit of the user".

The formulation of a user model and knowledge domain model are problematic also. The concern here is how accurately can this be done? How efficiently can knowledge be elicited from the user to build up a realistic user model or indeed a model of a group of users? How applicable can the application be to other areas?

2.5.3 Predictability & Transparency

The concept of predictability refers to the extent to which the user can forecast the effects of his/her actions. Transparency is the extent to which he/she can understand system actions and/or has a clear picture of how the system works.

In adaptive models unexpected change can hinder users' performance: "If adaptive systems make surprising changes users must pause to see what has happened. Then users may

become anxious, because they may not be able to preict the next change, interpret what has happened, or restore the system to the previous state" (Shneiderman (2004a), Shneiderman (2004b)). MicroSoft Word has many examples of this type of anxiety generating response.

It is questionable whether the design of the user model actually serves the users of adaptive systems. Fischer (2001) states that "past research has shown that there is often quite a difference between modelling certain aspects of a user's work and behaviour, and applying this knowledge for the benefit of the user".

Adaptive systems are not always predictable or transparent. They provide little or no feedback to the user and therefore this may be not conducive to user performance or efficiency. Predictability and transparency are important design tenets. Systems ought to convey a sense of predictability in the context of human – system interaction.

2.5.4 Controllability

Controllability refers to the extent the user can bring about or prevent particular actions or states of the system if he/she has the goal of doing so. Adaptation relating to task allocation and level of automation raises serious issues of control. Does the system decide when to switch from human control to system control automatically, or does the system monitor user performance and then decide to intervene?

In adaptable and especially adaptive systems, the safety of humans involved in the use of these systems becomes paramount. Such systems are often styled man-rated if the loss of human life may be a possibility due to system failure. An example might be a computer system used by an operator to control the amount of radiation administered to a cancer patient (Jacky (1997)).

Hook (1999) argues that adaptivity can sometimes cause more problems than solutions. Shneiderman (1997), Keeble and Macredie (2000), Hook (1999) argue that users of adaptive systems feel a loss of control and that the functionality of the

system through the user interface is not always transparent or readily comprehensible.

Shneiderman (2004a) argued that a critical issue is one of the reversibility of decisions.

2.5.5 Unobtrusiveness

Obtrusiveness refers to the extent to which the system places demands on the user's attention, which reduce the user's ability to concentrate on his/her primary tasks.

Adaptive systems can be overly conspicuous; they can interfere with task performance by providing distracting information that may hinder task efficiency and therefore put humans at risk depending on the task.

In adaptive models unexpected change can hinder users' performance: "if adaptive systems make surprising changes users must pause to see what has happened".... "users may become anxious" (Shneiderman (2004b)). Designers of adaptive systems according to Shneiderman (1997) "avoid discussing responsibility"; their designs rarely allow for monitoring the agents' performance, and feedback to the current user ... is often given little attention".

2.5.6 Privacy

Users are concerned with the extent to which the data gathered by a system about them can be put to inappropriate use that can cause them harm to a greater or lesser extent.

In cases where intelligent agents assimilate information about the user the security of this personal information can become problematic. Issues such as: security of confidential information and its digital storage become germane; types of precautions that can be taken to prevent the possible theft of this information are relevant. How secure is the information on the system? What measures have been implemented to insure against theft?

2.5.7 Breadth of Experience

Breadth of experience is concerned with the narrowing of experience due to handing over control to the system.

If the intelligent interfaces are treated like another human being by means of anthropomorphic agents, the issue of responsibility for poor decision making becomes more blurred. Shneiderman (1997) argues that these types of agents give the user the impression that the system acts like a rational being and thus will take responsibility for its actions e.g. Hal in the film "2001: A Space Odyssey" serves to illustrate this point. In cases where poor decisions are made who is legally culpable in these instances?

Adaptive system design may be problematic in terms of breadth of experience. It can present a narrow view of the domain. Adaptive systems may not facilitate a holistic view of the subject domain given the rigidity of algorithmic search engines.

The use of predictive text, characteristic of many of today's mobile phones, illustrates this point also. In some cases, the algorithm predicts the incorrect word and it can be quite difficult to correct it and prevent it from happening again. "Some designers continue to be seduced by anthropomorphic scenarios even though they have been repeatedly rejected by consumers" (Shneiderman (1995)).

2.6 Progress to Date in Adaptive Systems

Shneiderman who is opposed to agent-based scenarios states that there is "much discussion of user models but little empirical evidence of their efficacy". Furthermore few adaptive systems have been successful and generalizing from these systems has "proven much more difficult than advocates have hoped" (Shneiderman (2003)).

Shneiderman (2004a) argued that a critical issue is one of reversibility of decisions. Shneiderman (2004) states, "that anthropomorphic representations of computers have been unsuccessful in bank terminals, computer assisted instruction, talking cars, or postal service stations".

Hook (1998a) refers to some of the problems pertaining to direct manipulation design. These include information overflow and cognitive overload. Others have referred to the problem of user exclusivity (Benyon, McCrerar and Wilkinson (2001)).

Shneiderman favours direct manipulation system design as opposed to user-adaptable or user-adaptive design.

However Jameson's review clearly shows that adaptive systems do not provide the panacea as expected by the advocates of this approach. Jameson (2000) also reports that "adaptive interfaces, agents ... do not represent a smooth and easy road" and "it is often tricky to prove empirically that user adaptivity has actually added any value". He further adds, "it is no wonder that some computer experts believe that the interests of computer users are better served by continued progress within more familiar paradigms of user – centred system design".

Some researchers argue that here is a need for hybrid systems to implement the strengths of both adaptable and adaptive system designs (Bunt, Conati, McGrenere (2004)).

2.7 Conclusions

This Chapter reviewed the various types of system (Non-Adaptive, Adaptable and Adaptive) which could be considered for accommodating individual differences such as cognitive styles. Given the difficulties pertaining to Adaptable and Adaptive Systems, it appears that Non-Adaptive Systems are to be preferred at this point of time.

Hook (2000) reports "that the very few intelligent user interfaces that have succeeded commercially" have been implemented upon "very simple adaptations based on simple knowledge of the user", or have "created its adaptations based on what other users do" as opposed to a "complex inferred model of user ".

Solutions to the problems of adaptive systems may include a hybrid approach as advocated by Bunt, Conati and McGrenere (2004). They report based on their research that "adaptive support could be added to an adaptable interface to improve the effectiveness of the users' customisation".

Future developments in the area of adaptive design may include the use of genetic algorithms, paradigms that model swarm intelligence and other biologically inspired design models (Rothrock et al. (2002). Jameson (2000) points out that there is a need for such models as described above due to the following

factors: the diversity of users and their contexts, the scope of information to be dealt with, and the complexity of interactive systems. Work on the brain – computer interface may also prove fruitful in this regard.

Some researchers are not convinced about the benefits of adaptive systems since they infringe upon usability principles. Jameson (2000) highlights these, stating that adaptive systems "can lead to usability goals that may outweigh the benefits of adaptation to the user". Keeble and Macredie (2000) comment that "adaptation may make the interface less usable, rather than enhancing the interaction by supporting the user in his task ".

Hook lists the impediments to the design of good intelligent interfaces as: "a need for methods for how to develop them ... demands on better usability principles for them .... a better understanding of the possible ways the interface can utilise intelligence to improve the interaction and finally ... the need to design better tools that will enable an intelligent system to survive the life cycle of a system" (Hook (2000)).

Rothrock, Koubeck, Fuchs, Haas, Salvendy (2002) reviewed eighteen adaptive systems (Vid. Table 2 Appendix 2-2.2.B) and concluded that, "we have yet to find an interface or interface design methodology that provides real time adaptability."

The numerous difficulties in adaptable and adaptive systems are apparent in the context of usability goals and user models (Jameson (2000)). Jameson's review indicates that there has been little or no success with real time adaptive systems. It can be difficult to establish whether adaptive systems actually perform better than non-adaptive / adaptable systems; there is a dearth of literature in this regard.

The human computer interface affects how people use computers efficiently, effectively and creatively. To conclude, this Chapter has examined the adaptability of the three main types of human computer interface. The next chapter gives an overview of the field of cognitive styles, an important factor in the human use of computers.

Chapter 3:
A Review of the Cognitive Styles Domain

This chapter reviews the cognitive style domain. Cognitive Styles are a type of Individual Difference that affects how humans use computers in an efficient, effective and creative way as postulated.

This review considers definitions and some models of cognitive style. The chapter also discusses some of the most popular cognitive styles and evaluates them. A consideration of some definitions of style and a brief review of the major style labels will be presented. Also examined will be some of the proposed models in the field, which have sought to unify style, and their contribution to the field of style research will be assessed

The chapter will conclude by highlighting what may be considered the main impediments to style research.

3.1 Distinctions between Cognitive, Learning and Thinking Styles

There is a multiplicity of style definitions. Terms, like cognitive style and learning style, have been used interchangeably by some (Entwhistle (1981)). Other researchers like Das (1988) see learning and cognitive styles as being different and have attempted to define them as being so (Riding and Cheema (1991)).

Riding and Cheema (1991) distinguish between cognitive style and learning style in the following way: "cognitive style is a bipolar dimension, learning style entails many elements and are usually not either - or extremes". "One either has or does not have the element in one's style; similarly, the absence of one element does not necessarily imply the presence of the opposite element".

Head (1996) makes no distinction between the terms cognitive style, thinking style and conceptual style. On the other hand Sternberg (1997) distinguishes between thinking style, cognitive style and teaching style.

Kirby (1979) in Hashway and Duke (1992) believes that the term "learning style" was used by researchers when they started to "look for specific strategies for combining course presentation and materials to match the particular needs of each learner". Hashway and Duke (1992) also maintain that learning style is "a broader term that includes cognitive style "although they suggest that the more inclusive the term is, the more it "loses some of the specificity and precision that definitions of cognitive style possessed". Sadler-Smith (2001) maintains that cognitive style and learning style should be treated as separate constructs.

Others have used one style label (Witkin's Fd/Fi, Witkin et al. (1971)) to refer to cognitive style due to the enormous amount of research done using this construct (Adey, Fairbrother and Williams (1996)).

The field of style research has also been plagued by difficulties in accurate measurement of these style dimensions. "What is a style?" and "how can it be accurately measured?" remain as the most persistent problematic issues, that dominate style research. Style is an attractive construct to research in the field of individual differences since it is claimed by a majority of researchers to be independent of intelligence (Kirton (2003), Sternberg (1997), Riding & Rayner (2000)).

Sternberg & Grigorenko (1997) believe that style serves as the bridge between cognition and personality; others believe that it is the missing link between cognition and personality e.g. Riding and Rayner (1998).

The earliest applications of style research concentrated on educational contexts. More recently researchers have examined occupational settings and team selection. Some studies on the application of cognitive styles to the design and study of performance in eLearning environments are now starting to emerge. Many studies in these areas are based on the matching style hypothesis. This assumes that if an individual's style is matched to particular aspects of the environment then that individual is more likely to succeed at a particular task. This theory has been studied in the context of human relationships (Witkin (1976) and Armstrong (1999)).

3.2 The Importance of Cognitive Style

Cognitive style research has been thought to offer much to educational settings. Witkin (1976) states that "it is already clear that cognitive style is a potent variable affecting a number of areas: the students' academic choices and vocational preferences, the student's continuing academic development, how students learn and teachers teach, and how teachers interact in the classroom". Witkin (1976)) provides references to research carried out by others, which support this view. One can conclude that cognitive styles had, and still have, important implications for educational theory and pedagogy. Since then Riding and his co-researchers have provided further research in educational settings (Vid. Chapter 7).

During the last decade or so there has been interest in the application of cognitive style to organizational behaviour (Hayes and Allinson (1994), Allinson and Hayes (1996), Kirton and McCarthy (1998), Sadler-Smith and Badger (1998), Armstrong (1999)). Kilmann and Mitroff (1976) reported, for example, that people with different cognitive styles preferred different types of organizations. Some researchers have examined cognitive style in the context of team selection e.g. Kirton (2003). It is crucial that large organizations appreciate cognitive style with the ever-increasing importance attached to teamwork and the need for their workforce to function effectively in teams.

For example, Kirton, Bailey and Glendenning (1991), in a study of 182 British schoolteachers, found that Innovators had the preference for procedures loose in structure where aims and assessment were poorly defined. In contrast, "Adaptors showed a preference for procedures containing a tighter structure, more definable aims and more precise methods of assessment." With respect to the contentious issue of streaming in academic and social contexts, there were significant differences between Adaptors and Innovators. For instance, in academic contexts, Adaptors were more in favour of streaming than Innovators ($p <= 0.01$). In social contexts, Innovators were more opposed to streaming than Adaptors ($p <= 0.05$). Kirton concludes: "There is no best organisation, adaptive or innovative, and no single effective mix of Adaptors and Innovators within it; the positive qualities of both are needed if an organisation is to attain its goals".

Interest in cognitive style research has spread to the design of multimedia educational products and eLearning environments and the accommodation of cognitive style in the design of these flexible systems (Ford and Chen (2001), Parkinson and Redmond (2002a, 2005)).

3.3 Definitions of Cognitive Style

Allport (1937), one of the first to define cognitive style, comments that, "a cognitive style is an individual's typical or habitual mode of problem solving thinking, perceiving and remembering". Allport (1937) was probably the first to be credited with using the style construct in the context of cognition (Grigorenko and Sternberg (1995)).

Coop and Sigel (1977) define cognitive style as: "to denote consistencies in individual modes of functioning in a variety of behavioral situations".

There are numerous definitions of cognitive style. Goldstein and Blackman (1978) define cognitive style as being a "hypothetical construct that has been developed to explain the process of mediation between stimuli and responses". They would argue that cognitive style refers to "the characteristic ways in which individuals conceptually organize the environment". They refer to Harvey's (1963) view of cognitive style as being the "way an individual filters and processes stimuli". Messick (1976) also defined cognitive style as "consistent patterns of organizing and processing information". He also understands cognitive styles to be constant and dependable.

Guilford (1980) takes another view "it is generally agreed that what writers call cognitive styles are in the general family of personality traits, where traits are commonly conceived as variables or dimensions along which individuals of a population differ".

Hunt, Krzystofiak, Meindl and Yousry (1989), in Van den Broeck, Vanderheyden and Cools (2003), defined cognitive style in terms of how people process and organize information and arrive at judgments or conclusions based on their observations. Kirton (1989) also assumes that cognitive style is independent of "cognitive capacity, success, cognitive techniques (strategies)

and coping behaviour (functioning temporarily outside ones habitual style".

Tennant (1988) defines cognitive style as being "an individual's characteristic and consistent approach to organizing and processing information and experience". Sternberg & Zhang (2001) says that term cognitive style "refers to an individual's way of processing information".

Riding and Cheema (1991) describe how cognitive /learning style has been perceived, "cognitive/learning style have been viewed in three main ways, as a structure (content), as a process or as both. These authors then state "if cognitive style is viewed as a structure, then the focus is on stability over time, if cognitive style is viewed as a process then the focus is on how it changes" and finally "for others cognitive style is viewed as both a process and structure". The authors explain that in this instance that cognitive style "may be relatively stable, not changeable like a liquid with no form of its own yet at the same time always in a state of flux". They conclude by saying that "in such a view, style structure is continually modified as new events influence it directly or indirectly".

The definitions of style referred to above support Riding and Cheema's (1991) view that perhaps the definition of style may depend on one's perception of it. Such discrepancies in definition of cognitive style only add to the bewilderment amongst those who seek to define this intangible construct in absolute terms.

Leonard and Strauss (1997) support this view and believe that cognitive style can be affected by one's life experiences and thus are not inflexible.

Sternberg's definition of style is "a preferred way of thinking" (Sternberg 1997)). Sternberg outlined his own views on style. What follows is a summary of some of these views. Sternberg says that: "Styles are preferences in the use of abilities, not abilities themselves "and that "people have profiles (or patterns) of styles, not just a single style". Essentially Sternberg is arguing that, individuals cannot be labeled by a single test score, as is the case, for instance, with the GEFT. He also argues that, "people

differ in stylistic flexibility" and also "in the strength of their preferences"

He also believes that "style can vary across the life span "of an individual. In essence what is being questioned is the temporal stability of style that contrasts with Riding's view of it as being a fixed construct. Is it possible that as an infant develops that there is a transition from field dependency gradually to field independence, reversing to field dependency as ones natural frailties begin to become apparent?

Sternberg (1997) believes that that "styles are teachable "and that people"acquire their styles through socialization". The point of view being taken here is that thinking style is malleable rather than being inflexible. He also believes that gender, culture, age, parenting styles type of education received and occupation can affect the development of thinking styles.

Riding and Rayner's (1998) definition is: "Cognitive style is seen as an individual's preferred and habitual approach to organizing and representing information".

Riding and Rayner (1998) further categorizes style as being independent of intelligence, gender and personality. Cognitive style is structural; according to Riding, it has temporal stability. This construct therefore is stable and does not change over time. However there are others who have different views.

Riding and Rayner (1998) further state that cognitive style should "be related to observable behaviors such as: behaviour, learning performance, learning preferences, subject preferences, motor skills, social behaviour problems, stress and aspect of cognition and physiological measures.

Van den Broeck, Vanderheyden and Cools (2003) in an interesting review, give some definitions of cognitive styles and in light of these definitions propose their own. Cognitive style is defined "as the way an individual perceives environmental stimuli, and organizes and use information".
Kirton (2003) believes that cognitive style is bipolar, non-pejorative and non-evaluative. He assumes that cognitive style is

related to numerous personality traits that manifest themselves early in life and are temporally stable.

3.4 Cognitive Style and Ability

Cognitive style is bipolar and non-pejorative whereas ability is quite the opposite. Abilities are value directional, the more the better it is for the individual. Effectively performance on a task improves with style and ability. If however one considers style alone the effect will be positive or negative depending on the task in question (Riding (2000b)). Head (1996) states, that "measurements of cognitive style should be orthogonal to those of cognitive ability".

Tiedemann (1989) describes how Messick distinguished between abilities and styles as in Table 3.1.

Table 3.1 Distinguishing between Ability and Style (Adapted from Tiedemann (1989) and Messick (1984)).

Ability	Style
Content /level	Manner /form
Competencies	Propensities
Maximal	Typical
Unipolar	Bipolar
Value directional	Value differentiated
Domain /functional- specific	Pervasive/organising
Enabling	Controlling

Riding and Rayner (1998) comment on McKenna's (1984) delineation between ability and style as follows

"Ability is more concerned with level of performance while style focuses on the manner of performance"

"Ability is unipolar while style is bipolar"

"Ability has values attached to it ... while for a style dimension neither end is better overall "

"Ability has a narrower range of application than style".

These two sets of distinctions are broadly similar and clearly distinguish between style and ability. Riding & Rayner (1998) believe that cognitive style may be more important for individuals of lower ability than for those of higher ability Despite there being conflicting views about the nature of the relationship between personality and cognitive style, there

seems to be wide scale agreement on the distinction between ability on one hand and cognitive style on the other. It has been argued by Riding and Rayner (1998), that when both style and task characteristics match, performance is catalyzed for the individual.

3.5 Cognitive Style, Learning Strategy and Learning Skill

The distinction between style, strategy and learning skill is necessary since these terms are often confused. These terms /concepts have been clearly delineated by the following researchers:

Hayes and Allinson (1998) comment "while styles may produce consistent behaviour across a variety of situations over the short and medium term, strategies are much more specific and essentially represent the result of the conscious decisions an individual makes to cope with immediate cognitive tasks".

Adey, Fairbrother, Wiliams, Johnson and Jones (1999) use the term learning strategies and maintain that they can be changed by "development and /or teaching". They maintain this is the crucial difference between style and strategy". Learning strategies are "groups of skills that a learner uses together for a particular purpose".

They believe that there are different types of strategies: some which are involved in information processing and "other strategies help manage time, motivation, and emotions". They also argue that some of these strategies are "performed mentally and some are external tactics such as note taking". They also argue that learning strategies have certain characteristics such as:

- "They are deliberate actions performed to attain particular goals
- They are invented or generated by the person and involve agency and control rather than compliance and mindless rule following
- They are selectively and flexibly applied and involve cognitive skill and motivational will
- They are often socially assisted tactics for problem solving that become independent, especially when related to academic learning tasks
- Although strategies are important trouble-shooting tactics and are often consciously applied or shared, their development

should lead to them being applied automatically and to being transferable".

Riding and Rayner (1998) comment that "Style probably has a physiological basis and is fairly fixed for the individual" in contrast with strategies which "are ways that may be learnt and developed to cope with situations and tasks and particularly methods of utilizing styles to make the best of situations for which style is not ideally suited".

Riding and Rayner (1998) also state that style is "probably present at birth or at any rate fixed early on in life and it is thought to be deeply pervasive affecting a wide variety of individual function".

Learning skills can be considered as methods by which learning can be enhanced. An example would be the use of mnemonics. Other examples might be the use of advance organizers presented at the beginning of a chapter of text to tell what's coming next.

Strategies therefore can be learnt and are changeable depending on the situation in contrast with style, which is fixed. Strategies may be thought of as being collections of learning skills. Finally styles can be conceptualized as being programmed while strategies are liable to be changed.

Cognitive styles by definition should be value free; in other words it shouldn't really matter at what point an individual is on the continuum. While in theory this may be fine, in practice being free of value may not always hold. In some cases research indicates that it seems to be better to be field independent than field dependent. One would suspect that it would be better if flying an airplane to be field independent when devoid of a reference to the horizon as opposed to being field dependent.

Also take for instance, the research reviewed by Jonassen & Grabowski (1993) describing studies, which show that field independent students tended to perform better than their field dependents in traditional academic settings. More recently studies by Parkinson & Redmond (2001, 2002a) and Weller, Repman & Rooze (1995), report that the field independent

learner is at an advantage when studying in computer based environments.

Kagan's Impulsivity – Reflectivity cognitive style could be analyzed in much the same manner (Kagan, Rosman, Day, Albert and Philips (1964)). Those who are Reflective make less mistakes and tend to do better when compared with those who are Impulsive.

Zhang & Sternberg (2005) examine three controversial issues in the area of style research: "is style a matter of state or trait?", "is style value laden or value free?" and finally do "style constructs from theoretical models overlap". They present a thorough review of the research and conclude that, "thinking styles largely represent states, not traits". With respect to the second theme they conclude, "that thinking styles are value laden rather than value free.

When testing a new cognitive style, the external convergent validity of the style should be examined. Effectively one compares the performance of the new test of cognitive style with a well-known, similar, reliable and validated test. Zhang & Sternberg (2005) conclude that "clearly there are similarities among these style constructs" and they further state that: "However, by no means do these substantial overlaps and similarities give us reason to assert that that any of these two constructs are identical".

3.6 An Overview of the Style Field

Desmedt and Valke (2004) present a refreshing overview of the field of cognitive and learning styles, based largely on citation analysis. They acknowledge the limitations of this approach and make suggestions as to how the citation analysis could be improved. Their review of style assists in distinguishing between learning styles and cognitive styles.

They also reported that there are seven main areas of research interest in the style domain. Their research presents an overview of these disparate aspects and depicts overlap between learning style and cognitive style research (Vid. Fig. 3.1).

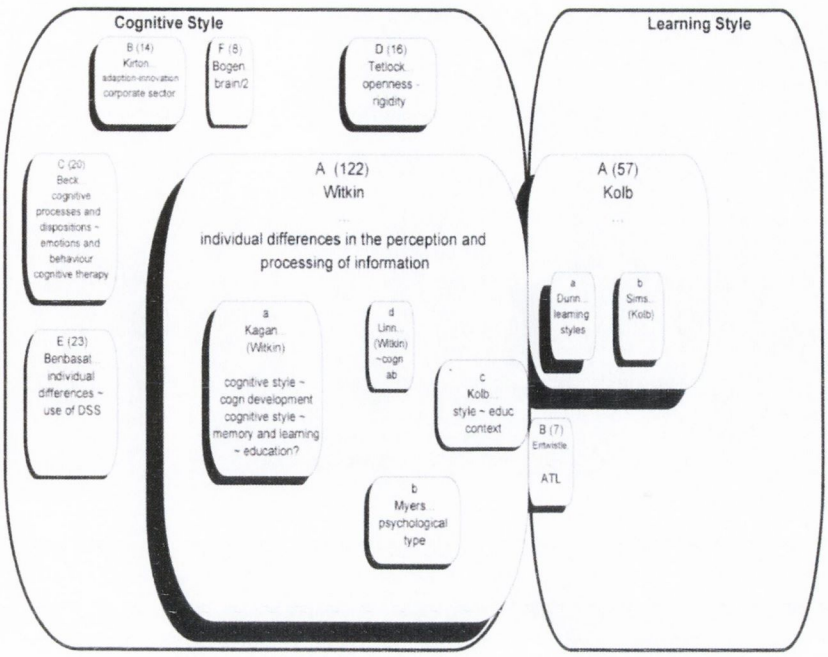


Fig. 3.1 Taken from: Mapping the Learning Styles "Jungle" (DeSmedt & Valcke (2004))

Sternberg's construct of thinking style is not mentioned. Citation analysis has its limitations and this is acknowledged in their paper. Particularly relevant from the review is the propensity of journal editors to publish only positive results from a statistical viewpoint. By setting criteria the authors' subjectively limited their review scope. This overview was innovative as it provides a bird's eye view of the extant style landscape.

Using citation analysis the review established that there were seven main groupings of research in the cognitive styles domain were:

- the study of individual differences in perception and the processing of information,
- cognitive processes and dispositions
- emotions and behaviour therapy,
- individual differences in the use of Decision Support Systems,
- Kirton's Adaptor – Innovator cognitive style in business,
- Brain studies and
- Tetlock's openness - rigidity (Adapted from Desmedt and Valcke (2003)).

Most popular was work relating to that of Witkin et al. (1971) - Group 1 - which was sub-divided into four smaller groups, which contain the work of Kagan, Myers, Kolb and Dunn & Dunn (1989).

3.7 The Coffield Report (2004)

The Coffield report reviewed 13 of the most important learning style models used in the style domain. This review is probably one of the most comprehensive and critical reviews to date in the area. This is welcomed as it clearly delineates the style field into what instruments are worth using and those that are not. Table 3.2 gives Coffield's categrisation of the style domain into families of styles.

Table 3.2 Coffield Report's Families of Styles

Learning styles and preferences are largely **constitutionally based** including the four modalities: VAKT	Learning styles reflect deep-seated features of the **cognitive structure** including 'patterns of abi;ity'	Learning styles are one component of a relatively **stable personality** type	Learning styles are **flexibly stable learning preferences**	Move on from learning styles to **learning approaches, strategies, orientations** and **conceptions of learning**
Dunn and Dunn	**Riding**	**Apter**	**Allinson and Hayes**	**Entwhistle**
Gregorc	Broverman	**Jackson**	**Herrmann**	**Sternberg**
Bartlett	Cooper	**Myers-Briggs**	**Honey and Mumford**	**Vermunt**
Betts	Gardner et al.	Epstein and Meyer	**Kolb**	Biggs
Gordon	Guilford	Harrison-Branson	Felder and Silverman	Conti and Kolody
Marks	Holzman and Klein Hudson	Miller	Hermanussen, Wierstra, de Jong and Thijssen	Grasha-Riechmann
Paivio	Hunt		Kaufmann	Hill
Richardson	Kagan		Kirton	Marton and Saijo
Sheehan	Kogan		McCarthy	McKenney and Keen
Torrance	Messick			Pask
	Pettigrew			Pintrich, Smith
	Witkin			Garcia & McCeachie
				Schmeck
				Weinstein, Zimmerman & Palmer
				Whetton and Cameron

However some models of style that have been influential in style research were not included. Two notable exceptions are Witkin's

Field dependence - Independence and Kagan's Impulsive - Reflective model.

The Report assessed each of the 13 models chosen using psychometric measures such as a) internal consistency, b) test retest reliability, c) construct validity, d) predictive validity. It relied mainly on external validation by other researchers as opposed to those who developed the tests themselves.

The Coffield Report (2004) concluded, based on an extensive literature review, that only one model met the four criteria: the CSI developed by Allinson and Hayes (1996). This report also found that neither Riding's CSA nor Sternberg's model of style Theory of mental self-government) met any of their four criteria. Therefore there is a tremendous psychometric weakness associated with many of the most influential instruments in use today. Coffield (2004) reports, "that too much is being expected of relatively simple self report tests". It appears that developers of such tests have ignored the dangers of designing such self reporting instruments. Riding and Rayner (1988) have already mentioned factors such as social desirability, and an inability to self characterize as inaccuracies in self-reporting measures. Coffield (2004) reports on the appropriateness and demands on the respondent when designing some of the test items used in these questionnaires. They cite Honey and Mumford's questionnaire (Honey & Mumford (1992)) as an example: here respondents are forced to answer (agree/disagree) in eighty questions like "People often find me insensitive to their feelings". The design of test items such as these and selections of response options is quite absolute and are not sometimes a true reflection of what happens in reality.

This may raise concerns about the accuracy of the interpretation of the vast quantities of literature reviewed and the conclusions based on them. Peterson (2005), in her review of the report, comments on the use of value-laden terminology in this report. She believes that it "unnecessarily biases readers against a style rather than allowing them to make their own informed decision based on the evidence provided".

One might wonder why the Coffield Report review didn't include Kirton's KAI (1976); this is a serious omission as Kirton's

KAI has been reported as having acceptable reliability and validity. If this report now functions as a starting point for other new researchers interested in the style area Kirton's model may now be unfortunately overlooked by them.

Putting these criticisms aside, this report has certainly pruned the cluttered style landscape and has identified the psychometric deficiencies of some of the most influential style models. It is plausible to suggest that a new model of style may be formulated based on some of the more positive characteristics of the thirteen models identified by this report. This report may also serve to foster debate and subsequent advancement in the style area.

3.7.1 Existing Difficulties in the Style Area

In a recent review of the literature Armstrong (1999) identified at least thirty style labels. This pluralism of style labels has resulted in much confusion. Accordingly, it has been exceedingly difficult to present a coherent view of the cognitive style / learning style domain.

There is a need to develop a definite model of style, which will enhance the credibility of the field and resolve the confusion that exists within the style domain. Previous attempts have shown this has proven to be a difficult task.

Styles have generally tended to be developed in relative isolation from each other. The psychometric properties of some of the instruments used to measure lack statistical rigour. Some test manuals fail to provide evidence of a) internal reliability, b) convergent validity c) discriminant validity. In some cases it is not evident as to how the style relates to a current underpinning psychological theory.

Curry (1990) highlights some issues of concern regarding learning style use, which include "(1) confusion in definitions, (2) weakness in reliability and validity of measurements." She states that, "test users should expect evidence indicating that the instrument meets minimum standards for use and interpretation".

Users should only buy Psychometric Tests which are certified as meeting the Standards requirements of the appropriate professional bodies such as the American Psychological Association, the American Educational Research Association and the National Council for Measurement in Education and/or the British Psychological Society. Standards should cover the following areas at least: Validity, Test development and revision, and reliability. Unfortunately this has not always been the case in practice. Tiedemann (1989) states that, "requirements of objectivity, reliability and validity are seldom fulfilled."

Riding's manuals for both the CSA and the IPI (Working Memory Test) are examples of many lacking this type of information (Riding (2000a, 2000c)). Some researchers have on the other hand been quite thorough in their test publication material; one of the best examples being Kirton (1987).

3.7.2 Three Style Approaches - a Model

Grigorenko and Sternberg (1995) suggest that work on a style model can take one of three approaches: a) a cognition centered approach, b) a personality based approach and c) an activity centered approach (Riding & Rayner (1998), Sternberg & Grigorenko (1997)).

The cognition centered approach focuses on the relationship between cognition and cognitive style with particular emphasis on perceptual and cognitive functioning. Styles in this domain "most closely resemble abilities" (Zhang & Sternberg (2005)). The personality-based approach examines styles, which most closely resemble personality traits. This grouping would include the well-known MBTI (Briggs-Myers (1990)). Finally the activity centered approach focuses on style in relation to environments, settings and activities. Zhang & Sternberg (2005)) argue that activity centred styles are most closely related to learning strategies than are styles from the other two approaches. Example styles pertaining to this categorization are Kolb (1976), and the Dunn and Dunn model (1989).

Sternberg & Grigorenko's (1997) model identifies three distinct traditions within the style domain and groups them based on similarity. However this model hasn't been thoroughly validated yet: "the validity of Grigorenko and Sternberg's integrative

model of three kinds of styles has not been tested" (Zhang and Sternberg (2005)).

3.7.3 Curry's Style Model

Curry (1983) suggested that learning style/cognitive style could be integrated into three layers or strata. She states "learning behaviour is fundamentally controlled by the central personality dimension, translated through the middle strata information processing dimension and given a final twist by interaction with environmental factors encountered in the outer strata".

The inner-most layer is termed the cognitive personality layer. Many of those cognitive styles categorized here are examples of the cognition-centered approach. This layer can be defined as "the individual's approach to adapting and assimilating information which does not interact with the environment" (Curry (1983)). Curry suggests that the inner-most layers should "be more stable than that of the others". "This inner-most cognitive layer does not interact with the environment, but is an underlying and relatively permanent personality dimension that is expressed indirectly and is apparent only when an individual's behaviour is observed across many learning instances" (Riding and Cheema (1991).

The next layer in the model is the information-processing layer. Curry (1983) believes that measures of information processing are comparatively more stable than instructional preferences and can be changed by learning strategies. According to Curry (1983) this layer "is considered as the individual's intellectual approach to assimilating information and, because this processing does not involve the environment these measures are more stable than instructional preferences but can be modified by learning strategies." LSI (Kolb (1976)) and the Inventory of Learning Processes (Schmeck (1977)) are examples of some learning styles, that are thought to be categorized by this level in Curry's model.

The outermost layer in Curry's model is referred to as the Instructional Preferences layer. As this layer interfaces directly with the environment it is considered to be the least reliable of the three. This is because the layer: "interacts most directly with environments, learner expectations, teacher expectations and

other external features" (Curry (1983)). Instructional preference can be thought to reflect an individual's selection of learning environment in which to learn. Curry believes that this layer of her model is the least stable as it is affected by other extraneous factors such as the environment, teacher and learner expectations. In her original model this outermost layer included such measures as the Learning Preference Inventory of Rezler and Rezmovic (1981).

3.8 Riding and Cheema's Model (1991)

This model was proposed in an attempt to categorise the large amount of style labels reported in the literature at that time. Messick (1976) reported 19 labels while Riding and Cheema (1991) chose over 30 labels.

Riding and Cheema (1991) hypothesized that there were two super-ordinate dimensions that could be used to categorise cognitive styles. Based on the arguments of other researchers who argued that "a number of cognitive style labels are, but different conceptions of the same dimension", they suggested the Wholistic – Analytic and Verbaliser - Imager dimensions. They presented some empirical evidence using factor analysis in support of their model.

Riding and Rayner (1998) included other cognitive style labels not previously included by Riding and Cheema (1991). This categorization aligned nine cognitive style labels along the Wholist – Analytic dimension and two along the Verbaliser - Imager dimension.

Table 3.4: Adapted from Riding and Rayner (1997, 1998)

Wholist	Analytic	
Field dependence	Field independence	Witkin and Asch (1948)
Leveling	Sharpening	Gardner et al. (1959)
Impulsivity	Reflectiveness	Kagan et al. (1964)
Adaptors	Innovators	Kirton (1976,1987)
Assimilator	Explorer	Kaufmann (1989)
Concrete - sequential	Abstract- sequential	Gregorc (1982)
Holistic	Serialist	Pask (1976)
Intuitive	Analytic	Allinson and Hayes (1996)

Table 3.5 Riding and Cheema (1991)

Wholists	Analytics
Field dependence	Field independence
Impulsive	Reflective
Levellers	Sharpeners
Divergers	Convergers
Holists	Serialists

This model attempted to provide a cohesive account of cognitive style labels. It is not without its problems. A peculiar inclusion in this categorization is Kirton's Adaptor -Innovator cognitive style as his KAI dimension correlates with some measures of personality (Table 3.4) (Isaksen, Lauer & Wilson (2003)).

Jones (1997) argues that Impulsivity – Reflectivity should not be included as part of the Wholist – Analytic dimension. She refers to studies that suggest that Impulsivity - Reflectivity is "value laden" and "is more likely to be susceptible to change by internal and external influences" thereby raising some concern about its acceptance as a cognitive style. She suggests that it may be positioned in "Curry's Onion Model" as "an information processing style".

Study 3.1 (Parkinson & Redmond (2001)) provides some evidence to suggest that field dependence-field independence and Kirton's KAI fail to support the Riding & Cheema model empirically. Sixty students were administered the Geft, KAI and the CSA. The test scores for the Geft and the KAI correlated poorly (0.173, -0.02 respectively) with the Wholist - Analytic dimension of Riding's CSA (Table 3.6). A correlation matrix for Study 3.1 is given in Table 3.6

Table 3.6 Pearson Product Moment Correlation Matrix for Some Cognitive Style Scores.

Pearson Product-Moment Correlation

No Selector

	wholis_	verbali_	adapto_	fd_fi
wholist_anal_	1.000			
verbaliser_v_	-0.092	1.000		
adaptor_inov_	-0.020	0.159	1.000	
fd_fi	0.173	0.089	0.392	1.000

One might have expected that there would have been a greater correlation between these measures. The model also assumes that these grouped cognitive style labels and their instrumentation have much in common. This is a questionable assumption, and a similar reservation has been expressed by Armstrong (2005).

To conclude, the Riding & Cheema model (as referred to by Riding and Rayner (1998)) has not been validated by means of a detailed factor analysis.

Cassidy (2004) reviewed the field of learning style and developed a taxonomy based on Sternberg's classifications, as outlined previously, along with Curry's model and Riding & Cheema's (1991) model. This review provides a description of the main cognitive styles and learning styles and evidence supporting their reliabilities and validities. This is an interesting perspective as it attempts to provide an integrated view of the style area. It is yet another taxonomy developed to provide an alternative perspective on the style domain. However it is an open question as to how beneficial this approach is and whether it has provided greater insight into this style area?

It is interesting to note that the other orthogonal dimension of Riding & Cheema's model, the V-I dimension, is missing from Cassidy's classification.

The models and taxonomies reviewed serve as interesting ways to gain a holistic way of organizing the style field. As mentioned above very seldomly have these models been thoroughly substantiated. Models of style and individual style conceptualizations, if worthwhile, should by definition bear up statistical scrutiny. This approach would then lend more credibility to the style domain.

3.9 Coffield's Continuum of Style Families

Coffield et al. (2004) developed a continuum that consisted of five categories of style families. As the authors state, this is a "classificatory system to impose some order on a particularly confusing and endlessly expanding field". The authors also acknowledge the weaknesses of this approach "for example it

may not over emphasize the differences between families and cannot reflect the complexity of the influences on all 13 models". The five families of style identified were -

1. Constitutionally - based learning styles and preferences
2. Cognitive structures
3. Stable personality type
4. Flexibly stable learning preferences
5. Learning approaches and strategies

There are other models of cognitive style based on hemispheric preference which are not considered in this review as they tend to be based on the assumption that the right and left hemispheres of the brain are involved in different cognitive functioning and processing of information (Sperry (1964)). Research reviewed in Chapter 8 would seem to suggest that this perspective can be no longer sustained. Therefore models such as Herrmann's brain dominance theory (Herrmann (1944)) are not reviewed. Furthermore this model has not been thoroughly validated (Allinson and Hayes (1996)).

3.10 Review of some of the main cognitive styles

This section reviews some of the main cognitive styles in the field at the moment. This review is based on the reviews of Tiedemann (1989), Riding & Cheema (1991), Riding & Rayner (1998), Hashway & Duke (1992), Jonassen & Grabowski (1993), Coffield (2004), de Smedt & Valcke (2004) and Cassidy (2004).

3.10.1 Field dependence – Field independence

This is the most researched cognitive style with over 3000 studies reported in the literature (Moran (1985)). It describes the "the extent to which the organization of the prevailing field dominates perception of any of its part" (Witkin, Oltman, Raskin and Karp (1971)). An individual who finds it difficult to disambiguate objects from a complex field is field dependent. On the other hand an individual who is field independent finds it easier to distinguish objects from the surrounding perceptual field.

This construct was initially measured using a rod and frame and more popularly today the Embedded Figures Test (EFT) and the Group Embedded Figures Test (GEFT). Some researchers argue

that there are unresolved issues pertaining to the construct (Moran (1985)).

The Rod and Frame test comprises a luminous rod within a rectangular frame. The participant and his apparatus are situated within a darkened room. The luminous rod is rotated from the vertical with respect to the frame by the experimenter. The subject is required to rotate the rod until it is positioned vertically again. The index of field dependence is the angle between the rod and the frame. The more the angle that the subject rotates the rod the more field-dependent is the testee (Hashway & Duke (1992). A portable version of the rod and frame test was developed by Oltman (1968) which has reported correlations of r = 0.89 between these two measures (the Rft (i.e the rod and frame test) and Prft (the portable rod and frame test). Later, this construct was assessed using the Embedded Figures Test (EFT) and the Group Embedded Figures Test (GEFT). The EFT had been considered to be a suitable alternative to the Rod and Frame test due to ease of administration. However Goldstein & Blackman (1978) report that there are poor correlations between the GEFT (a competence test) and the Rod & Frame Test (a preference test). These correlations range from 0.3 to 0.65.

McKenna (1984) reviews studies which indicate that the Embedded Figures Test (EFT) correlates with general levels of ability such as the Scholastic Test and the Wechsler Adult Intelligence Scale (Wais) in a range from 0.08 to 0.50. The studies referred to include Witkin et al. (1977) and Bieri et al (1955). Some of the earlier studies indicated that the EFT tended to correlate higher on the non-verbal component of the Wais (0.37 -> 0.66) and poorly on the verbal component (0.16 -> 0.27). The EFT was also found to correlate with Raven's Matrices in a series of studies (McKenna (1984)).

Moran (1985) for instance, refers to methodological problems pertaining to research on Field dependence/ Field independence such as:

- a proliferation of measures
- inadequate control of intelligence
- inadequate construct validation of tests.

There have been other criticisms of this test. Some researchers report that the GEFT correlates with some tests of IQ. Other researchers have argued that it only measures the field independent end of the continuum. The Geft is an accuracy-based and time-based measure with the emphasis on competence as opposed to preference.

There may be other factors that might account for a low score on the GEFT (Parkinson and Redmond (2001)) including: sight defects for instance: Daltonism, short sightedness and short-term memory impairment.

When the testee is completing the GEFT it is necessary to disembed a figure from within a complex field. However the shape that requires disembedding is listed at the back of the booklet. The subject doing the test has to be able to remember the shape and then extract this shape from the complex field. Perhaps there is a degree of imaging ability involved here. The test also requires fine psychomotor skill. The subject's psychomotor skill has to be well developed in order to trace around the shape, which isn't controlled for in the test.

Despite these issues there is a consensus that the construct is valid whereas the measure (the GEFT) may be problematic. The reliability of the Geft is reported as 80% in Witkin et al. (1971).

3.10.2 Leveling and Sharpening (Holzman and Klein (1954))

Jonasssen and Grabowski (1993) state that "This cognitive style .... describes how individuals perceive and memorize images. Levelers ... miss changes or inconsistencies in sequentially presented stimuli, strong sharpeners .... tend to retain discrete and clearly differentiated images and the small differences between them". Central to this dimension is the concept of assimilation. This is what distinguishes the Sharpeners from the Levelers. The Leveler assimilates, details are merged together and specifics are lost. The case of the Sharpener is in deep contrast. Little or no assimilation occurs, each detail is remembered, and each aspect is a unique and a distinct event.

The Schematizing Squares test was developed initially to measure this style (Holzman & Klein (1954)). No reliability data

have been documented in the literature nor are available for this measure. Tiedemann (1989) comments, that "the concept of leveling-sharpening has come to grief due to measurement problems". He refers to the studies of Vick and Jackson (1967) and Pritchard (1975) that raise doubts about the validity of the Schematizing Squares test. Other measures were then developed to assess this construct as discussed by Santostefano (1964) and Guthrie (1967) in Jonasssen and Grabowski (1993). The different measures designed by these researchers to assess the Leveling - Sharpening dimension lack substantive psychometric investigation.

3.10.3 Convergent – Divergent Style (Guilford 1967) and Hudson (1966, 1968))

Riding & Rayner (1998) describe the Converger as being "narrow, focused, logical and deductive with regards to their thinking style". The Diverger can be best described as being "broad, open ended, displaying associational thinking to solve problems".

Hudson (1966, 1968) has reported a correlation between intelligence and the measure designed to assess this style dimension. Tiedemann (1989) provides some supporting evidence. "A critical point in this cognitive style dimension is the emphasis on ability measures in assessing a preference construct".

Cassidy (2004) asserts that "Assessment of convergent thinking is the more straight forward of the two, using standard intelligence tests, multiple choice questions as well as being inferred from performance on the EFT and MMFT".

Riding & Rayner's (1998) review and Cassidy's review (2004) do not take cognizance of this important point. If this measure is so closely associated with standard intelligent tests then how could it be considered to be a measure of cognitive style, which is supposed to be independent of intelligence?

This casts some doubt on the validity of this style measure. Tiedemann (1989) reports, that "there is insufficient evidence of a cognitive style dimension of convergent thinking versus divergent thinking as a bipolar preference dimension".

3.10.4 Concrete (Sequential/Random) – Abstract (Sequential/ Random) Gregorc (1982)

Gregoric (1982) describes four different types of behaviour: Concrete Random, Concrete Sequential to Abstract Sequential and Abstract Random. This style is assessed using the Style Delineator. The ends of the continuum range from Concrete to Abstract and bear a resemblance to the work of Piaget.

3.10.5 Intuition – Analysis Style (Allinson and Hayes (1996))

This style is a preference for a type of reasoning. Allinson and Hayes (1996) believe that "Intuition, characteristic of right brain orientation, refers to immediate judgment based on feeling and the adoption of a global perspective. Analysis characteristic of left brain orientation, refers to judgment based on mental reasoning and a focus on detail"

This style is assessed by means of a questionnaire (Cognitive Style Index - CSI) containing thirty-eight test items. Respondents choose between three options when answering each of the questions: true, false, and uncertain.

The instrument has been reported as having good test retest reliabilities (Armstrong (1999)), Allinson & Hayes (1996) and internal reliabilities (r ranges from 0.84 - 0.92). Allinson and Hayes (1996) report a test-retest reliability for the CSI of $r = 0.90$, even though the sample size might be considered small in this instance ($n = 30$).

The validity of this style has been the subject of some debate in the literature recently (Hodgkinson & Sadler-Smith (2003), Hayes et al. (2003)).

The recently published Coffield Report (2004) comments that "of all the instruments we have evaluated, the cognitive style index (CSI) of Allinson & Hayes has the best psychometric credentials despite the debate about whether it should be scored to yield one or two measures of intuition and analysis". It also notes that it is an appropriate instrument for use in educational and business environments for the identification of entrepreneurs.

What is of concern is the assumption by these authors that Intuition is characteristic of right brain function and Analysis of

left-brain function. As Coffield (2004) points out "they follow Mintzberg (1976) in linking right brain intuition with the need of managers to make quick decisions ... while left brain analysis is seen as the kind of rational information processing that makes for good planning". In the early seventies Sperry's Hemispheric Processing model was very popular but is considered by some to be just a useful metaphor today (Gazzaniga (2002)).

3.10.6 Impulsivity – Reflectivity (Kagan (1964))

This dimension identifies two types of people. At one end of the continuum there is the Reflective type. This individual has the tendency to think about the accuracy of the answer before answering. The Impulsive type of individual responds very quickly with little or no inhibition at all making many more mistakes in comparison with the reflective.

The MFFT (Matching Familiar Figures Test - (kagan (1964)) is used typically to measure the Impulsive –Reflective continuum. It is an accuracy-based measure. Some researchers have questioned its validity (Block, Block & Harrington (1974), Block, Gjerde and Block (1986), Jones (1997)).

Jonassen and Grabowski (1993) refer to studies that report high internal reliabilities and test retest reliabilities for this instrument. Lajoie and Shore (1987) established that the MFFT was independent of intelligence. Others dispute these findings (Brannigan, Ash and Margolis (1980)).

3.10.7 Holists and Serialists

Holists use a global approach to learning; they are conceptually orientated and have low discrimination skills (Jonassen and Grabowski (1993). The Serialist works sequentially, emphasizes details and has high discrimination skills.

Pask (1973) assessed this style dimension using the Spy Ring History test (Pask & Scott (1972)). The test is preference based and there are no reliability data available for this test. Pask later on developed the Caste and Intuition test (Pask (1976)). This was computer based. Again no reliability data were presented.

Riding and Cheema (1991) point to problems with the sample size used to investigate the Holists - Serialists construct: "It is

based only on a small sample of students; secondly all the students were above 15 years of age". However the authors state "despite the outcomes of Pask's research, it "is suggestive of individual differences that could be very important in the educational system consequently this has led to the widely recognition of the Serialist as a cognitive style". Such a claim provides a dubious basis for the inclusion of any construct for consideration as being a cognitive style.

3.10.8 Adaptor – Innovator

This describes two types of individual: the Adaptor and the Innovator. The Adaptor is characterized by "precision, reliability efficiency, prudence and conformity". The Adaptor also seeks to "find solutions to problems in tried and trusted ways". Adaptors are authorities within a given structure whilst "rarely challenging the rules". In contrast, the Innovator is seen as "undisciplined, approaching tasks from unsuspected angles". With respect to problem solving, the Innovator "queries ... assumptions, manipulates problems". Innovators often challenge rules and would have little respect for past customs" (Kirton (1976)).

When the Adaptor is collaborating with the Innovator, he/she supplies "stability, order and continuity to the partnership". In comparison, the Innovator when collaborating with the Adaptor "supplies the task orientations, the break with past and accepted theory". In interacting with others, the Adaptor is "sensitive to people, maintains group cohesion and co-operation" while the Innovator is quite the opposite (Kirton (1976)). Kirton (1976) also states, that "it is assumed that Adaptors and Innovators bring incommensurate viewpoints and different solutions to administrative and organisational problems".

The KAI is a preference based measure and is used to assess this dimension using a questionnaire. The KAI is a self-report questionnaire, which consists of 32 test items. These test items ask the testees to indicate how easy or difficult they find it to behave consistently over a long period of time. Tullett (1977) reports that the KAI "is a stable cognitive process within the individual which is largely uninfluenced by national culture". He based his findings on numerous studies undertaken in a variety of different countries where the KAI was translated from

English into the local language (French, Dutch, Slovak and Italian).

The KAI assesses the Adaptor - Innovator continuum by providing a range of scores from 32 to 160. The Adaptor scores below 96 and the Innovator scores above 96. It has been shown to be very reliable (Clapp (1993)) and has been extensively validated (Kirton (1982)). This cognitive style has been extensively researched in corporate environments (Desmedt and Valcke (2004)). Clapp (1993) reports a test retest correlation of r = 0.82 for 69 UK managers, with a test retest interval of 41 months Tullett & Kirton (1995) report Cronbach's alpha as ranging from 0.87 to 0.89.

3.10.9 Models NOT Reviewed

There are some models not included in this review, such as the Hermann Brain Dominance Instrument (Hermann (1988)), since they are based on the questionable notion of hemisphericity (Vid. Chapter 8). Springer and Deutsch (1998) refer to pencil and paper tests which are developed with goals of assessing hemisphericity. They refer to the Herrmann Brain Dominance Instrument (HBDI). Springer and Deutsch (1998) are dismissive of such instruments as they state, "before the idea of hemisphericity can be evaluated "we need good measures of differential hemispheric activity". They refer to the problems of assessing this even using new neuroimaging techniques and in this context they comment "even if the techniques prove more positive it would still remain to be seen how effectively the results of these paper and pencil tests would reflect this hemispheric activity". They are not convinced about the legitimacy of hemisphericity. "All in all, hemisphericity currently remains an interesting but fundamentally untested hypothesis and techniques asserting that they are based on the concept of hemisphericity are built on a shaky foundation".

3.10.10 Riding's CSA (1991): Wholist – Analytic dimension Verbaliser – Imager dimension

This assesses two dimensions of cognitive style the Verbaliser – Imager dimension and the Wholist – Analytic. These two dimensions are orthogonal (Redmond, Mullally & Parkinson (2002), (Parkinson, Mullally & Redmond (2004c)). The assessment is done using the CSA and is computer based. There

has been much evidence to support its construct validity and neurophysiologic underpinnings (Riding & Sadler-Smith (1992), Riding & Douglas (1993), Riding, Glass & Douglas (1993), Riding & Reid (1996), Riding & Agrell (1997), Riding, Glass, Butler & Pleydell (1997)).

The Verbal - Imagery dimension of cognitive style has been investigated by Riding and his co-researchers since 1976. Riding & Taylor (1976) developed a test of this dimension, which led in turn to the Verbal - Imagery Code computer-based test (Riding & Calvey (1981)). This research was extended to lead ultimately to the CSA (Cognitive Styles Analysis) in 1991.

Until relatively recently there were no studies which attempted to examine its reliability. Recently this issue has become quite topical in the literature, Redmond, Parkinson & Mullally (2002) being the first to demonstrate and report the unreliability of the CSA. These findings were supported by Rezaei and Katz (2003). Parkinson, Mullally & Redmond (2004b) investigated the unreliability of the CSA again, using different test retest intervals and again demonstrated its unreliability (See Chapter 5 for further details).

Peterson, Deary and Austin (2003a) have also supported these findings indirectly by raising some doubts about the CSA's reliability using their own re-engineered model of the CSA. Her results provide evidence that her approximate model of the CSA was not reliable. In evaluating Peterson et al.'s findings, it is also necessary to ask how good a model of the CSA was this? These authors provide no evidence of this as they never compared the performance of their model with the performance of Riding's own copyrighted version whether on the same or a different sample by means of an empirical study. They could not compare their approximate model of the CSA with Riding's CSA directly since Riding has never released the code of his test.

3.10.11 Peterson, Deary and Austin's Work

Peterson, Deary and Austin (2003a) report that by doubling the number of test items in their simulated form of Riding's CSA they increased the test retest reliability of the W-A dimension (combined form) from 0.3 to a more acceptable level of about 0.53.

3.10.12 Individual Differences Questionnaire (Paivio (1971))

This questionnaire was designed to overcome the inadequacies of self-report measures of imaginal ability such as the Betts' Inventory (1909), an abbreviated form of the Betts' Inventory (Sheenan (1967)) and the Gordon Scale of Imagery Control (Gordon (1949)). These inadequacies are discussed by Pavio & Harshman (1983) who comment "they often fail to show correlations with objective performance on tasks where good images would presumably be helpful". These authors also comment that these self-report questionnaires were deficient in "the lack of internal controls for response sets such as acquiescence and social desirability".

The Individual Differences Questionnaire was designed to measure Verbal and Imaginal thinking habits and skills (Pavio and Harshman (1983)). It contains eighty-six test items to cover a continuum: high scores reflecting visual tendencies and low scores verbal tendencies. Riding and Cheema (1991) reviewed the IDQ and mentioned that there was not any evidence available that supported its validity. Riding and Rayner (1998) refer to studies, which examined the reliability of the Individual Differences Questionnaire and stated "that the measure carried a moderate level of reliability". They did not present any correlation coefficients in their review of this test but do refer to the work of those who did (Richardson (1977); Irwin (1979); Cohen and Saslona (1990)).

Paivio and Harshman (1983) report that the internal reliability for the Imagery scale was $r = 0.80$ and $r = 0.83$ for the Verbaliser scale.

3.10.13 Verbaliser – Visualiser Questionnaire – Richardson (1977)

This is a fifteen item questionnaire containing true / false questions. Eight of the questions refer to the visual domain, seven to the verbal domain. The VVQ was based on subjects' correlated scores between lateral eye movements and Pavio's (1971) Individual Differences Questionnaire. Riding and Cheema (1991) report that the VVQ "has been used little in instructional research". The test retest reliabilities have varied from $r = 0.48$ to $r = 0.91$ (Jonassen and Grabowski (1993)). Richardson (1977) reports test retest reliability for the VVQ of $r = 0.91$. The test

retest interval was seven days and the sample sizes were n = 20 males, n = 17 females.

Antonietti & Giorgetti (1998) reviewed studies which used Richardson's VVQ. They report that (a) "the Verbaliser - Visualiser Questionnaire does not measure a unidimensional construct", and (b) it fails to "predict the actual use of imagery in thinking" and finally (c) "a lack of long-term reliability of the questionnaire emerged". The authors of the review conclude that the "use of the questionnaire to assess the Verbal - Visual style remains questionable".

3.11 Conclusions

This brief review raises some important issues in the fields of cognitive style and learning style. The various cognitive styles reviewed above either use performance, self-report questionnaires, or to a lesser extent preference measures. Self-report instruments have several disadvantages.
Riding and Rayner (1998) refer to the disadvantages of the use of self report measures including: "an individuals' possible inability to process and objectively report their behaviour, unwillingness to make the necessary effort to respond accurately, bias due to the pressure of social desirability in making responses and inclination to contrive their responses".

In some cases there have been concerns about reliability and validity of the cognitive styles reviewed.

Hashway and Duke (1992) refer to "a very serious problem ... concerning instrumentation" stating, that "The instruments tend to lack concurrent validity". They also refer to the tendency to have used University students as the sample in testing the psychometric tools. They comment on the generalizability of findings using this type of sample as applied to the population as a whole. Hashway and Duke (1992)) comment "Researchers have generalized the results achieved from the analysis of these limited samples to the universe of all people".

The accurate assessment of style dimensions is problematic. Some style labels have not been thoroughly investigated before release into the field.

The Coffield Report (2004) states that there is just one test of Cognitive Style that is reliable and valid after completing a comprehensive review of seventy psychometric tests in the area. These style measures may provide the basis for a new model of style in this area.

This Chapter reviewed definitions of cognitive style, models of style and finally some of the most popular style labels. Some conclusions that can be drawn from this review of the literature are:
- there still exists a multiplicity of style models, labels and definitions.
- Many of the Style instruments in use today are psychometrically deficient
- the accurate assessment of the Verbaliser - Imager dimension is still proving to be difficult.
- To date, the Coffield Report (2004) along with Desmedt & Valcke (2003) and Cassidy (2004) provide the most comprehensive analyses of the Style domain. It would appear that Allinson & Hayes' CSI and Kirton's KAI are the most sensible psychometric test options at the moment.

Cognitive Styles may be an important feature in how the user makes use of the human computer interface efficiently, effectively and creatively. The cognitive style literature has been comprehensively reviewed in this Chapter. The next chapter discusses how the human computer interface should be designed so as to accommodate Individual Differences such as cognitive styles. Cognitive styles may also be an important influence in how designers structure the human computer interface.

Chapter 4:
Cognitive Styles & Human Computer Interface Design

4.0 Introduction

In Chapters 2 and 3 flexible systems and cognitive styles were reviewed since they can be of importance in improving the user's performance in the human computer interface. Following on from these two chapters, some questions arise from the merging of these themes.

Some studies which have implemented various flexible systems ranging from non-adaptive (static) to flexible systems which have taken cognisance of individual differences (cognitive styles) are discussed.

Firstly there is an important question as to what affect cognitive styles in a computer-based environment have on user performance? The second question is whether, by the inclusion of better navigational aids, for instance, the field-dependent's performance can be improved to that of the field-independent learner? Thirdly, could user performance be improved by adapting the human computer interface based on the user's cognitive style score from Riding's CSA so as to improve user learning performance? The chapter then discusses how others handled cognitive styles in the design of computer system interfaces. This involves a discussion of the matching hypothesis, which underlies the whole idea of adapting the HCI to the user's individual differences. Ways of matching individual differences to a Human Computer Interface has evolved in this thesis from cues through to Cognitive Style assessment over time and is clearly shown in this Chapter and in Chapter 9.

This chapter begins with a description of a study which investigated the affect of a variety of cognitive styles in computer based environments.

There are three exploratory pilot studies described in this Chapter which are used to pave the way to more in-depth investigations. Study 4.1 investigates whether cognitive styles affect learning performance under different computing media.

Study 4.2 explores whether the performance of the field-dependent can be brought to the same level as a field-independent on a website. Study 4.3 investigates whether Verbalisers - Imagers could be accommodated by presenting a matched set of pages depending on their CSA scores.

Finally in this Chapter there is a discussion of the matching/mismatching hypothesis since it underlies the various studies that have attempted to accommodate cognitive style in interface design.

4.1 Introduction to Study 4.1

Shneiderman (1998) has emphasized the importance of cognitive style commenting: "A clear understanding of personality and cognitive styles can be helpful in designing systems for a specific community of users". Cognitive Styles research is a significant predictor of an individual's educational success. Cognitive Styles provide valuable insights into how individuals apparently learn. Sternberg and Grigorenko (1997) suggest, "cognitive styles seem to have serious implications for educational theory and practice".

In Study 4.1, Cognitive Styles and their impact on learner success were investigated in three different educational computer assisted environments in which text presented via floppy disk, multimedia Cd-Rom and the Internet. The cognitive styles investigated were: Riding's CSA (W-A and V-I dimensions), Witkin's field dependence-independence, Kirton's Adaptor - Innovator continuum.

Computer Assisted Environments provide opportunities for discovery learning and the facilitation of individual differences such as Cognitive Styles. To quote Bruner (1962):
"Emphasis on discovery learning has precisely the effect on the learner of leading him to be a constructionist to organize what he is encountering in a manner not only designed to discover regularity and relatedness, but also to avoid the kind of information drift that fails to keep account of the uses to which information might have to be put".

4.1.2. Study 4.1 Design

The first study addresses the issue whether cognitive styles do actually affect user performance in computer-based environments. Before 1994 few studies had investigated this aspect of computer environments (Ayersman and Minden (1995)).

Jonassen (1989) argues that a well implemented and designed hypermedia system can provide an educational environment that (a) is computer based, (b) exploits multi-nodal approaches to learning and (c) facilitates user decision making opportunity. Central to the concept of hypermedia is the assumption that the mind learns by association (Bush (1945)) and that the objective of a well-designed hypermedia system is to allow individuals to create their own associations.

Computer learning environments can be ranked in terms of their historical evolution, which shows more and more use of hypertext. In the first case an educational product is delivered via floppy disk - an Introductory Course on Artificial Intelligence. There may be some use of hyperlinks and hypermedia but this media option is generally limited because of the inherent limitations of floppy disk (one floppy typically holds 1.44 Mbytes).

The second media option historically presents the material via Cd-Rom. A Cd-Rom has a capacity of the order of 700 Mbytes and as a consequence allows much more scope for the use of graphics, audio and video. The Introductory AI course could be presented as text with the addition of graphics, sound and video.

The third media approach historically is the Internet site. The Internet uses hyperlinks and hypermedia much more intensively than the previous two modes. Indeed it is hard to conceive of the Web working without hyperlinks.

Text on floppy disk(s) was chosen to represent the closest computer medium to a conventional book with no hyperlinking. This medium is also severely restricted by its memory limitations. An example would be a cookery book or a book on Manet; the memory limitation means it would be difficult to have many colourful pictures.

A Cd-Rom can handle text similarly to floppies but because of its much greater memory can also present graphics, audio and video to enhance the presentation of the textual material. In addition hyperlinking is usually highly used. In our cookery book example graphics, pictures, audio and visual material can be used to demonstrate the putting together of a recipe with the various ways of handling the individual constituents. The main problem with a Cd-Rom is that the material is frozen at a point in time.

An Internet site could be regarded as a Cd-Rom with much more hyperlinking possible to a very large number of pertinent sites. Some sites may not facilitate good audio - visual functioning. The web also has the advantage of almost infinite memory and it is also up-to-the minute, not just frozen in time as much as the other two media. It is easier to update and be immediately available. Disadvantages of the Internet may be inadequate access at times and low download speed, but these are improving over time as well as possible contention for a site with other users.

While the course content for each of the three media remained the same, the presentation was adjusted and extended to take account of the natural advantages of each medium.

4.1.3. Study Implementation

Forty-seven final year Information Systems students agreed to participate in this study. They were taught an Introductory course on Artificial Intelligence. The course material was presented in three computer-assisted formats, Text via Floppy Disk, Multimedia Cd-Rom and an Internet-based environment. The students attended over a six-week period, totaling seven hours in duration. The students were randomly allocated to one of the three treatments.

4.1.4 Treatments

The subject matter for all three treatments was an Introductory Course in Artificial intelligence that was run on identical computers. While the basic text remained the same the presentations were enhanced to allow for the natural advantages of each medium. All treatments used in the study were run on the same machine type with identical CPU speed and memory.

The students being IT students were all computer-literate and had no problems using a floppy disk, a Cd-Rom or the Internet to access the appropriate treatments. The treatments used are briefly as follows:

4.1.5 Treatment 1 – Text via Floppy Disk

A set of course notes were presented to the students in a textual, electronic format produced in Microsoft Word on a floppy disk. The students loaded, into main memory, the disk, that they were given at the commencement of each session and interacted with the text in a passive manner. The speed at which they accessed each page of the text was determined by the speed of the machine. The only linking was between modules.

4.1.6 Treatment 2 – Multimedia Cd-Rom

Essentially the same set of course notes were presented to the students on a Cd-Rom that they loaded at the beginning of each session. The multimedia package was designed using Authorware 4 and contained additional text, audio recordings and video clips relative to the floppy disk approach. Hyperlinking existed to a fair extent between and within modules.

4.1.7 Treatment 3 – Internet Web-site

An educational Internet site was developed using Microsoft Front Page. At the commencement of each session the students logged on to this site using a password. The core subject matter was identical to that presented in Treatments One and Two with the addition of numerous links where relevant to other pertinent sites on the Internet. This site was interactive and multimedia in nature containing video, audio and additional textual information. A high degree of hyperlinking was available within the course module and, of course, to the Internet itself.

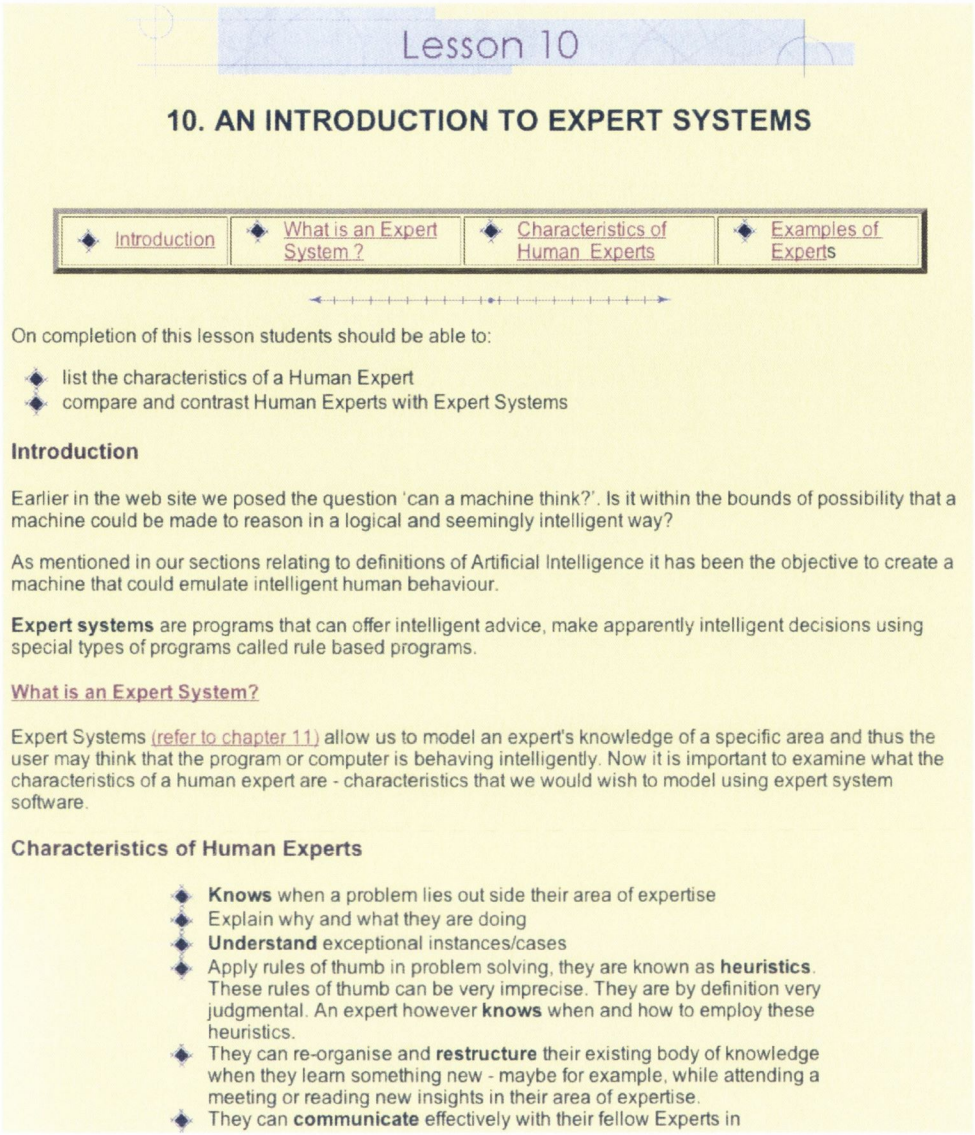


Fig 4.1 Sample web page for web site for Pilot Study (Study 4.1)

This screenshot is part of a typical page of the website used in Study 4.1. It shows hyperlinking embedded in a linear, sequential piece of text which need scrolling by the user to access the contents of the whole page.

4.1.8 Learning Performance – Evaluation

The students' knowledge of the material was examined at the end of six hours of instruction in a test that lasted one hour. The assessment test for the students was made up of two sections – Section A (60%) consisted of thirty multiple-choice questions while Section B (40%) contained five objective-style question. All these thirty-five questions assessed the students' knowledge and

understanding of basic concepts in AI (e.g. forward-chaining, expert system development and application) at both higher and lower cognitive levels.

4.1.9. Psychometric Tests

All participants were administered the GEFT – Group Embedded Figures Test (Witkin (1971)), Kirton – Adaptor-Innovator Inventory KAI (Kirton (1976)) and Riding's Cognitive Style Analysis Riding (Riding (1991)).

4.1.10. Aim of Study 4.1

The general focus was to investigate whether the four Cognitive Styles (GEFT, KAI, Riding's CSA (W - A & V - I)) affected learner performance in the three educational environments.

4.1.11. Results of Study 4.1

The data were analyzed via multiple regression analysis using two packages Data Desk V9 and Minitab V6. Multiple regression analysis of the data allows the quantification of the predictive affect of the independent variables (Cognitive Styles) on final score. A more comprehensive analysis of the results is given in (Parkinson & Redmond (2001), (2002a)) (Vid Appendix 3-4.1.A).

4.1.12. Discussion of Study 4.1

The overall results are summarized in Table 4.1.

Field Dependence-Field Independence

Field Independent students in the study performed better in the Internet treatment than in the Cd-Rom and Text treatments. They also performed better overall.

Verbaliser–Imager

This Cognitive Style was found to have no statistically significant predictive effect on final score in the Cd-Rom treatment in the study.

Wholist-Analytic

It was found that the Wholist–Analytic Cognitive Style did have a predictive effect on final score in the Cd-Rom treatment.

Adaptor–Innovator

Kirton's Adaptor–Innovator continuum was not a statistically significant predictor of learner performance at the $p = 0.05$ level (Kirton (1976)).

Table 4.1 Summary of the Affect of Cognitive Style on Final Score by Treatment

	OVERALL	TEXT	CD-ROM	INTERNET
Field Dependence-Field Independence	**Significant (0.005 level)**	Not Significant	Not Significant	**Significant (0.005 level)**
Verbaliser-Imager	Not Significant	Not Significant	Not Significant	Not Significant
Wholist-Analytic	Not Significant	Not Significant	**Significant (0.005 level)**	Not Significant
Adaptor-Innovator	Not significant	Not significant	Not Significant	Not Significant

4.1.13 Design Recommendations

In catering to field dependent individuals, environments can be designed to provide appropriate cueing techniques and helpful navigational routes through the learning material. Perhaps in a well-designed system the individual learner may be able to re-arrange the material that best suits his or her Cognitive Style, thus maximizing the potential of the medium. As mentioned already, the Web is predominantly a text-based medium, best suited to Verbalisers and probably Field-independents. The Internet has the potential to provide Imagers with highly interactive graphical, audio, visual learning materials despite current difficulties with technical implementation and bandwidth.

"What I see as the real contribution of digital media to education is a flexibility that could allow every individual to discover their own personal paths to learning. This will make it possible for the dream of every progressive educator to come true: In the learning environment of the future every learner will be special" (Papert (1996).

4.1.14. Conclusions of Study 4.1

(A) Cognitive Styles do affect learner performance in different computer media.

(B) Witkin's Field dependence–Field independence is the most consistent predictor of final score irrespective of treatment.

(C) Witkin's Field dependence–Field independence is the most consistent predictor of final score in the Internet environment.

(D) Riding's Wholist-Analytic cognitive style was shown to be a significant predictor of final score in the Cd-Rom environment.

In conclusion, properly-designed, and well-motivated, Internet-based learning environments have tremendous potential and much to offer from an andragogical point of view if proper account is taken of the individual's Cognitive Style.

A more detailled discussion of this Study is given in Parkinson & Redmond (2002b).

4.2 The Matching / Mismatching Hypothesis

The matching hypothesis has its advocates and opponents. Its assumption is that when there is a match between an interface, learning materials and an individual difference such as cognitive style successful outcomes are more likely. The hypothesis has been studied predominantly in educational contexts. There have been other instances where the matching hypothesis has been examined. Armstrong (1999) for instance investigated this hypothesis while studying relationships between supervisors and their students.

4.2.1 The Evolution of the Matching / Mismatching Hypothesis

Allinson and Hayes (1996) review some of the literature on the matching hypothesis. They state, that "Cognitive style can have important implications for task design and "might have implications for team composition and conflict management". They cite Kirton (1980, 1989) who suggests, that "innovators perceive their work environment as more turbulent than adaptors who, in turn regard innovators as disturbers of the peace". These differences have obvious consequences for team development. They state in this article that "Cognitive styles might also have important implications for training and development "and "there is a widely held shared view that people will learn more effectively when the learning environment matches their cognitive style". They refer to work undertaken by Ash (1986) and Halpin & Peterson (1986) who substantiated this claim. They conclude that "more research is required to replicate and confirm the results of existing studies..." and "It is, however a concept which has the potential to make an important contribution to management practice".

Many studies have investigated the accommodation of individual differences in various types of system design. It is critical to realize that these studies have been based on the matching hypothesis. An important concern is the validity and the reliability of this matching premise. Unfortunately in many cases these studies have not been replicated so it is is difficult to establish the truth or otherwise of the matching hypothesis.

4.2.2 Early Discussion of the Matching Hypothesis

Witkin, Moore, Goodenough and Cox (1977) explore matching cognitive style and the teacher pupil relationship. They believed that the effect of match/mismatching of cognitive style, reflected in behaviour; underpinning the teacher pupil interaction may have important effect/ consequences in classroom learning.

Saracho & Dayton (1980) and Frank & Davis (1982) reported that cognitive style matching had no effect in their studies. The conclusion from these studies was that "a cognitive style match may have been beneficial for field independent students but a cognitive style mismatch was beneficial for Field dependent students. Kagan and Buriel's (1977) study collaborates these findings.

4.2.3 Riding and the Matching Hypothesis

Previous research completed by Riding and Taylor (1976), Riding and Calvey (1981), Riding and Anstey (1982) found that "Imagers find concrete and readily visualized information easier than semantically and acoustically complex details with the reverse applying to Verbaliser" in Riding and Sadler-Smith (1992).

Riding and Ashmore (1980) studied whether the mode of presentation of information interacted with cognitive style and had an effect on learning. They examined the Verbal – Imagery cognitive style as assessed by the CSA in a sample of seventy-four eleven year old school children. They established that "there was a significant interaction ($p < 0.05$) between the Verbaliser - Imager learning style and mode of presentation in their effect on the number of questions correctly answered."

Riding and Sadler-Smith (1992) conclude, on the basis of their study, "that the identification of cognitive styles is of immediate

relevance to teachers and trainers since it can be used to (a) predict learning difficulties, (b) inform the discussion of the learning process amongst professionals (c) enlighten interpretations of learner evaluations of instructional programmes, (d) make possible the design of instructional treatments which may be congruent with an individuals habitual modes of thinking and (e) improve the effectiveness and efficiency of instruction".

Riding & Calvey (1993) suggest for example that Imagers as assessed by the CSA recall visually presented textual material better than complex and unfamiliar text while the converse is true for Verbaliser.

Riding and Rayner (1995) propose an outline model of a learning system with the World Wide Web (WWW) in mind. They argue that such a model has to be psychologically based with cognitive style being one of the key elements in its design. They state, "Trainees taught with materials and methods which do not match their learning style will therefore find difficulty in learning". They also state, that "the style of the verbaliser would appear to be more appropriate to coping successfully with learning from text and definitions, than that of the imager". They suggest in the design of learning control system that the verbaliser should be supported by "verbal versions of pictorial and diagrammatic material". The Imager should be supported in the following way: verbal material should be converted into "pictorial form and "choose concrete analogies "instead of abstract ideas". The authors mention the need for the development of a model of individualized learning. This they say "will facilitate the further development of the intelligent Agents and Control Systems".

Riding and Read (1996) investigated cognitive style and pupil learning preferences in a sample of seventy-eight twelve-year old pupils. They administered the CSA, which determined the cognitive style position of each student on the W-A and V-I dimension. One of the authors' main findings was that Imagers tended to prefer to use images in comparison with Verbalisers who preferred writing. This study was limited by the fact the study asked the students questions which involved subjective ratings of themselves. Limitations of the experimental approach

in this study included: lack of reliability factors being reported for the questionnaire used and difficulty in establishing the relationship between response given and actual response. There appeared to be little evidence of a control for false responses. One might expect greater tendencies on the part of the sample to comply and appease the adults who were conducting the study. There was no attempt to correct for error between the two people rating raters: the study isn't explicit in how all of the information was gathered. This study might have been enhanced if the questionnaire had been administered to the students for them to do, and then if the students had been given some practical tasks and observed and maybe interviewed afterwards.

Riding and his co-researchers have examined the effect of cognitive style on preferred format of instructional material. Riding and Watts (1997) investigated the cognitive style and preferred format of handout amongst a sample of ninety fifteen to sixteen year old female pupils. These students were asked to choose a handout, which was presented in three modes: Unstructured – Verbal, Structured - Verbal and Structured - Pictorial. Cognitive Style was assessed using the CSA. They report that "there was a significant effect of the Verbal - Imagery dimension ($p = 0.002$) with the majority of Verbalisers choosing the Structured – Verbal sheet and Imagers opting for the Structured – Pictorial version. The authors of this study report "that pupils have clear preferences, in terms of their cognitive styles for certain formats of instructional materials". They also suggest, that "the appearance of instructional materials have important implications for the attractiveness of a subject to the pupil and its consequent effect on motivation".

Peterson (2003) failed to replicate these findings but says that the experimental conditions may have been at variance to what was reported by these authors. Peterson's sample size was also different consisting of fifty psychology students probably less naïve than the sample used in Riding's study.

4.2.4 Recent Discussion of the Matching Hypothesis

McKenna (1990) reviews some of the earlier literature on the matching/mismatching hypothesis. Some of the earlier research concentrated on using Witkin's concept of Field dependence /Field independence. McKenna (1990) traces the origins of this

hypothesis back as far as Witkin, Lewis and Weil's (1968) work on the patient – therapist relationship. During the 1970's it was hypothesized that Mexican – American students performed poorly in educational institutions, as there was a mismatch between teacher and student (Ramirez (1973)). This was substantiated by Packer and Bain (1978).

McKenna (1990) in this review concludes that "the evidence for cognitive style matching is very poor" and claims that "at present therefore, no convincing demonstration of the matching hypothesis has been published".

Liu and Reed (1994) have suggested that understanding the variables associated with successfully matching cognitive styles to instruction is important in the instructional design process. It is implied that instructional treatments supported by design features related to cognitive styles may prove to have a significant impact on learning achievement.

Hayes & Allinson (1996) reviewed twenty studies on the matching hypothesis. They found that thirteen of these studies suggested that matching learning styles with "learning activity would have a positive effect on learning performance".

Pillay (1998) reports on a study investigating the effect of Riding's Cognitive Style dimensions. She suggests that her results "indicate a potential for cognitive style to influence learning outcomes measured by performance on test tasks."

Saracho (2000) defines a match as "when an agreement on certain characteristics exists between two individuals or an individual and his/her environment", and refers to Witkin (1965) who proposed the matching hypothesis initially in the context of student and teacher interactions. Witkin believed that when students "are matched with their preferred instructional mode, achievement and satisfaction with learning will be enhanced".

Armstrong (2000) believes that "mismatching on cognitive style is more likely to result in conflict as differences in style yield differences in interests, values and problems solving techniques that may handicap a working relationship" (in Riding and Rayner (2000)).

Saracho (2000) in her review of the matching/ mismatching hypothesis makes a few relevant points: "insufficient studies support the importance of identical matching in cognitive style" and she believes "that research evidence suggests that the matching hypothesis is still ambiguous". She also considers that mismatching cognitive style may have important consequences for the learning process: "the mismatch of cognitive style may be considered a performance matching style, because it offers enrichment, challenge and stimulation for both teachers and students". In substantiating this claim she refers to some of her previously published work (Saracho (1993)), which demonstrated that "FI teachers acquire higher achievement scores than those students with FD teachers".

Ford and Chen (2001) claim that matching cognitive styles with instructional presentation strategies may have important potential in enhancing student learning.

The Coffield Report (2004) refers to Ford's empirical studies on the topic of matching. Ford's earlier studies (1985, 1995) and Ford & Chen (2001) demonstrated that improved performance was linked to matching cognitive style. Coffield et al. (2004) also refer to Ford & Chen's 2001 paper, which states that the "effects of matching and mismatching may not be simple and may entail complex interactions with other factors such as gender and different forms of learning". The authors of the Coffield Report also believe that another neglected factor is subject matter. The Coffield Report gives a comprehensive review of the Matching Hypothesis in Section 8 and concludes that the evidence for it is "at best equivocal and deeply contradictory at worst".

Armstrong (1999) came to similar conclusions to the Coffield Report after a comprehensive review of the Matching Hypothesis and states: "However, the dearth of empirical studies into the actual effects of matching/mis-matching cognitive styles on interpersonal relationships, coupled with the fact that where there is reported evidence of effects, there is often confusion as to whether these are concerned with performance outcomes or interpersonal behaviour, or with both of these, highlights the need for a clearer and more systematic research programme".

4.2.5 Cognitive Style and Human Computer Interface Design

Cognitive style has been considered as a factor in the design of the interface and in the design of Information and Management Systems. Some researchers have questioned the usefulness of computer technology in learning (Clark (1985), Kohli (1995)).

Zmud (1979) reviewed the literature that examined the affect of individual differences on management information design. While accepting that the review is now quite dated, the authors concluded, that "individual differences do exert a major force in determining MIS success". They conclude, "that potential payoff from further research investigating the relationships between MIS success and the personal characteristics of MIS users is high".

Zmud (1979) views are in broad agreement with Bariff and Lusk (1977). They suggest, "that the measurement and evaluation of user's cognitive style and related personality traits may provide an effective means for attaining successful MIS modifications".

Huber (1983) reviews the literature that underpins the assumption that a user's cognitive style should be a factor when designing Management Information Systems and Decision Support Systems. He refers to numerous studies, which reinforced this belief, for example Benbasat and Taylor (1978) and Driver & Mock (1975). This review critically examines the literature and concludes, that "the current available literature on cognitive styles is an unsatisfactory basis for deriving operational guidelines for MIS designs and DSS designs". His justification in this instance is (a) " inadequately developed theory of cognitive style, (b) multitude of measuring instruments with inadequately established properties and (c) faulty research designs used in empirical investigations of cognitive styles".

Huber (1983) also argues further research studies in the area of "cognitive styles are unlikely to lead to operational guidelines for MIS and DSS designs". The reasons why he believes this are, firstly, there is a multitude of other individual differences that may be relevant to the design of these systems. Secondly, there is the issue of the usefulness and credibility of cognitive style as a factor in DSS design from a managerial perspective. Thirdly, the development of "empirically based bodies of knowledge are very slow"and, perhaps before they emerge, technology may be

able to provide very flexible systems that may surpass the current systems. Finally he wonders, even if it were possible to create such systems, would this be a good idea anyhow? He argues that this would lead to reinforcing idiosyncratic tendencies, which may not be to the advantage of the organization, as such a design would not facilitate alternative cognitive styles. This arises if others have to use the DSS or MIS. Huber's view is perhaps limited as it pertains to the area of Decision Support Systems and may not be applicable to the wider domain.

There appears to be a divergence in the literature as to the importance and effectiveness of the matching hypothesis. Bostrom, Olfman and Sein (1988) report on a series of studies which examined novices' learning style whilst learning electronic and spreadsheet packages. They concluded that learning style is an important predictor of learning performance and that it also interacted with training methods. They suggest, that "the four studies reported in the paper demonstrate the importance of individual difference variables in the learning of computer systems". They also concluded, that training methods need to be customized for the individual, as the same size doesn't fit all. Some of the studies reported upon are limited by very small sample sizes, a point which is acknowledged by the authors in this paper.

Other researchers have found that when computer based instruction matches the individual's characteristics, learning can be quite effective (Kozma (1991), Lynch (1991), Sternberg (1997), Atkinson (1998)). Allinson and Hayes (1998) report that when technology and cognitive styles are compatible that individuals are in a favorable position to learn effectively. Some researchers argue that individuals may disregard information if it is presented to them in a manner that doesn't match their cognitive style (Churchman and Schanblatt (1965)).

Spence and Tsai (1997) investigated cognitive style as a factor in the explanation of human interaction with computer systems. In assessing the literature in the field they suggest that the findings concerning the feasibility of cognitive style as a guideline for IS research have proved inconclusive. Spence and Tsai (1997) found that cognitive style does not affect a user's action in a specific

task environment. They also found that individuals with varying cognitive styles did not perform any differently. They conclude "consequently, further research aimed at differentiating decision behaviour with respect to task performance that is based on cognitive assessment would be fruitless".

Boles and Pillay (1999) investigated cognitive style using Riding's CSA, subject content and the design of computer based instruction. The authors report that there were no significant differences between the matched and mismatched groups on total test score. What was interesting though was the fact that in the matched conditions the task was completed quicker when compared with the mismatched treatment. The authors, whilst not explaining this finding in any great depth, said, "it may be plausible to suggest that there appears to be a pattern (although not significant) in favor of the matched group".

McKay (1999) investigated the effect of graphical metaphors on the performance of learning computer programming concepts. They report that Verbalisers tend to learn new programming concepts better if presented with textual plus graphical metaphor format than with just text plus textual metaphor format. These results confirm earlier findings reported by McKay (1999). Again, these results are in conflict with those reported by Riding and Ashmore (1993).

John and Boucouvalas (1999a) compared user performance with interfaces designed to suit testee's cognitive styles and stated "the results suggest that the style of presentation did not affect the subjects' performance of the task and subjects performed consistently on all tasks whatever cognitive style the task was designed to suit".

In another study (John and Boucouvalas (1999b)), they concluded, that "in a majority of tasks cognitive style did not have a significant influence on the performance of subjects". However they do report that where cognitive style was found to have a small statistically significant effect, it was in the expected direction for Verbaliser but not for Imagers. In this study the individuals were asked questions in both textual and image forms.

They also suggest that that task may have had "more of an influence on the individuals' performance". This study also compared user cognitive style as assessed by Riding's CSA and performance with text user interfaces and image user interfaces. This result contrasts with research reported by Riding and Douglas (1993) and Riding and Ashmore (1980). Chou (2001) reports that cognitive style should be taken into account in the training of individuals on IT- related courses.

Torenvliet, Jameson and Vincente (2000) found that cognitive style was a constant predictor of successful individual performance when using an Ecological Interface Design (EID). They report, that "the strongest and most consistent predictor of performance was the interaction between a holistic cognitive style score and interface based on the principles of ecological interface design".

The individuals in this study were using an EID interface to assess fault diagnosis. The measures included fault detections, fault detection time, diagnosis accuracy and diagnosis time. Pask and Scott's (1972) spy ring history test was used to assess cognitive style in this instance. The sample size was effectively forty five. The study is somewhat limited given the number of other predictor variables (gender, educational level, educational relevance, training, goal tolerance, and fault order).

Piombo, Batatia and Ayache (2003) proposed an adaptive system to match the student's cognitive style. The instruction adaptation module was designed so that it matched the student's cognitive style. They propose a Bayesian model consisting of a content, and student, model. However they do not provide any empirical evidence as to its subsequent performance.

Triantafillou, Pomportsis, Demetriadis (2003) report that "an Adaptive Hypermedia system can be designed to accommodate a variety of individual differences including cognitive style". Similar to adaptive systems already described in Chapter 2, this system consisted of a domain, student module and an adaptation module. The adaptation module presented information to the user in a manner that matched his/her cognitive style. The goal of the system design was to improve the performance of the field dependent learner by provision of a

concept map and graphic path indicator. The individuals also had the opportunity to switch between program and individual control. The system was designed to provide very clear and obvious instruction to the field dependent learner and the bare minimum to the field independent learner. The results indicated that the differences "between means of the pretest and post test showed that FD students were improved more than non-FD students". The Fd students are reported as stating that the "concept map and graphic path indicator were very useful in order to organize the structure of the knowledge domain". Again the results ought to be treated with caution given the very small sample size (n = 10). This study demonstrates that it is possible to design an adaptive system to facilitate cognitive style. This could also have been successfully achieved by adopting a non-fixed design approach as was demonstrated by Parkinson and Redmond (2003), Redmond, Walsh & Parkinson (2003), Parkinson, Redmond & Walsh (2004a) and Parkinson & Redmond (2005).

Workman (2004) reports that cognitive styles do make a significant difference in computer based education delivered via CD – ROM and computer based education delivered over the World Wide Web.

4.3 Study 4.2 Can both ends of a cognitive style continuum be supported to give equal performance in a HCI?

Field-independent learners performed significantly better than Field-dependent learners in a Web-based environment in the previous pilot study - Study 4.1. This leads to the question as to whether the disparity in performance can be reduced between the field-dependent and field-independent.

Consequently the Web user interface for this preliminary study was modified to accommodate the Field-dependent learner to try to reduce the disparity in performance with Field-independents in two studies - Study 4.2a and Study 4.2b. Seventy-two Information Systems students were initially assessed for Field-dependence/Field-independence cognitive style using the Group Embedded Figures Test (GEFT) (Witkin, Oltman, Raskin & Karp (1971). Thirty-four of these students studied Modules 1 - 9 of an Introductory Course in Artificial Intelligence in a Web environment for one hour per week for six weeks (Study 4.2a).

They were then assessed in a one-hour exam. Similarly the remainder of these students studied Modules 10 - 18 under identical conditions and assessment (Study 4.2b). Here the performance of field-dependent learners in a web-based environment, which has been adapted for them, is examined.

4.3.1 Support for the Field-dependent Learner in a Web-based Environment

Ford and Chen (2001) claim that matching cognitive styles with instructional presentation strategies may have important potential in enhancing student learning. Liu and Reed (1994) have suggested that understanding the variables associated with successfully matching cognitive styles to instruction is important in the instructional design process. It is implied that instructional treatments supported by design features related to cognitive styles may prove to have a significant impact on learning achievement.

Studies conducted by McLeod, Carpenter, McCormack and Romuladas (1978), Adams & McLeod (1979), Wilborn (1981) and Hansen (1983) confirm that Fd's perform worse where the learning environment is unstructured. Jonassen and Grabowski's (1993) brief overview of Field-dependent/Field-independent research carried out in traditional classroom environments supports this earlier research.

Examples of this are numerous attempts to model and match cognitive style characteristics in traditional and computer-based presentations of learning material (Riding & Sadler-Smith (1992)), (Hayes & Allinson (1994)), (Riding & Watts (1997)), (Ford & Chen (2001)). There is, however, an alternative view which suggests that exposure to mismatched learning environments can help the learner develop a wider range of coping behaviours or learning strategies (Hayes & Allinson (1994)).

The purpose in this study was to introduce the field of Artificial Intelligence to student groups, not to directly develop their coping behaviour or learning strategies. Consequently it was decided to accommodate the Field-dependent learner in this instance using the matching hypothesis paradigm (Vid. later in this Chapter) by designing the web interface to match the Field dependent user's cognitive style for better examination performance.

Chen (2002) in a review of quite a number of recent studies of Field dependence/ Field independence comments that most of the "... studies revealed that Field Independent individuals performed better than Field Dependent individuals in hypermedia learning".

Ayersman and Minden (1995) suggest that Hypermedia have the capacity to accommodate cognitive style differences. Learning effectively in a web environment demands that the individual

- Develop successful navigational strategies,
- Be accustomed to learning on his/her own and
- Have a high internal locus of control.

These characteristics best describe Field-independent learners thereby explaining why they are at an advantage while learning in such an environment (Table 4.1) and this is what was found in the pilot study - Study 4.1 (Parkinson & Redmond (2002b)).

Table 4.1. Differences between Field-dependent and Field-independent learners (adapted from Jonassen & Grabowski (1993)

Field-dependent	Field-independent
Accepts structure	Imposes structure
Wholistic / Global	Analytic
Influenced by format / structure	Less influenced by format/ structure
Externally directed	Internally directed
Affected by stress	Ignores external stress
Accepts ideas as presented	Represents concepts through analysis
Sociable & gregarious	Individualistic
Extraction of elements from field is difficult	Easy extraction of elements from field

The Field-dependent learner finds it difficult to extract pertinent pieces of information from a hypermedia environment. Korthauer and Koubek (1994) found that experienced Field-dependents performance was less than that of experienced Field-independents particularly when explicit structure was not provided. The Fd learner needs to be externally directed through the use of cueing (Satterly & Telfer 1979) (Ford & Chen 2001). Kim (2001) found that, in a Web based online searching task, Fd students preferred a well-structured format, as Field-dependents do not easily impose a structure themselves. Chen and Macredie (2002) provide an extensive review of studies that have investigated the performance of the Fd/Fi individuals in Web-based environments. It may be concluded that Field-dependent

learners may not be ideally suited to learning in typical web-based environments.

4.3.2 Motivation for Study 4.2

In the previously completed pilot study (Study 4.1) where learners' individual differences such as cognitive styles were not explicitly taken into account, it was found that Field-independent learners significantly outperformed Field-dependent learners.

The pilot study website had been designed with care, influenced somewhat by the recommendations of Shneiderman (1998). A screenshot of a typical website page is given below in Fig. 4.1. The results of Study 4.1 had shown that a typical web environment suited Field-independents (Vid Table 4.1.A in Appendix 3-4.1.A). These results from the preliminary study (Study 4.1) showed clearly that the more Field-independent a respondent was, the higher the exam score. This can be clearly seen when the simple linear regression model is examined in Table 1 in Appendix 3-4.1.A.

This outcome raised the question as to whether the Web site interface could be redesigned in order to raise the Fd performance so there would be little difference in performance between Field-dependents and Field-independents while not hampering Field-independents. To explore this idea it is necessary to look in more detail at how the characteristics of the Field-dependent as discussed above might be supported.

4.3.3 Method for Study 4.2

Participants: The students, being third year Information Systems students, were all considered computer-literate and were experienced in using the Internet. The students had not studied Artificial Intelligence before. The Geft was used to measure field dependence/ independence instead of Riding's W-A dimension because it had been widely used and reported upon in the literature, and also because it had good reliability values (Witkin, Oltman, Raskin & Karp (1971)).

4.3.4 Study 4.2 Design

The course content for all three studies (preliminary pilot study (Study 4.1) and Studies 4.2a & 4.2b) was an Introductory Course

in Artificial Intelligence. Study 4.1 was the original study that raised the question: can a user interface be adapted so as to accommodate both Field-dependent learner, while not disadvantaging the Field-independent learner?

In Study 4.1 (the pilot study), the final year students volunteered to participate in an Introductory Course in Artificial Intelligence, which was supplementary to the main AI course, scheduled for later on in the term. As a consequence of this study, an Introduction to AI course was fitted into the structure of the overall degree course but was scheduled for two years earlier. This new course was the basis for Study 4.2a and Study 4.2b.

4.3.5 Study 4.2a

A group of 34 third year Information Systems students studied Modules 1 - 9 of an Introductory Course in Artificial Intelligence in a Web-based environment for 6 weeks, one hour per week (Study 1). Their mean age was 27.9 years (median 27) with s.d. = 4.4. The website was only available to the students for a one-hour window each week which corresponded to a one hour allocated lecture slot. At the end of 6 weeks their learning was assessed by a one-hour exam (again given during the allocated slot).

4.3.6 Study 4.2b

A second group of third year Information Systems students (n = 38) similarly studied Modules 10 - 18 under identical conditions (Study 2). The mean age was 29.1 years (median 29) with s.d. = 4.0. (These students had already covered Modules 1 - 9 in a similar fashion to those in Study 4.2a but using hard-copy Text only). The content in Modules 10 - 18 was farther along in the Course and therefore somewhat more advanced. Each Module was relatively independent and distinct.

4.3.7 Measures

The students were initially assessed using the Group Embedded Figures Test (GEFT) as discussed below (Witkin, Oltman, Raskin & Karp (1971)). The basic measure of student performance in the Introduction to Artificial Intelligence course was the exam score. After the students had been assessed using the GEFT, they were randomly allocated to either Study 4.2a or Study 4.2b.

4.3.8 Psychometric Test

All participants were administered the GEFT (Group Embedded Figures Test) (Witkin, Oltman, Raskin & Karp (1971)) beforehand (vid. Appendix 4-4.2.B for the summary statistics on GEFT for these Studies). This test is a twenty five-item test, (seven items of which are practice items and are not counted), which contains three time sections of 2, 2 and 5 minutes, respectively. The test comes in booklet form and the individual is required to trace one of eight simple figures embedded in a visual background of greater complexity. The GEFT is a measure of Field-independence/ Field-dependence. The measurement scale extends from zero to eighteen; the better the score, the more field-independent the person is. Very acceptable reliability coefficients have been reported for Witkin's GEFT (Witkin, Oltman, Raskin & Karp (1971)).

4.3.9 Apparatus

Both studies were run on the same desktop PC machine types with identical CPU speed and memory and all students accessed the same website. The data were analysed using a simple linear regression model via Data Desk V6 (2001) software.

4.3.10 Stimulus Materials for Study 4.2

The website with the course content provided the stimulus material.

4.3.11 Procedure for Study 4.2

At the beginning of each session the students logged on to the website using a password. A specially designed piece of software designed using CGI locked the users out of the site after exactly one hour. The one-hour time slot was dictated by the academic timetable. The students were encouraged to take notes if they so wished while studying the course material. They were not permitted to print any of the course material.

4.3.12 Learning Performance - Evaluation

All exams were set at about the same relative level of difficulty while the exam format for Studies 4.2a and 4.2b differed from that of Study 4.1 (vid. Appendix 5-4.3.C for the summary statistics on the examination scores for the Studies). Study 4.1 had been examined by a multiple-choice questionnaire. The students' knowledge of the material was examined on both

occasions at the end of 6 hours of instruction (one hour per week) in a test that lasted one hour. The assessment test for the students contained twenty-two objective style questions and was marked out of 22. These questions were objective style questions and assessed the students' knowledge and understanding of basic concepts in AI (e.g. Turing test, expert system development and application) at both lower (fact recall) and higher (synthesis and application) cognitive levels. Examples of the 22 objective style questions were: "List 4 disadvantages of the use of robots, "Compare and contrast an intelligent machine with a human being", "Write a short note on Alan Turing".

When Studies 4.2a and 4.2b were each completed, the students were given a presentation by the authors on the results and their implications. Students were then presented with their own individual results and any resulting queries were addressed.

4.3.13 Website Design

An educational Internet site was adapted from the pilot system website using Java and Common Gateway Interface (CGI) scripts. The site design was intended to support the characteristics of Field-dependent learners as discussed above (Vid. Fig. 4.1).

Based on the literature reviewed above the course content and interface as used in the pilot study were adapted as follows. The Field-dependent learner was supported principally by a site map. Present also, as in Study 4.1, were the following: internal links appeared on every page consistently and clearly spaced, examples being a link to the discussion forum, a site search engine, home page and help menu. In addition, to further help avoid the Field-dependent getting lost on the Website, scrolling was reduced wherever possible. The site map (tree hierarchy) was developed using free software available from Joust (1999). The three-level map (arranged as a hierarchical tree) had three tiers as follows:

- an overview of module titles
- if a module title was clicked on it expanded to give a view of the sub-module titles for that module and also changed colour
- if the user clicked on a sub-module title he/she was brought to a leaf node which contained some course content.

The Field-independent learner also had the opportunity to explore the site in a non-linear way mainly facilitated by the use of a site index, and the use of the interspersed hypertext links and links to the WWW. An example of the Website adapted for Field-dependent learners is shown in Fig 4.1:

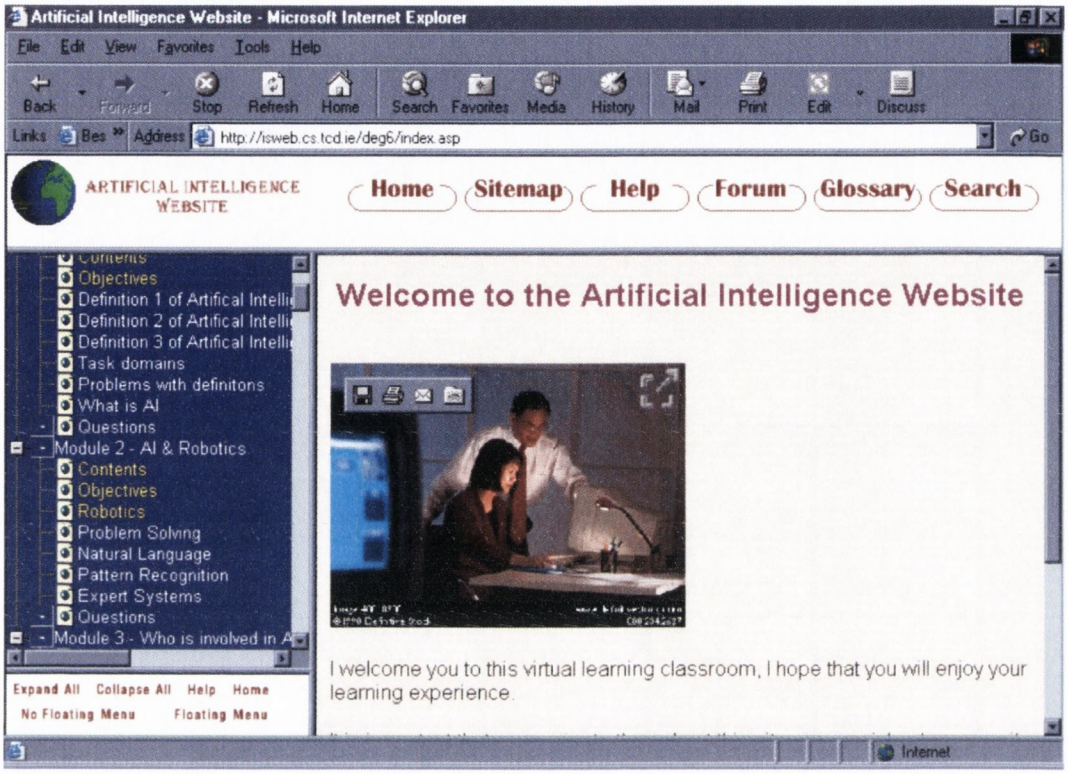


Figure 4.1: Sample screenshot of the user interface

A user can browse the site and go to any of the 18 modules given in the site plan. The user is then ready to access any of the modules listed on the left hand side of the above screen. By clicking on the "+" symbol, the directory structure is expanded and the user sees all the contents of that module. He can then go to visit whichever section he is interested in. The information from each particular section is displayed on the right hand portion of the screen and the user is often given the opportunity to visit external links.

4.3.14 Results of Study 4.2

Further statistics for Study 4.2 are given in Appendix 4-4.2.B and Appendix 5-4.3.C.

Study 4.2a (Modules 1 - 9)

The data were analyzed using Data Desk V6.1 (2001) and are shown in Table 2. The results show that the Fd / Fi Study 4.2a

score was not statistically significant at the p = 0.05 level (since p = 0.73 for Fd/Fi Study 4.2a). This is confirmed by an R squared of 0.4% which shows that less than 1% of the variability in the model is accounted for by Fd/Fi.

Table 4.2.1. Simple Linear Regression Model of Exam Score vs. Fd/Fi in Study 4.2a (n = 34)

Dependent variable is: **Exam1**
No Selector
R squared = 0.4% R squared (adjusted) = -2.7%
s = 3.475 with 34 - 2 = 32 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	1.45304	1	1.45304	0.120
Residual	386.437	32	12.0761	

Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	14.3026	1.312	10.9	≤ 0.0001
Fd/Fi 1	-0.042552	0.1227	-0.347	0.7310

Study 4.2b (Modules 10 - 18)

The results, given in Table 4.2.2, show that the Fd/Fi Study 4.2b score was not statistically significant at the p = 0.05 level (since p = 0.23 for Fd/Fi Study 4.2b). This is confirmed by an R squared of 3.9% which shows that less than 4% of the variability in the model is accounted for by Fd/Fi.

Table 4.2.2 Simple Linear Regression Model of Exam Score vs. Fd/Fi in Study 4.2b (n = 38)

Dependent variable is: **Exam2**
No Selector
R squared = 3.9% R squared (adjusted) = 1.2%
s = 3.408 with 38 - 2 = 36 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	17.0465	1	17.0465	1.47
Residual	418.197	36	11.6166	

Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	11.0675	1.315	8.41	≤ 0.0001
Fd/Fi 2	0.141588	0.1169	1.21	0.2336

Consequently field dependence or independence is not a factor in the learning performance of the students using this adapted web site for two different groups of sample modules for two different groups of students. In other words, the adaptation of field-dependents has been successful.

4.3.15 Discussion of Studies 4.2a and 4.2b

The results for Studies 4.2a and 4.2b show that Witkin's Field-dependence/ Field-independence is not a significant predictor of student examination score when the Web interface has been adapted to cater for the Field-dependent. This shows that our redesign of the environment so as to accommodate the situation

for Field-dependents has been successful. In other words, those who were Field-dependent performed similarly to those who were Field-independent. In the previous study (Study 4.1) Field-dependents had performed significantly worse than Field-independents where the learning environments had little support for the attributes of the Field-dependent learner (Parkinson & Redmond (2002b)).

The literature clearly indicates that one of the characteristics of the Field-dependent learner is an inability to distinguish salient cues from a complex field. The map provided the Fd learner with links to each module and its sub-modules. This assisted the Fd learner by highlighting the type of content in each module and sub-module resulting in more efficient learning for the Field-dependent individual.

In Studies 4.2a and 4.2b, Fd learners were supported by means of a map, which essentially provided feedback as regards to their position in the learning environment and perhaps facilitated the more efficient generation of their own mental model of the learning environment. This may in turn have assisted the field-dependent learner in the more efficient re-organization of their own learning schema each time they studied new learning material. It is generally accepted that Fd's are also externally directed so the provision of a map in this instance would also have been doubly beneficial (cf. Table 4.1). The map also gives the ability to reverse a navigational decision and is the basis for building a good mental navigational model of the website.

The map may have reduced the tendency for the Field-dependent individual to become lost in hyperspace (Chen and Macredie (2002)). Dvorak and Sommerville (1996) report "being disorientated or lost is one of the fundamental difficulties which users experience when trying to navigate within hypertext systems". The users end up not knowing where they are in the overall structure of the document. Otter and Johnson (2001) state "... orientation cues which are so important ... are absent in many hypertext systems". They also confirm the assumption that when an individual is lost, worthwhile learning doesn't take place in a web-based environment.

This study thus demonstrates that it is possible to accommodate a range of individual differences using a single interface. It was accomplished by enriching the environment as opposed to restricting the ways in which users interacted with the system. Tan and Lo (1991) report that "if the interface can be customized to suit the idiosyncrasies of the user, the effect that the cognitive style of the user has on the system can be reduced if not removed".

These Studies clearly demonstrate the importance of good design and what can happen to some types of learner when the design isn't particularly sensitive to their needs (Study 4.1 vs. Studies 4.2a and 4.2b). In other words when those who are field-dependent aren't taken into account the subsequent negative effect on their learning has been clearly demonstrated (albeit in reverse). Where mission critical interfaces are being designed, e.g. air traffic control, interfaces used by surgeons during neurological procedures, manual instrument control, accommodating individual differences such as cognitive style becomes more crucial.

The adaptation helped the field-dependents but did not disadvantage the field-independents because the field-independent learners had a choice as to whether to use the expandable map or not. They had opportunity to navigate through the learning material in an unstructured or structured fashion, as desired, in accordance with his/her cognitive style preference using the website internal search engine, hyperlinks and the map if desired.

Other alternatives to support the field-dependent learner might include the use of audio cueing from within the virtual environment; the use of a help/support chatroom aimed at field-dependent users and more use of summaries.

4.3.16 Conclusions of Studies 4.2a and 4.2.b

These studies demonstrate the importance of identifying and accommodating individual differences such as field-dependence/field-independence:

• It has been shown that it is possible to redesign an eLearning environment - a Web site - so as to balance the performance of

Field-dependents and Field-independents. This has been replicated for different groups of modules and students.

- It has been found that Field-dependent learners performed similarly to Field-independent learners on the adapted Website.
- Web-based content structure can be adapted to accommodate both types of learner. The adaptations used have been discussed. This has implications for Web designers and educational psychologists who may be specifying such designs.

The main adaptations of the Website and course content were:

- A three level map which proved to be a good navigational aid and a basis for the construction of a good mental model of the structure of the course material.
- The reduction of scrolling to a minimum or eliminated where possible.

These results are reliable in that a Web environment (which was initially biased towards Field-independents) has been adapted so that it became equally useful for both Field-dependents and Field-independents measured over two different sets of students. Further details are given in Redmond, Walsh & Parkinson (2003), Parkinson, Redmond & Walsh (2004a) and Parkinson & Redmond (2005).

4.4 Introduction to Study 4.3: Whether Verbalisers - Imagers could be accommodated by presenting a matched set of pages depending on their CSA scores?

The results from Redmond, Walsh & Parkinson (2003), Parkinson, Redmond & Walsh (2004a) and Parkinson & Redmond (2005) confirmed that a non-adaptive (static) system design could be implemented to ameliorate the deficiencies of the field dependent characteristics. This leads to the question as to whether matching a user's cognitive style with the interface would result in similar performance for imagers and verbalisers.

What is examined in this study is whether a non-adaptive design using cognitive style scores could be implemented effectively. What was envisaged in this system design was inputting an individuals cognitive style scores obtained from the CSA and then presenting a set of web pages to match the particular cognitive style characteristic.

A set of web interfaces was designed to accommodate an individual's cognitive style in a preliminary study. Fourteen third year university students were tested for cognitive style using Riding's Cognitive Style Analysis (CSA) which assessed whether a student was a Verbaliser or Imager or in-between (Bimodal). The students were given an Introductory Course on Artificial Intelligence (AI) on the Web; they were presented with the course material in a web-based format matched to best suit each individual's cognitive style score: Verbaliser, Imager or Bimodal. The adaptation was achieved by giving the students three alternative modes from textual to highly graphical depending on their cognitive style score. In previous research (Redmond, Walsh & Parkinson (2003)), where little consideration had been given to the V-I dimension, it had been found that Verbalisers performed better than Imagers. Verbalisers and Imagers did not perform in this new assessment test any differently from each other, which indicated that the adaptation had been successful. Hence it can be concluded that the disadvantages of the Web being largely textual in character can be overcome by explicit user interface design considerations.

Shneiderman (1998) comments "A clear understanding of personality and cognitive styles can be helpful in designing systems for a specific community of users". Where an interface can be improved for any one individual difference, efficiencies in performance, behaviour, and usability can also be much improved.

The purpose was to establish in a pilot study whether Verbalisers and Imagers would perform equally well when presented with Web-based learning environments that matched their cognitive styles.

During the last fifteen years, Riding's research has shown that Verbalisers tend to perform better than Imagers in learning environments where the material is in a textual and/or auditory format (Riding (1998)). Imagers similarly tend to outperform Verbalisers where the presentation mode is graphical and visual. Douglas and Riding (1993) reported that Verbalisers performed better than Imagers in a text-based environment. Imagers scored higher than Verbalisers in a similar type environment where the learning material was in text and picture format. The Web, being

largely a textual medium, has been shown to favour Verbalisers. In previous research where no allowance was made for the cognitive style of Verbaliser or Imager, Verbalisers were shown to perform better in examination performance in a Web-based environment because of the largely textual basis of the content (Redmond, Walsh & Parkinson (2003)). (This discussion must be treated with some caution given the results of the V-I dimension of the CSA to be discussed fully in Chapter 5).

4.4.1 Method for Study 4.3

Procedure

The students' cognitive style scores were assessed using Riding's CSA and used to categorize each student as being either a Verbaliser, Imager or Bimodal (in between) as determined by Riding's style labels (Riding (1991a)). The students were then given an Introductory Course in Artificial Intelligence (AI) covering sixteen short modules in one hour. The students had not studied Artificial Intelligence before. If the subject was a Verbaliser then the digit 0 was entered into the program interface and a textual presentation of the AI course was presented. If the student was classified as being a Bimodal – the digit 1 was entered and the program presented a learning environment that was both textual and graphical. Finally if the CSA designated the subject as being an Imager the digit 2 was entered and a highly illustrated presentation of the text was presented. The students studied the material for one hour and completed a short examination on the material.

4.4.2. Apparatus - Description of Website

The CSA measures an individual's position on two dimensions: the Verbaliser-Imager and Wholist-Analytic (Riding 1991a). The Verbaliser-Imager dimension only was used for the purpose of this study. The Verbaliser–Imager dimension (V-I) is indicative of whether the individual, while thinking, tends to represent information verbally at one end of the continuum, in mental pictures at the other or in-between (Bimodal). This test was used to assess the students' position along the Verbaliser-Imager continuum. The data were analysed using Data Desk 6.0 (Data Desk (1997)).

The website was developed using Active Server Pages (ASP) and Interdev for developing the front end (Microsoft, 1998). ASP was

used with HTML to create dynamic pages. The cognitive style score from Riding's CSA was entered on a log-on screen. This score is sent to the Web server resulting in the appropriate mode and dynamic HTML page being sent back to the browser. The interface was divided into three regions, facilitated by the use of frames. See Figures 4.2 and 4.3 for screenshots of the user interface.

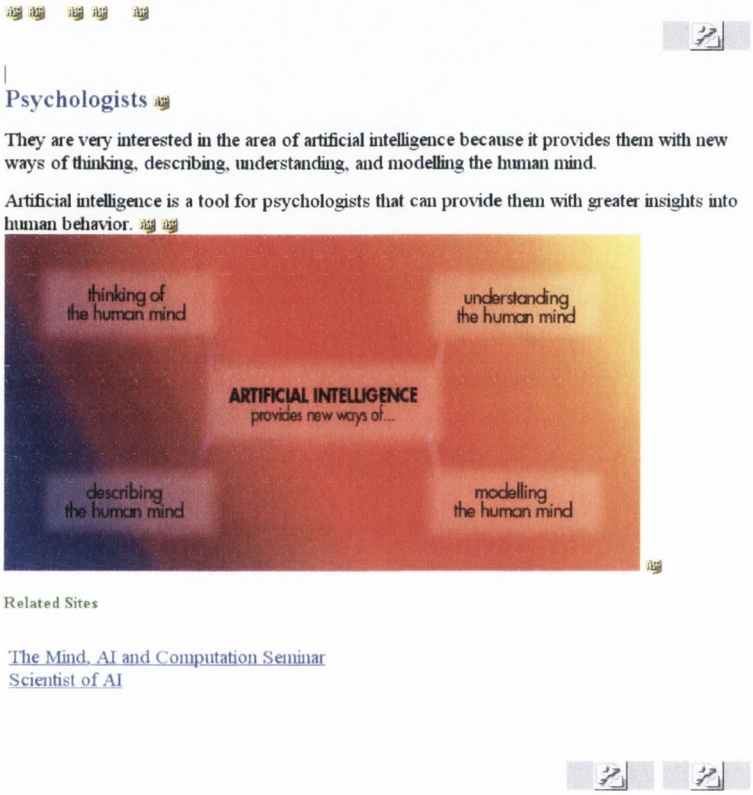


Fig. 4.2 Screenshot of website for Study 4.3 - Text

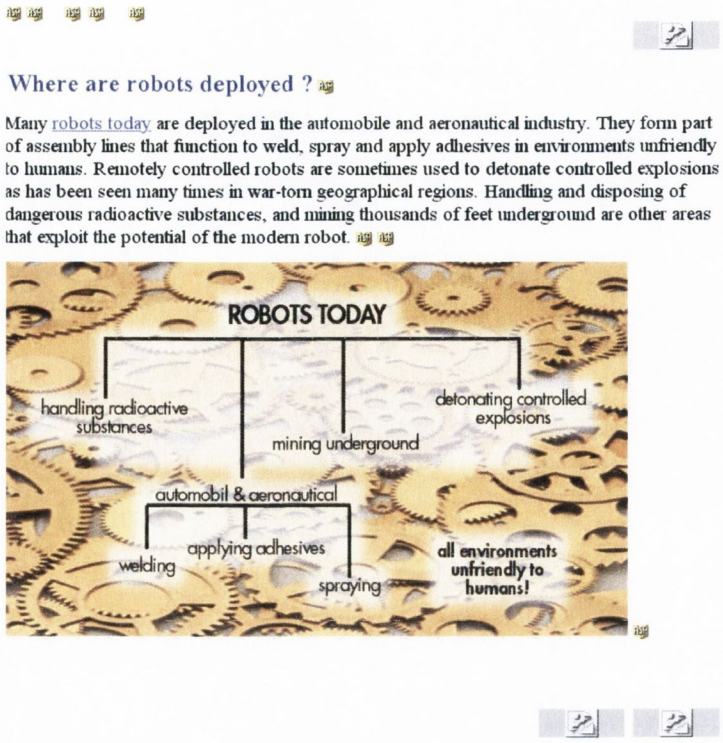


Fig. 4.3 Screenshot of website for Study 4.3 - Bimodal

A short, objective-style, ten minute test on the material was completed. The test was negatively scored for each incorrect response given and positively marked otherwise. The test assessed both the higher and lower cognitive domains (Bloom, 1976) e.g. factual recall versus synthesis and application.

Participants: Fourteen, mainly Engineering and Computer Science students, were administered the CSA in conditions suggested by Riding's CSA Manual (Riding, 1991a).

4.4.3 Results of Study 4.3

A simple linear regression analysis was performed on the data and this confirmed the fact that V-I did not predict examination score as is evident by the p value (0.493) being greater than 0.05 in Table 1.

Table 1: Regression: Exam Score vs. V-I

Variable	Coefficient	Standard Error of Coefficient	t-ratio	Probability P
Constant	39.0778	11.36	3.44	0.0049
V-I	-15.5191	9.984	-1.55	0.146

This table portrays the regression line for exam score to be = 39.0778 - 15.5191 * V-I. Ideally the slope B (= -15.5191) should be zero (to indicate complete independence of exam score and V-I in this linear relationship. In fact the slope is not statistically significant at the 0.05 level as is shown by the fact that 0.146 is greater than 0.05. In other words exam score and V-I are largely independent of each other. To provide more precision in this analysis a larger sample size is required. The results of this Study were presented in summary form in Redmond & Parkinson (2004a).

4.4.4 Discussion of Study 4.3

Web-based learning presents a paradigm where, rather than matching students' and teachers' cognitive styles as suggested by Witkin, Moore, Goodenough and Cox (1977), the individual's web interface can be personalised to his/her cognitive style by judicious use of design hence enhancing the learning experience. In the future, it may be possible for the individual to use a browser and/or software package that is customisable automatically or otherwise to his/her cognitive style.

Tan and Lo (1991) report, "if the interface can be customized to suit the idiosyncrasies of the user, the effect that the cognitive style of the user has on the system can be reduced if not removed. Jennings, Benyon and Murray (1991) argue, that "adaptive systems should be able to accommodate the preferred interface styles of different users". Their empirical evidence suggests that task performance is enhanced when this is the case.

In this study, the user interface was customised to accommodate Verbalisers and Imagers differently. For example, the Imagers were presented with a highly illustrated version of the AI course, whereas the Verbalisers were presented with a textual version of the same material. In this study Verbalisers and Imagers performed equally well since they were both in environments that matched their cognitive styles. Consequently the Verbal-Imager dimension was not statistically significant as a predictor of examination score. However the results must be treated with a cautionary note given the small sample size used in this pilot study.

The interface in Study 4.3 can be classified as either a static interface, which can be adaptively changed once into one of several discrete options (three here) via an Adaptor (e.g. CSA score). Otherwise the interface can be called a system adaptive interface that is adapted once into one of a number of discrete options.

The adaptive interface as a concept has been around for some time. The lack of useful, generally available, realistic adaptive interfaces underlines the difficulties in providing them. The testing of such interfaces is also difficult especially in terms of server and client configurations and the problems in assembling large sample sizes. The measurement of individual differences to be used in accommodating such an interface also presents severe problems. Without doubt, conceptualising how an interface should be designed to be dynamically adaptive presents quite a challenge. The implications of presenting information that matches the characteristic ways an individual processes and represents it are: improved usability characteristics leading to higher levels of user satisfaction and, consequently in the future, better interface design, more effective user performance. Negroponte's quote (1995) is still quite relevant:
"The challenge for the next decade is not just to give people bigger screens, better sound quality, and easier to use graphical input devices. It is to make computers that know you, learn about your needs ... That's good interface design".
Indeed Shneiderman believes "we can make a difference in shaping the future by ensuring that computers serve human needs" (Shneiderman, 1999).

4.5 Results from the Three Pilot Studies 4.1, 4.2 and 4.3

The results from the three pilot studies suggest that there is merit in in the idea of using the V-I dimension in HCI. The pilot studies will be extended futher to a more in-depth of the V-I dimension pending a more detailed examination of Riding's CSA's V-I dimension.

4.6 Conclusions

This chapter reviewed the different types of flexible systems and critically assessed the merits of each one. The basis of the matching / mismatching hypothesis was also examined and its

usefulness scrutinized. Flexible systems, which were designed to accommodate cognitive styles, were reviewed so as to provide a basis for the design of some prototype systems already discussed in Chapter 2.

Studies by Parkinson and Redmond (2003), Redmond, Walsh & Parkinson (2003), Parkinson, Redmond & Walsh (2004a) and Parkinson & Redmond (2005) have demonstrated that non-adaptive systems can be designed to accommodate cognitive style, and such systems are relatively uncomplicated in their design. The main requirement is that the designer have an accurate conceptualization of the individual difference to be catered for when designing the system. These Studies have demonstrated an evolution over time in the accommodation of cognitive style and interface design.

Arising from the discussion in the previous section, it appears that the issue of the matching / mismatching hypothesis is unresolved. It seems logical that the next step in interface design is the incorporation of a test of cognitive style within the user interface. This will be discussed further and demonstrated in Chapter 9.

Based on the literature reviewed in Chapter 2 and this Chapter, it seems appropriate to base the design of the prototype application on a variant of a non-adaptive system approach. A new test of cognitive style (VItrap - Ch. 6) will function as an adaptor as discussed by Egan and Gomez (1985) in the prototype discussed in Chapter 9. In other words, this new test of cognitive style will be integrated into the interface.

While the conundrum of the matching / mismatching hypothesis remains unresolved, the applications discussed in Chapter 9 will evaluate whether accommodating individual differences via the matching / mismatching hypothesis was worthwhile or not.

The key results from the empirical studies (studies 4.1. 4.2a, 4.2b and 4.3) in this chapter have been summarized in papers in the literature (Parkinson and Redmond (2003), Redmond, Walsh & Parkinson (2003), Parkinson, Redmond & Walsh (2004a) and Parkinson & Redmond (2005)).

In Chapter 4, the matching/ mismatching hypothesis which underpins the accommodation of cognitive style, was discussed in the context of a number of studies (Studies 4.1, 4.2 and 4.3) which demonstrated how cognitive styles of users could be accommodated so as the improve their use of the human computer interface.

In Chapter 5 the suitability of Riding's CSA for human computer interface accommodation will be examined psychometrically.

Chapter 5:
The Validity and Reliability of Riding's Cognitive Styles Analysis.

5.1 Introduction

The overall purpose of this thesis is to facilitate user performance by varying the human computer interface according to a user's cognitive style. Studies 4.2a, 4.2b and 4.3 in Chapter 4 suggest that this is feasible. The next logical step was to choose a psychometric test, ideally computer-based. The only available test to meet this requirement was Riding's CSA (1991). (It should be pointed out that Study 4.3 was done contemporaneously with reliability studies carried out in this Chapter into the reliability of the CSA).

The purpose of this chapter is to examine the psychometric properties of Riding's CSA. Riding's test has been commercially available since the early 1990's and in computerised form since 1991. Riding's test was assumed to have been extensively validated, given the large body of literature referring to it, and also its widespread worldwide use.

Riding's test was picked for a number of reasons: i. it was easy to use, ii. self-scoring for students, iii. it was relatively inexpensive for a one year license, iv. it did not require an extensive training programme to administer and use it (unlike, for example, Kirton's KAI which required a two day training program, v. it did not require the presence of a qualified educational psychologist to administer it or interpret the scores (unlike Witkin's Group Embedded Figures Test), vi. the test was readily available on an inexpensive 3.5 inch floppy disk and the results were written to the disk, vii. this test was the first of its kind to assess cognitive style in a computerised format, viii. the test took only about twenty minutes to administer and did not result in user fatigue and finally, ix. being computerised the test had the potential to be directly embedded in the human computer interface. Riding's test met the requirements for test construction and it generally well accepted as evidenced by the large literature based around it.

Initially it appeared that this test of cognitive styles was ideal as

a basis to guide the design of a computer interface. Upon closer inspection there was no available evidence to suggest its reliability despite numerous studies attesting to its validity. All useful psychological tests need good reliability and validity. It was necessary to investigate the reliability of the CSA before proceeding with any further interface design implementation.

Riding's CSA is a computerised test which assesses two dimensions of cognitive style Wholist – Analytic and the Verbaliser- Imager. This design of this test was innovative as it was computer based and easy to administer. It was one of the first of its kind to be computerised. It should be noted that the Coffield Report (2004) was published well after the empirical work was done for this thesis. In fact Coffield quotes the result from this empirical work in his Report (Redmond, Parkinson & Mullally (2002)).

5.2 Validity

There are numerous types of test validity. Rust and Golombok (1999) present an overview of test validity. They discuss what they consider to be the main categories of test validity. These categorisations comprise face validity, content validity, predictive validity, concurrent validity and construct validity.

They define Face Validity as "the acceptability of the test items, to both the user and the respondent for the operation being carried out". It is crucial that the respondent take the test seriously otherwise "the results may be meaningless".

These authors next define Content Validity as being "the extent to which the test specification, under which the test was constructed, reflects the particular purpose for which the test was being developed". They assert that "In an educational setting, content validation will generally involve a comparison between the curriculum design and the test design". (It is assumed that there will be a considerable overlap between them). This can also be known as "Criterion - Related or Domain Referenced Validity". They state that if the test specification is not reflecting the task specification, it must be reflecting something else, and all else is a source of bias".

Predictive Validity is defined "is represented as a correlation between the test score itself, and a score of the degree of success in the selected field, usually called success on the criterion". For example, one could assume that the final examination score at second level should predict a candidate's success at University. The authors state, that "Predictive Validity is the major form of statistical validity, and is used wherever tests are used to make predictions".

Concurrent Validity "describes the correlation of a new test with existing tests that purport to measure the same construct". The authors give as an example a new test of intelligence correlating with other older tests of intelligence. The authors conclude the section by saying that "if old and new tests of the same construct fail to correlate with each other then something is seriously wrong". (The authors may have picked an unfortunate example in this case as the definition of intelligence and what intelligence tests actually measure is difficult to establish anyway).

They also discuss Construct Validity and state that it is "the primary form of validation underlying the trait-related approach to psychometrics". They describe Construct Validity in the following way "the entity which the test is measuring is not measurable directly, and we are really only able to evaluate its usefulness by looking at the relationship between the test and the various phenomena which the theory predicts". The authors conclude that" Construct Validation is never complete but is cumulative over the number of studies available ".

In their review they describe the multitrait - multimethod approach (Campbell and Fiske (1959)). Campbell and Fiske argued that in order to establish the Construct Validity of a test it should not only correlate with other similar measures of same construct (Convergent Validity) but be shown to correlate poorly with other tests that measure different constructs entirely (Discriminant Validity).

5.2.1 Validity of the CSA

Riding formulated his own set of criteria, which he used to investigate whether a construct can be considered to be a style, or not. Riding (2000) has stated "For a style to be shown to possess a degree of construct validity, several conditions need to

be fulfilled; the style constructs should not duplicate other constructs and should relate to a range of behaviour". More specifically the dimensions should be: independent of intelligence, independent of personality, and be related to observable behavior. Cognitive style should also have a neurophysiological basis.

The cognitive styles as assessed by the CSA have been shown to be orthogonal (Riding and Douglas (1993) and Redmond, Mullally and Parkinson (2002)). The CSA was shown to be independent of intelligence (Riding & Pearson (1994), Riding & Agrell (1997)) and by Parkinson and Redmond in an unpublished study, which is detailed in Appendix 7-5.2.B).

Riding and Wigley (1997) investigated the interactive effect of the CSA two style measures with that of Eysenck's Personality Questionnaire (EPQ-R) (Eysenck and Eysenck (1975)) and the State Trait Anxiety Inventory (STAI) (Speilberger (1977)). In this article the authors proposed "a tentative model in which physiologically based personality sources are independent of cognitive style but are moderated by style in their effect on behaviour" (Riding and Wigley (1997)).

Riding and his co-researchers have provided a plethora of research during the last decade in claiming to establish the construct validity of the CSA. They have investigated cognitive style and the affect on school attendance (Riding and Baker (2000)), (Riding and Rayner (1998)). They investigated the affect of cognitive style on the preferred mode of material presentation (Riding and Ashmore (1980), Riding and Douglas (1993), Riding and Watts (1997), Riding and Al Sanabani (1998), Riding and Rayner (1998), Riding and Rayner (2001), Riding (2001) in Sternberg and Zhang (2001)). Riding also examined the affect of cognitive style on stress in both the teaching and nursing professions (Riding & Borg (1993), Riding & Wheeler (1995)).

Riding and his co-researchers (1997) provided physiological evidence for the existence of cognitive style. They investigated the relationship between information processing tasks and alpha suppression using the electroencephalogram. They concluded that "overall the results indicate a clear distinction between the two style dimensions in terms of brain functioning".

Evidence for the validity of Riding's CSA has been summarized extensively by Riding and Rayner (1998), Riding and Rayner (2001), Riding (2001) in Sternberg and Zhang (2001). Table 5.1 summarises the different types of Validity based on the Riding literature.

Table 5.1 Validity Categorisations and the CSA

Face Validity	Difficult to say
Predictive Validity	None
Concurrent Validity	None
Construct Validity	Yes but constrained by the low reliability of the CSA
Multitrait -Multimethod approach	None

Based on this categorisation of validity, studies completed by Riding would indicate that the validity of the CSA has not been extensively reported on and what has been reported, is questionable given the low test re-test reliability of the CSA (Vid. later in this Chapter).

Redmond, Mullally & Parkinson (2002), Parkinson, Mullally & Redmond (2004c), Peterson et al. (2003a), Rezaei & Katz (2004) and Massa & Mayer (2005), all have raised concerns about the validity of Riding's CSA. Redmond, Mullally & Parkinson (2002), Parkinson, Mullally & Redmond (2004c) state that "the low level of test retest reliability raises concern about the validity of the CSA". They also state, that "it is probable that the test items do not measure the W-A dimension or the V-I dimension with sufficient precision".

Peterson (2003) raises concerns about the validity of the CSA. She states, that "the extensive list of studies showing a relationship between the Verbal-Imagery and Wholist - Analytic style measured by the CSA is quite impressive and it raises the question of how an unreliable test can obtain all these associations". She comments on the lack of reporting of the effect size, i.e. the extent to which the null hypothesis is false (effect size) as opposed to the null hypothesis just being false (Significance Level). Another problem she raises, as has been mentioned elsewhere in this thesis, is the use of style labels/categories instead of style ratios when investigating "interactions and associations with behaviour". In some of Riding's own studies there is an anomaly. Riding allows

researchers to calculate the norms or the break points using his norms based on a standardised sample or they can divide their own sample size into two or three equally divided groups (Peterson et al. (2003a)). The consequence of this is an individual may be categorised as an Imager in one study and a Verbaliser in another. Peterson (2003) provides evidence of these anomalies in her Ph.D thesis (Chapter 11, Table 11.1).

Massa and Mayer (2005) assess the validity of the CSA under some of the headings referred to above and would argue that the test would appear to lack Face Validity. They state: "On its surface, the Verbal -Imager subset of the CSA does not appear to measure the test taker's primary mode of processing or representing information in thought". They also raise questions about the Face Validity of the underlying theory: "the Face Validity of the theory behind the test construction of the Verbal - Imager subset of the CSA is also lacking". The authors also refer to their previously published work (Mayer and Massa (2003)), which examined the Construct Validity of the CSA. They raise doubts about the Construct and Concurrent Validity in this study.

Massa and Mayer (2005) also raise doubts about the Predictive Validity of the CSA. They designed a computer based electronic program in which the learner could opt for material presented in visual or imagery form and recorded the number of times the learner opted for each of textual or image based screens. They state that: "if an instrument purporting to measure Verbaliser - Visualiser cognitive style has Predictive Validity then there should be a high correlation between the score on the instrument (CSA) and measures of processing during learning". They conclude, that "the Verbal-Imager Subset of the CSA does not appear to be a valid measure of Verbal-Visual cognitive style in that it fails to exhibit Face Validity, Construct Validity, or Predictive Validity". (Unfortunately their comment in relation to the CSA is factually incorrect as it attributes to Peterson, Deary and Austin's work as assessing the test retest reliability and split half reliability of Riding's CSA which this work did not do as discussed elsewhere).

5.2.2 Origins of the CSA

The origins of the Verbaliser - Imager dimension may be linked to Paivio's dual coding theory. This has also been supported by the Coffield report (2004), which suggests that "the verbal imagery dimension is related to Paivio's (1971) dual coding theory".

Riding and Buckle (1990) refer to the fact that the Wholist - Analytic dimension is founded upon Witkin's field dependence - field independence construct. As was referred to in Chapter Two this was investigated but little correlation was found between the CSA and the Group Embedded Figures Test both of which are used to assess these dimensions.

5.2.3 The Basis of the CSA - Inspection Time

The ratios that Riding uses to calculate the two cognitive style ratios for the Verbaliser - Imager and Wholist - Analytic dimension would seem to be based on inspection time.

Nettelbeck and Lally (1976) reported a significant correlation between psychometric intelligence differences and inspection time. Nettelbeck (2001) states, that "more than 25 years of research suggests that the measure inspection time (IT) does capture low level aspects of cognitive functioning that contribute to human intelligence". Cooper (2002) reports "most (of the studies) find that correlations between measures of ability and inspection time are of the order of –0.3 to – 0.5".

Grundnik and Kranzler (2001) performed a meta-analysis of over 90 studies, containing a total of 4100 participants. These studies all examined the relationship between inspection time and intelligence. The samples were drawn from a Normal populations and r was corrected for sampling error, error of measurement and range variation. They found that the degree of association between inspection time and IQ was r = -0.51. They concluded the "results of this meta analysis provide further support for the substantial relationship between IT and IQ".

The negative correlations reported in these studies occur because individuals with greater ability when compared with those of lesser ability will do the task quicker. Their result compared favorably with that of an earlier study carried out by Krantzler and Jensen (1989) who had previously reported a correlation of r

= - 0.49 between inspection time and IQ.

Inspection time is defined as "the minimum duration for which two different stimuli must be presented if they are to be perceived as different" (Irwin (2002)). Alternatively Cooper (2002) defines IT as "this simply measures how long a stimulus has to be presented in order to be perceived correctly".

Inspection Time can be measured in a straightforward manner. One way is to have the person sit down in a dimly lit room and face a computer interface. This individual will have as much time as he/she likes to make their response, the objective being to get the question correct without any time constraint.

He/she will see one of two shapes, which will then be followed by a mask. The mask has two thicker vertical lines of equal length. The task for the individual will be to state on which side of the shape is the vertical line the longest. The individual will report this to the person who is assessing the inspection time.

Riding's CSA calculates inspection times for the shape and type questions. As previously mentioned, an Imager would be expected by Riding to complete the colour questions faster than the type questions and vice versa for the Verbaliser. Peterson (2003) has supported this view. It would also be expected that the mean score of shape and type questions would correlate with the measures of intelligence used to validate the new test of the V-I dimension in Chapter 6.

However the total inspection times for the Verbaliser and Imager test items are not what are of interest. It is of critical importance to appreciate that the calculation of cognitive style score is a ratio. As a result of this, the affect of IQ is canceled in the CSA (essentially the IQ effect in the numerator is cancelled by the IQ effect in the denominator).

5.2.4 Description of the CSA computerized test

The CSA is a compiled, copyright, computer-presented test consisting of three sub-tests: the first assesses the V-I dimension, and the second and third assess the W-A dimension. Cognitive style analysis (CSA) measures an individual's position on its two orthogonal dimensions: V - I and W -A. The V - I dimension is

indicative of whether the individual, while thinking, tends to represent information either verbally or in mental pictures. On the other hand, the W-A style dimension describes whether an individual is inclined to organize information into wholes or parts).

The CSA consists of two main sections, the first one assesses the Verbaliser-Imager dimension; the other section containing two subtests each assessing different parts of the Wholist – Analytic continuum. The test consists of 48 questions that assess the Verbaliser-Imager dimension; half of these questions are incorrect. Test items which assess Imagers describe the colours of items: these test stimuli ask – "are x and y the same colour?". The test items which assess the Verbaliser dimension of the continuum are categorical, "are x and y the same type?". It is assumed that the Verbaliser will respond faster to the categorical questions, the Imager faster to the shape questions. The program records the individual's response times to the questions and calculates a ratio, the Verbal Imager ratio. A high ratio is characteristic of an individual being an Imager, a low ratio a Verbaliser.

The Wholist-Analytic dimension is assessed by means of two subtests. The first subtest presents pairs of geometric shapes as line drawings; the testee is required to establish whether the objects are similar or different. The second subtest presents pairs of objects, again as geometric line drawings, to the testee. The individual is asked whether one of the objects is embedded in the other. In each case twenty test items were presented to the testee, of which half were correct, and half incorrect.

It is assumed that the Wholist will complete Subtest 1 faster than the Analytic and conversely in Subtest 2.The test software calculates the ratio a Wholist – Analytic ratio. A high ratio corresponds to an Analytic, a low ratio to a Wholist. The CSA also calculated other test metrics, including a Speed index and percentage correct for both dimensions Wholist –Analytic, Verbaliser - Imager.

5.3 Study 5.1 of Riding's CSA - Ambiguity of Test Items

The purpose of this study was to investigate the ambiguity of the test items that assessed the V-I dimension. This issue has been

raised in Redmond, Mullally & Parkinson (2002) and Parkinson, Mullally and Redmond (2004c).

5.3.1 Method

A sample of 60 Information Systems students was given a questionnaire to complete. This questionnaire consisted of the list of the questions, that Riding used to assess the Verbaliser - Imagery continuum. The students were asked to tick the extent of their agreement / disagreement with each of the test items . The following histograms show the ambiguity of response for each item.

5.3.2 Results of Study 5.1 - Ambiguity of Test Items

The results of Study 5.1 are presented in the following histograms:

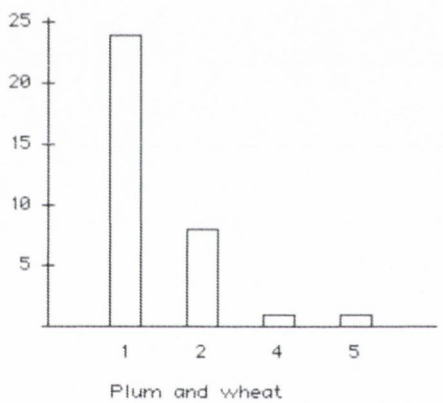


Fig. 5.1 Are plum and wheat the same colour?

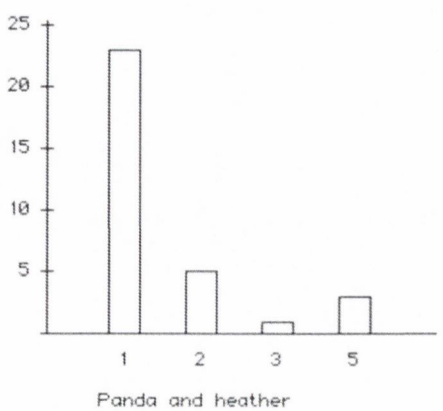


Fig. 5.2 Are panda and heather the same colour?

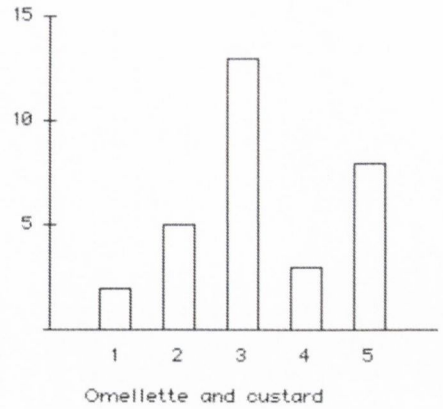


Fig. 5.3 Are omelette and custard the same colour?

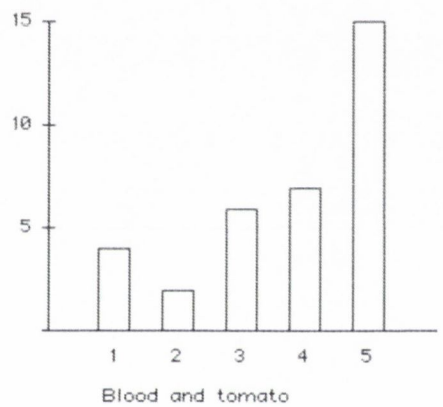


Fig. 5.4 Are blood and tomato the same colour?

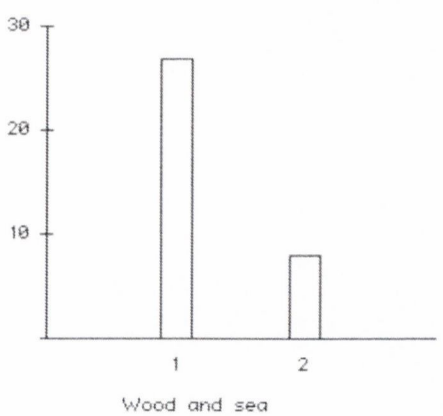


Fig. 5.5 Are wood and sea the same colour?

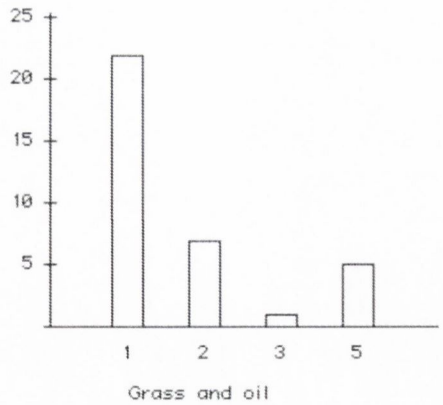


Fig. 5.6 Are grass and oil the same colour?

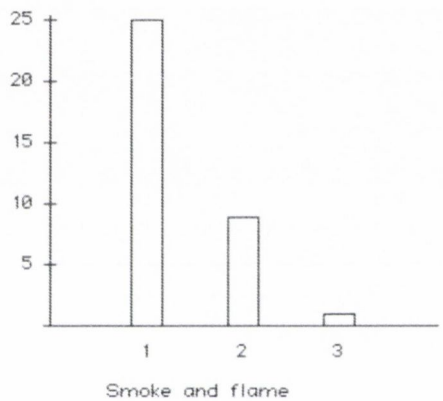


Fig. 5.7 Are smoke and flame the same colour?

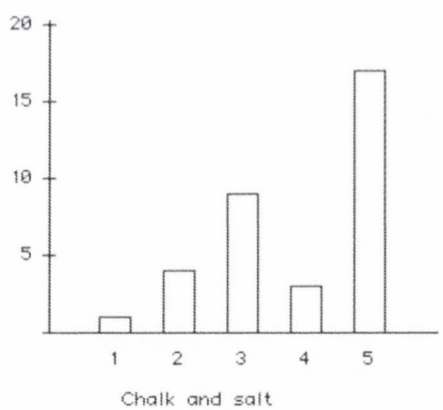


Fig. 5.8 Are chalk and salt the same colour?

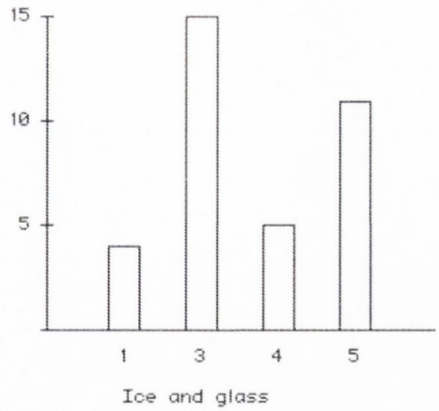


Fig. 5.9 Are ice and glass the same colour?

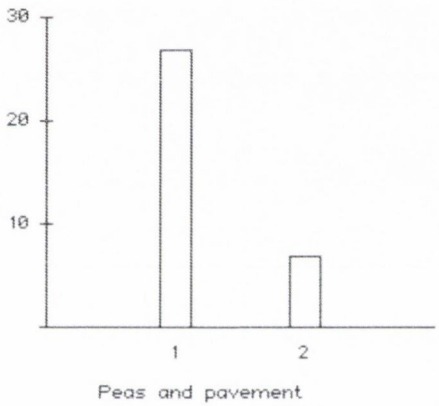


Fig. 5.10 Are peas and pavement the same colour?

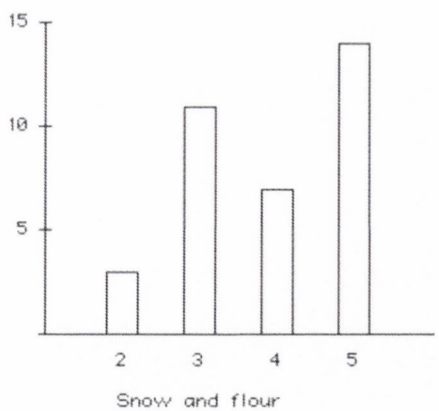


Fig. 5.11 Are snow and flour the same colour?

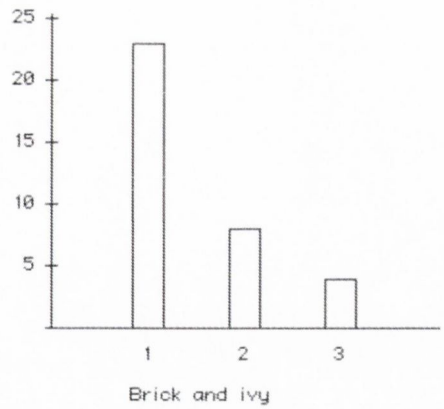


Fig. 5.12 Are brick and ivy the same colour?

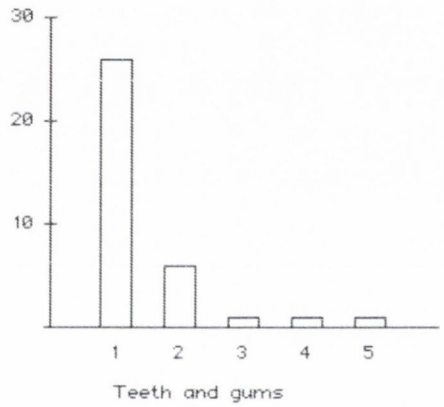


Fig. 5.13 Are teeth and gums the same colour?

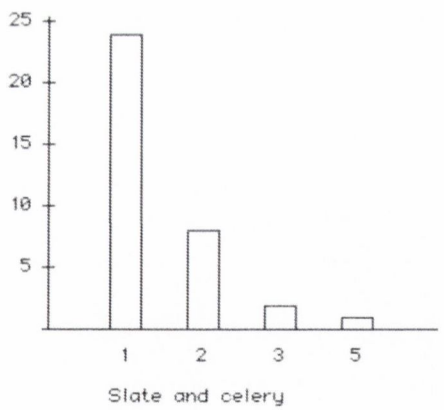


Fig. 5.14 Are slate and celery the same colour?

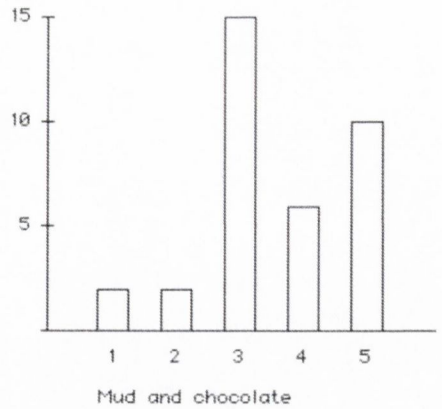


Fig. 5.15 Are mud and chocolate the same colour?

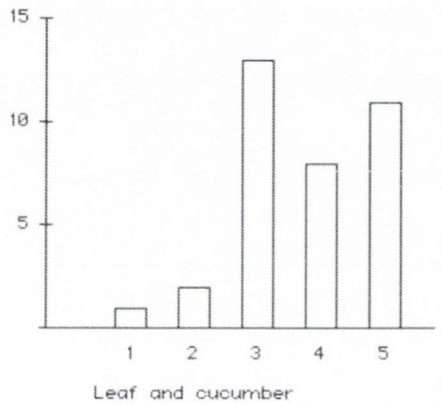


Fig. 5.16 Are leaf and cucumber the same colour?

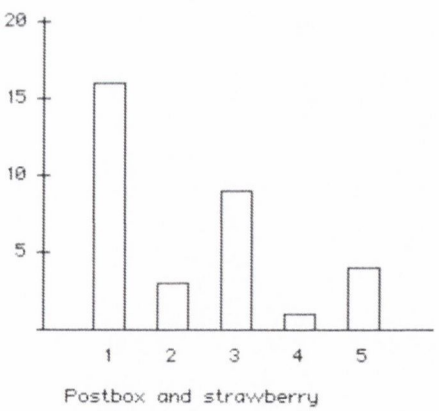


Fig. 5.17 Are postbox and strawberry the same colour?

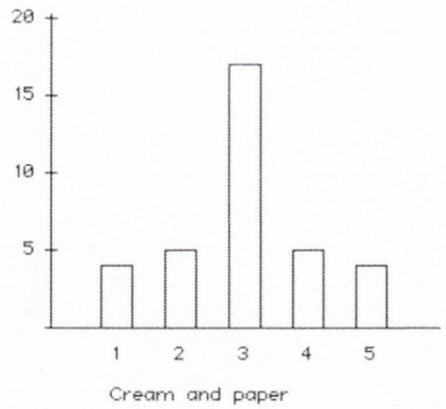


Fig. 5.18 Are cream and paper the same colour?

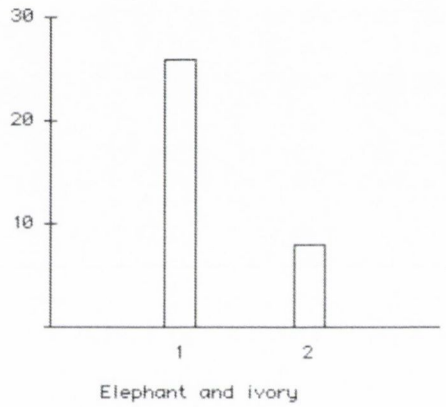


Fig. 5.19 Are elephant and ivory the same colour?

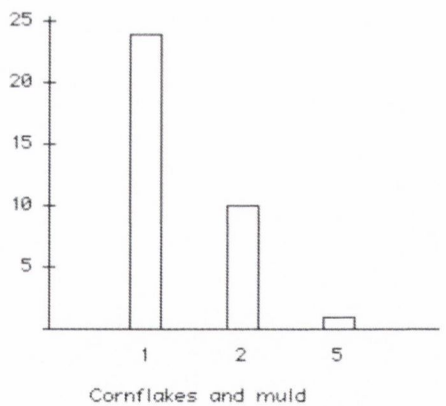


Fig. 5.20 Are cornflakes and mud the same colour?

The bar charts in these figures represent the extent of agreement /disagreement amongst a sample of 60 final year Information Systems students.

The key for the bar charts is:
1 = very strongly disagree
2 = strongly disagree
3 = neutral
4 = strongly agree
5 = very strongly agree.

Comment: these bar charts indicate the degree of ambiguity that existed amongst the CSA questions that assessed the Verbaliser - Imager dimension.

5.3.3 Conclusion of Study 5.1

These results demonstrated that some of the CSA test items were confusing for testees. In particular questions like "leaf and tar are the same colour?", "postbox and tomato are the same colour?", "Elephant and Ivory are the same colour?" and other test items caused much confusion for the respondents. If the questions in the CSA were unambiguous, there should not have been such a dispersion in the bar charts for each question. These issues were also raised in a previously published paper (Parkinson, Mullally & Redmond (2004c)).

5.4 Study 5.2 of Riding's CSA - Test - Retest Reliability

5.4.1 Introduction

The CSA was investigated as to its retest reliability. In all three individual studies of Study 5.2 the participants were psychologically naïve with regard to the purpose of the test.

Study 5.2.1 of Riding's CSA (n = 38, 12 day interval)

Thirty-eight students mainly from the disciplines of Computer Science and Engineering participated in these two 15 minute sessions (n = 38, average age 20.7, standard deviation 1.106). The test retest interval was exactly twelve days (Redmond, Mullally and Parkinson (2002)).

Study 5.2.2 of Riding's CSA (n = 51, 14 day interval)

51 students (mainly third year) from the disciplines of Computer Science (mostly), Engineering and some other disciplines including Arts (but none from psychology) (n = 51, M = 20.7 years, SD = 1.12) participated in these test–retest sessions of Riding's CSA with a mean retest interval of 14 days (T = 14 days,

SD = 4.6) (Parkinson and Redmond (2004c)).

Study 5.2.3 of Riding's CSA (n = 27, 23 month interval)
A different cohort of ninety-six (third year) Information Systems students participated in this third study on the first occasion. Approximately twenty-three months later twenty-seven of that original cohort (n = 27, M = 30.2 years, SD = 4.8) completed the CSA again in identical conditions (Parkinson, Mullally & Redmond (2004c).

5.4.2 Methodology

Pearson's product-moment correlation was used to calculate the correlation between test and retest sessions of Riding's CSA in Studies 5.2.1, 5.2.2 and 5.2.3. All studies were run on the same desktop PC machine types with identical CPU speed and memory. The students were all considered computer-literate. The data from the CSA tests were analyzed using Data Desk software (2006).

When Studies 5.2.1, 5.2.2 and 5.2.3 were completed, the students were given a presentation by the authors on the results and any implications. Students were then presented with their own individual results and resulting queries were addressed.

The possibilities of establishing the reliability of Riding's CSA test were limited by:
(i) The lack of psychometric information given in Riding's manuals,
(ii) A shortage of information on calculations and design decisions in his test,
(iii) The source code for his test is not available for inspection.

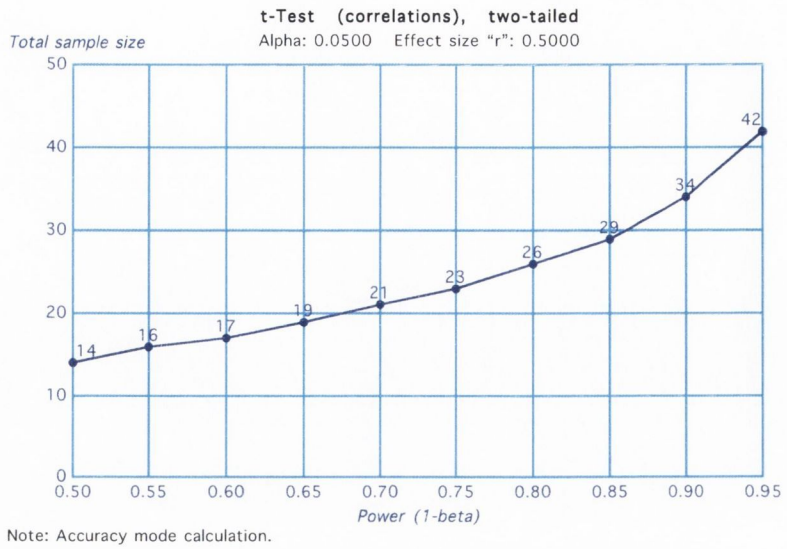
As a result any attempt to come up with a split-half or parallel form design would face the same rebuttal as Peterson, Deary & Austin (2003a) received from Riding (2003) i.e. that their version of his commercial test is not a valid copy of his test.

Consequently establishing the reliability of Riding's CSA test using a test–retest approach was the only approach that could be adopted but it must be acknowledged that this has its own limitations as discussed for example by Kline (1993).
Peterson, Deary and Austin's test–retest interval was 8.5 days

(Peterson (2003a), which Riding considers to be too short (Riding (2003)). Kline (1993) suggests that to be trustworthy the two testings should be separated by at least a three-month gap. To quote Riding (2001): "There is the need for a long-term test–retest reliability study over an interval of say one year between testing in a sample who are likely to be fairly naïve as to the purpose of the test". Study 5.2.3 exceeds Riding's requirement for a long retest interval.

Riding's cognitive style dimension ratios, style labels, and other test metrics are calculated automatically by the CSA package. An explanation of how the W-A and V-I ratios are calculated may be found in Peterson et al. (2003a). The data satisfied the criteria for accepting a participant's attempt as being valid as outlined by Riding in his CSA Administration Manual (1991b). The Pearson product-moment correlations were calculated for Session 1 (test) vs. Session 2 (retest) for each of the dimensions W-A and V-I for all three studies. When the data were log-transformed to adjust for skew, the calculated correlations were similar. Consequently only the non-transformed data are reported on.

All sessions ran without obvious problems. No grounds could be found for omitting any individual's results. Examination of scatter-plots showed no outliers. It was appropriate to power the studies to detect correlation coefficients in excess of about 0.5 given that the reliability of a psychological measure was being studied (Cohen (1977)). Thus with alpha set at 0.05 (two-tailed) and with $r = 0.5$, Study 5.2.1 ($n = 38$), Study 5.2.2 ($n = 51$) and Study 5.2.3 ($n = 27$) all had a power of in excess of 80% (Vid. Fig. 5.1). Adequate power is usually considered to be 80% or above (Furlong, Lovelace & Lovelace (2000), Cohen (1988)).



Note: Accuracy mode calculation.

Fig. 5.1 shows the relationship between Power, Total Sample Size with alpha set to 0.05 and effect size = 0.5.

5.4.3 Results of Study 5.2

Table 5.2: Results of Test-Retest Studies on CSA or Replica

	Sample size Test interval	W - A Ratios Test vs. Retest	V - I Ratios Test vs. Retest	W - A Style labels Test vs. Retest	V - I Style labels Test vs. Retest
Study 5.2.1 Redmond et al. (2002)	n = 38, T = 12 days	r = 0.56	r = - 0.21		
Study 5.2.2 Parkinson et al. (2004c)	n = 51, T = 14 days	r = 0.34	r = -0.19	r = 0.45	r = -0.03
Study 5.2.3 Parkinson et al. (2004c)	N = 27, T = 23 months	r = 0.30	r = 0.36	r = 0.54	r = 0.03
Peterson et al. (2003a) CSA-A Replica (not exact version) of CSA)	n = 50, T = 8.5 days	r = 0.30	r = 0.20	r = 0.34	r = 0.10
Rezaei & Katz (2004) Study 1 Study 2	n = 73, T= 7 days n = 36, T = 1 month	r = 0.42 - 0.45	r = 0.30 - 0.35		

Studies 5.2.2 and 5.2.3 present the correlation results from grouping the data into Riding's three style label categories for each dimension.

For comparative purposes results from Peterson et al. (2003a) (a

reprogrammed but not exact version of Riding's CSA), are included. The three Studies 5.2.1, 5.2.2 and 5.2.3 have been summarized in (Parkinson, Mullally & Redmond (2004c)). Additional results are given in Appendix 6-5.1.A.

5.5 Discussion of Studies 5.2.1, 5.2.1 and 5.2.3

The documentation, which accompanied the commercial version of the CSA, provided evidence of its validity but none as to its reliability. The reliability of Riding's CSA had not been reported in the literature to date.

These were the first data to be published on test–retest reliability of Riding's CSA test at mean intervals of (i) 12 days (ii) 14 days and (ii) 23 months. These results clearly show that Riding's cognitive styles analysis dimensions are unreliable for the sample populations given these sample sizes and test–retest intervals.

Rezaei & Katz (2004) investigated the reliability of the CSA using a test retest of 7 days and 28 days. They report correlations ranging from r = 0.42 to 0.45 for the W-A dimension and r = 0.35 to 0.30 for the V- I dimension. These results (also given in Table 5.2) support the figures reported by Redmond et al. (2002) (2004c). Rezaei & Katz (2004) in this study do not report whether there was a learning effect or practice effect. They also potentially biased the third experiment by telling the participants that reaction time was important as a variable in this third study. This is a weakness in the third study and violates Riding's instructions for administering the test.

All these correlation coefficients fall far short of the commonly accepted value of r = 0.8 (or at least 0.7) for test–retest, for instance as given by Kline (1993). What can be concluded is that the reliability of the W-A dimension is stable but low whereas the V-I dimension by contrast is quite erratic. The test items, which assess the V-I dimension, are of a more subjective nature, perhaps partially explaining this evident instability. The W-A test – retest correlation was stable over time at about 0.34 for both Studies, but the V-I dimension was not, with values of –0.19 and 0.36, which raises questions about the performance of CSA in assessing this dimension. The V-I dimension was particularly unstable where the test–retest interval was of a shorter duration.

While not a valid copy of the CSA, Peterson, Deary and Austin's (2003a) test gives similar test–retest values for the W-A dimension for a short time interval (Table 5.1).

These test–retest studies were designed to address the contentious issue of the reliability of both dimensions of Riding's cognitive styles analysis (CSA) test (1991). In particular, Peterson, Deary & Austin (2003a) have attempted to assess the reliability of Riding's CSA test using an approximate model (CSA-A) of his test. Riding, in rebuttal (Riding (2003)), has stated that "a reliability study surely requires the use without modification of the test of which it claims to assess the reliability", and "... in fact the test used in the study was significantly different from the CSA ... " and later "Consequently the study is not a reliability test of the CSA" (Riding (2003)). Kline (1993) has commented: "A test is said to be valid if it measures what it claims to measure". Riding (2003) gives an example of a test, which is reliable but not valid (measuring personality using a tape measure). He also gives an example of a Startle test which is valid but whose reliability diminishes over time.

Peterson, Deary and Austin's (Peterson et al. (2003a)) test–retest interval was 8.5 days, which Riding considers to be too short. Kline (1993) suggests that to be trustworthy the two testings should be separated by at least a three-month gap. To quote Riding (2001): "There is the need for a long-term test–retest reliability study over an interval of say one year between testing in a sample who are likely to be fairly naïve as to the purpose of the test". Study 5.2.3 exceeds Riding's requirement for a test retest interval.

The low level of test–retest reliability raises concerns about the validity of the CSA. It is probable that the test items do not measure the W-A dimension or the V-I dimension with sufficient precision. Rezaei & Katz (2004) make this point "Riding's claim about the validity of the CSA was doubted immediately" and then they also state "that reliability is a necessary condition for any kind of validity".

When some of the research conducted by Riding is closely examined, it is noticed that style labels are used instead of actual CSA scores; this abstraction may lead to inaccuracy given that it

is a summary measure. In some cases this results in an individual being categorised as a Wholist in one study and as an Analytic in another depending on how the Style Labels are calculated. This may also be true for the categorisation of an individual as Verbaliser or Imager in the other dimension. Peterson (2003) has also commented also on this anomaly.

Generally Riding approaches establishing the validity of the CSA from a construct validity perspective. There are a number of different approaches to verifying test validity. Benchmarking the CSA against other tests (VVIQ, IDQ and VVQ) would have been interesting. The CSA wasn't accurately measuring its own constructs given the low test- retest values. How can such extensive validation be reported for the CSA in the literature? Rezaei & Katz (2004) may also provide the answer "the low correlations coefficient between the scores for separate constructs may be more of a function of low reliabilities rather than the independence of the constructs."

The Coffield Report (2004) reviews Riding's CSA as well as the literature that has been published on it. Peterson, Deary and Austin (2003a) never tested Riding's copyrighted version of the CSA. This point was made quite clearly, by Riding in his reply to these authors (Riding (2003)). Peterson et al. (2003b) replied to Riding's comments but did not address the point that they did not test the CSA. As has previously been mentioned in this thesis, these authors tested their own model of the CSA based on how they imagined it was constructed. Riding (2003) pointed out the shortcomings of this approach.

Riding (2003) states that "reliability per se although desirable and convenient is not the most important characteristic of a test" and discusses the Startle effect as an example. The Startle test is not useful in a discussion of CSA reliability since the respondents are not expected to decline in reliability on retest, certainly over a short time interval. The Coffield report supports this point of view and states, "yet he (Riding) offers no reasons for suggesting that the CSA is valid when first administered but not on later occasions".

The pathological cases of no validity and full reliability, or the converse, are not useful in practice. Conceptually, maximizing

the product of validity and reliability is more useful, since it then becomes possible to choose between different approaches of differing validities and reliabilities. Riding's example of no validity with full reliability, and its converse, would both have a validity–reliability product value of zero.

Riding's CSA is a computerised test, perhaps the first of its kind, which assesses two dimensions of cognitive style namely the Wholist – Analytic and the Verbaliser - Imagery. The design of such a test should have many advantages when compared with traditional pen and paper forms: ease of use, efficiency and accuracy in correction, ease of interpretation, assessment of larger samples simultaneously over computer networks with immediate calculation of cognitive style score. However it would have been better for the test developer to provide greater transparency and information, which would have helped answer research questions. Computerized psychometric development can raise ethical issues such as confidentiality and security of test data.

These 5.2 Studies which investigated the reliability of Riding's CSA also emphasize the need for regulatory bodies to ensure that any psychometric test, before it becomes available, should be rigorously tested with respect to set agreed criteria. Suggested criteria would include comprehensive reliability statistics, different measures of test validity, accurate reporting of samples used, acknowledgement of outliers and accurate power calculations being presented.

Should measures like these be adopted, researchers could feel confident about the robustness of the psychometric instruments used. This would add greater certainty to the legitimacy of the results that they report.

There were some difficulties with the design and implementation of the CSA. Some of these have been discussed under the methods section. The test required that two stickers be placed on specified keys on the keyboard. These allowed the testee to enter the correct /incorrect response respectively. Remembering the function and colour of which key to press may have also been a contributory factor in accounting for this test's unreliability. The red colour of the sticker may not have been easily visible for

those who were colourblind.

It may have been the case that having to remember which key to press may have caused hesitancy and therefore added to increased inspection times. Dyslexia can be manifested in individuals in a few ways. The most commonly known form expressed in the phenotype is a difficulty in reading. Interestingly enough it can also be expressed in a milder form as a difficulty in identifying left from right.

Another source of possible error is the use of the ratio in determining the two cognitive style scores on this test. A ratio is a summary measure and is an abstraction from the data; Lohman (1999) discusses this in greater detail. These studies have shown that the test is unreliable so it is possible that each ratio consists of two sources of unreliability, the numerator and the denominator. It also may have been the case that for the retest session this practice effect would have become a learned response resulting in faster inspection times.

The Coffield Report (2004) states that "few tasks in everyday life make exclusive demands on either verbal or non-verbal processing which are more often interdependent or integrated aspects of thinking" so a logical query from this might be how useful is the Verbaliser or Imager label if most people are in-between? The authors of this report contend that people who are very competent verbally or spatially do not tend to avoid other forms of thinking based on Carroll's (1993) work. Coffield while discussing the design of the CSA states that "the onus must be on the test constructor to show that consistent results are obtainable with different types of task and with verbal and non verbal presentation" and that there "are serious problems in biasing the assessment on a ratio measure as two sources of unreliability are present instead of one".

Another study conducted by Mayer and Massa (2003) report very low correlations for the CSA with measures such as Richardson's Verbalisers - Visualiser Questionnaire ($r = 0.2$). In this study the authors reported that "we are unable to validate the Verbal - Visual scale of the Cognitive Styles as a measure of cognitive style ... in short, we are not able to specify what the Cognitive Styles Analysis measures but it does not seem to

measure what other instrument designers think of as cognitive style or learning preference".

There are a few inaccuracies in the Coffield Report: for instance, the authors have inaccurately stated that Peterson, Deary and Austin (Peterson, Deary & Austin (2003a)) have demonstrated that the CSA was unreliable. As has been pointed out elsewhere in this thesis, these researchers never examined the CSA, but a re-engineered copy of it (which is not an accurate copy of the CSA as Riding stated), as to its test retest reliability. So therefore attributing this contribution to these researchers is incorrect.

Coffield did base his negative review of Riding's CSA on a paper written by the Author (and co-authors) in preparation for this thesis (Redmond, Mullally & Parkinson (2002)) which was supported by the later work of Rezaei and Katz (2004)).

5.6 Conclusions

Study 5.2.1 is the first study reported in the literature on the poor test retest reliability of Riding's CSA. This study has demonstrated the unreliability of the CSA. Subsequent Studies 5.2.2 and 5.2.3 authenticate these findings. Rezaei & Katz (2004) have verified these findings subsequently also.

The test-retest reliabilities of Riding's CSA (Riding (1991)) failed to satisfy Kline's criterion of $r = 0.7$ Kline (1993)). The W-A dimension proved to be more stable than the very erratic V-I dimension, over two test-retest intervals with two different cohorts.

Peterson, Deary and Austin (2003a), while claiming to have tested the reliability of the CSA, actually tested a simulated form of the CSA test, as Riding (2003) subsequently commented. However the broad results of this approximate are supportive of our results.

These findings cast doubt upon other studies supporting the validity of this measure and as such their findings should be treated with caution. This is particularly relevant when considering studies that have reported significant results, with respect to the Verbaliser - Imager dimension and the Wholist - Analytic dimension.

If the difficulties pertaining particularly to the unreliability of Riding's V-I dimension can be successfully addressed in a new test then the opportunity exists to re examine and replicate those studies already discussed in order to establish the validity of the new test.

The limitations of the studies in this Chapter are that the samples are probably not representative of the population as a whole, being third year technical university students.

In Chapter 6 a new test for the V-I dimension will be developed and examined as to its reliability. A psychometrically sound test is obviously necessary if it is to be in the accommodation of cognitive style in the human computer interface.

Chapter 6:
The Design of a New Test of the V-I Dimension - VItrap

6.0 Introduction.

In Chapter 5, it has been shown that Riding's CSA is not reliable particularly in the V-I dimension. The logical next step is to design a better new test for the V-I dimension with the intention of integrating it into the human computer interface. This approach will facilitate the accommodation of the use of individual differences such as cognitive style in the Human Computer Interface (HCI).

This chapter describes the design, implementation and reliability of a psychometric test, VItrap, of the Verbaliser - Imager dimension which has been influenced by Riding's CSA. It was concluded at the end of Chapter 5 that Riding's test of the V-I dimension was very erratic and that the test items were subjective and vague. It was hypothesized that this may have added to the unreliability of the V-I dimension.

The test was designed to remedy the deficiencies apparent in Riding's CSA test to do with the V - I dimension. The new test assessed the V-I dimension using both word and picture forms. The input device was changed from the keyboard to keyboard and mouse. The number of questions used in word form to assess both dimensions remained similar to Riding's CSA. When the Spearman Browne (Appendix 8-6.1.A) formula was used to estimate the number of test items required in Riding's CSA to improve the reliability of the V-I dimension (assuming a desired r of 0.7), it was found to be impractical because the number of test items would need to be 672 (Vid. Appendix 8-6.1.A). This would present a huge fatigue problem for testees and so was impractical. Consequently for the design of the VItrap, this number of test items would also be too fatiguing so therefore greater emphasis was given to the design of non-ambiguous test items. In other words each test item was chosen to provide a larger contribution to reliability by dropping potential test items, which were found to be ambiguous for schoolchildren and university students.

6.1 Test Item Design

Prior to selecting the one hundred and twenty test items for inclusion approximately two hundred candidate test items were sampled amongst two different populations - university students and school children (from 12 to 16 years of age). The final choice of test item was determined by strong sample response (Vid. Appendix Table 12-6.5.E).

The verbal questions in Riding's CSA involve comparing the categories of two objects e.g. "Are 'Skiing' and 'Cricket' the same type?". Riding & Rayner (1998) argue that this is a verbal task because it requires identification of the semantic conceptual category, which is verbally abstract in nature and therefore cannot be represented in the visual form. The imagery questions in Riding's CSA involve comparing the colour of two objects (e.g., "Are 'Lettuce' and 'Lawn' the same colour?"). Riding argues that this is an imagery task because to compare the colour of two objects requires the generation of a mental picture. For the VItrap test the same assumptions that Riding made for his CSA are made.

The test consisted of two main types of test item, which assessed the Imagery and Verbaliser end of the continuum. The test item "are X and Y the same shape?" assessed the Imager end of the continuum (Fig. 6.1).

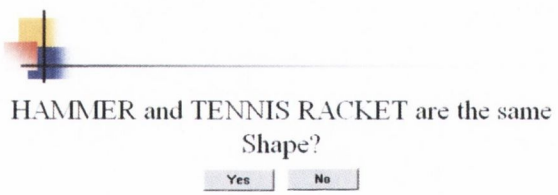


Fig 6.1 An example of a Test Item (word form) used to assess the Imager dimension

Similarly "Are X and Y the same type?" assessed the Verbaliser end of the continuum (Fig. 6. 2).

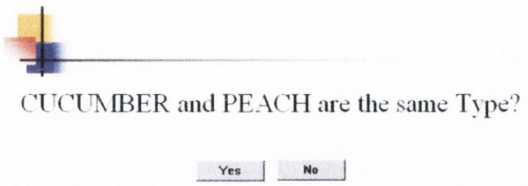


Fig 6.2 A test item (word form) which assesses the Verbaliser dimension.

The structure of this test item has been taken from Riding's CSA. The testee was required to indicate which of the statements were true or false.

The test also had a tutorial, which comprised sample questions pertaining to each of the forms and test items. Detailed feedback was given in the tutorial (Fig 6.3).

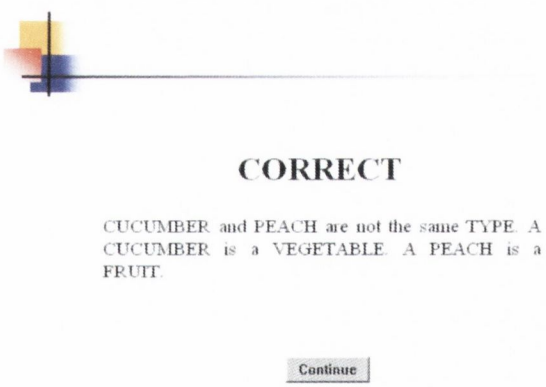


Fig 6.3 An example of feedback given in the tutorial

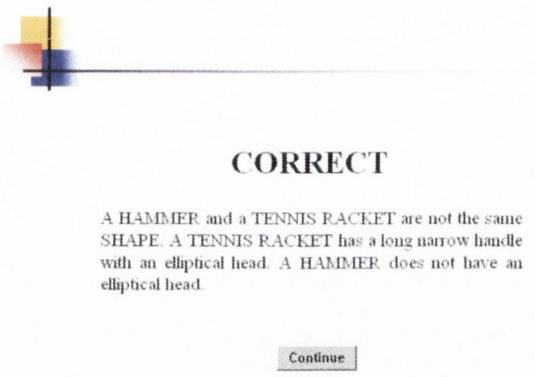


Fig 6.4 A further example of feedback in the tutorial

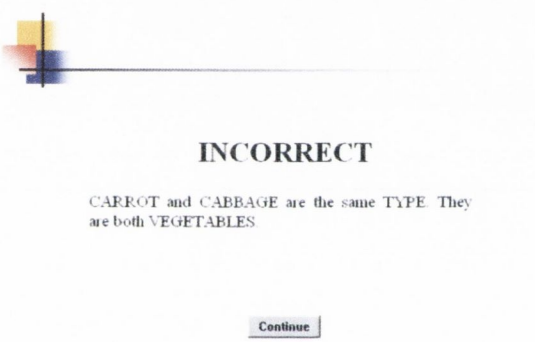


Fig 6.5 Negative feedback in the tutorial

The test included two different forms of test item: word and picture. The word items were presented first. Next the picture items were presented in exactly the same order i.e. there was no randomisation of test items. If the test item "Tiger and lion are the same shape?" is presented as test item 5 in word form it will also be presented as test item 5 in pictorial form. The test contained thirty questions of each type in both forms, one hundred and twenty in total. The testee was given feedback as to the correctness of his/her response immediately after each question as was done in Riding's CSA.

The following data were recorded for each test item: Inspection Time and, whether the response was correct/incorrect. The average Inspection Time for the Verbaliser test items (word form) was divided by the average Inspection Time for the Imager

test items (word form), to give Vq/Iq. This is the ratio of interest in this study. Other ratios were also examined:

(Vq +Vp)/(Ip + Iq)

Vp/Ip

where

Vq = word form (verbaliser dimension)

Iq = word form (Imager dimension)

Ip = picture form (Imager dimension)

Vp = picture form (Verbaliser dimension)

Examples of four of the test items, two from each form is given next:

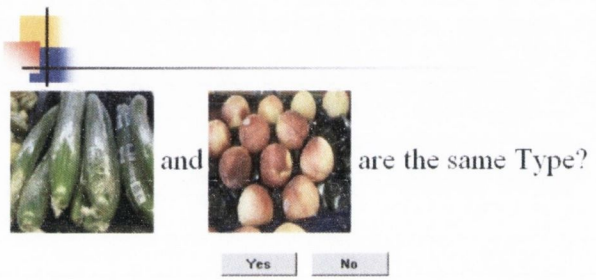


Fig 6.6 An example of a test item (picture form) which assesses the Verbaliser dimension

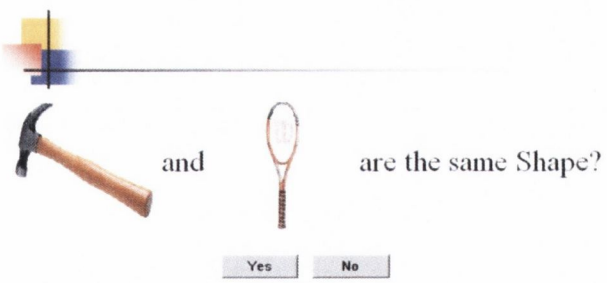


Fig 6.7 An example of a test item (picture form) which assesses the imagery dimension

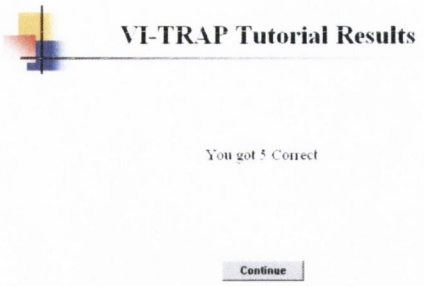


Fig 6.8 An example of the Tutorial results screen

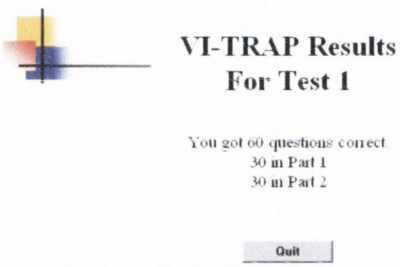


Fig 6.9 An example of the overall results screen

6.2 Implementation

The VItrap, designed by the author, was implemented using Authorware 6.5. The design was coded by a final year B.A. (Mod), student, Declan Murphy, with initial coding by the author of this thesis. Later another program, Filereader, was developed, by postgraduate student Colm Moore in Authorware (Authorware v6.5 (2000)), to help present the raw test files in a more suitable form for analysis. This implementation is also discussed in Redmond, Parkinson & Mullally (2004b).

6.3 Studies to Estimate VItrap Reliability

Two studies were conducted to examine the test-retest reliability of VItrap. A number of further studies to examine the validity of VItrap are discussed in Chapters 7, 8 and 9.

6.3.1 Study 6.1 To test the reliability of VItrap

Sample

A total of one hundred and fifty four students were paid to participate in this study. They were mainly from the disciplines of Computer Science and Engineering at Trinity College, Dublin. This sample consisted of a mixture of first year students to final year. The sample also contained 31 non-native English speakers.

Procedure

The test sessions were held in the Computer Science laboratories. The students completed the tests under conditions similar to that outlined by Riding (generally under relaxed conditions all students were tested in the same types of room). The test was published to Cd-Rom and the results file saved to a 3.5 inch floppy disk. The results were analyzed using Datadesk v6.0.

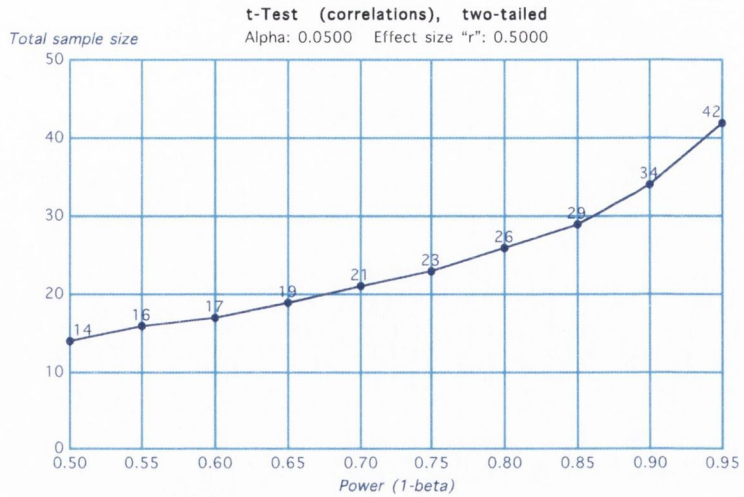
Power

Using G*power (http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/) it was possible to estimate the power of this study at 86.5% as follows:

A priori analysis for "t-Test (correlations)", two-tailed:

Alpha: 0.0500
Power (1-beta): 0.8500
Effect size "r": 0.5000
Total sample size: 29
Actual power: 0.8501
Critical value: t(27) = 2.0518
Delta: 3.1091

Adequate power is usually considered to be 80% or above (Furlong, Lovelace and Lovelace (2000), Cohen (1988)). This illustrated by Fig 6.10.



Note: Accuracy mode calculation.

Fig 6.10 Sample size vs. Power for test retest of VItrap (Study 6.1) using G*Power.

6.3.2 Results of Study 6.1

The sample size n for the Test Retest sessions was 154. The Pearson moment correlation was calculated for the different ratios for Session A (Test) versus Session B (Retest).

These results indicate that there is a very slight, negligible difference between the test retest values calculated for the ratio VQ/IQ and the ratio VQ+VP/IQ+IP. As a consequence, the picture images did not contribute anything useful to the VItrap test. Therefore the rest of the study concentrates solely on Vq/Iq and calculations were done using the 60 word test items while the sixty picture test item results were ignored.

On both sessions for VItrap, the internal reliabilities were quite acceptable, r = 0.88. These results have also been discussed in Parkinson, Mullally & Redmond (2005). Summary statistics for the VItrap may be found in Appendix 13-6.6.F.

Table 6.1: Test-retest Correlation of Session A vs. Session B - Study 6.1

		Session A (Test)	Session A (Test)
	Ratios	Vq/Iq	(Vq + Vp)/(Iq + Ip)
Session B (Retest)	Vq/Iq	0.316	
Session B (Retest)	(Vq + Vp)/(Iq + Ip)		0.314

6.4 Study 6.2 Second test retest of the reliability of VItrap (new sample)

Procedure:

The test was given to a group of Final Year Information System students (n = 40). The students were naïve as to how the test functioned. They were instructed to complete the test in relaxed conditions. Six weeks later they were given the test to do again. The average age of the sample was twenty-eight years of age. The tests ran without difficulties. There was no reason to omit any of the data sets. The test items, which assessed the Verbaliser and Imager dimensions in word form, were exactly the same as in the first test retest study already described above.

6.4.1 Results of Study 6.2

Table 6.2 The test retest reliability of VItrap.

	Session A test
Session B Test	0.306

This is a second test retest of VItrap. The sample size was smaller and the test retest interval greater than in the previous Study 6.1. Both test retest correlations compare favorably at about 0.31. This indicates that VItrap is very stable, and is more stable when compared with Ridings test retest of the V-I dimension as assessed by the V-I dimension. The remaining issue is how can the reliability be improved to get it close to 0.7.

6.5 Study 6.3 Investigates the Effect of Incorrectly Answering Test Items on Subsequent Test Item Inspection Times

Sample

The sample size used was n = 154 (the same sample used in Study 6.1). The test data gathered in Study 6.1 during both Sessions A (test) and B (retest) were used to investigate the effect of incorrectly answering test items on subsequent test item inspection times.

Statistical Analysis

A Repeated Measures model was applied using Data Desk V6.

6.5.1 Session 6.3.A (Test Results)

A linear repeated measures model being applied to investigate the effect of incorrect/correct questions on subsequent inspection times to questions. Another factor examined by this model is the effect of question type on subsequent inspection times. Where necessary post hoc tests are included to investigate a significant p value in the general analysis of variance table.

Table 6.3 ANOVA Analysis for Repeated Measures - Session 6.3.A

General Results

18480 total cases

ANOVA

Analysis of Variance For **LogTSA**
No Selector

Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Const	1	2285.48	2285.48	82890	≤ 0.0001
LCt	1	13.573	13.573	492.27	≤ 0.0001
PRN	153	146.88	0.960001	34.817	≤ 0.0001
QTy	3	174.642	58.214	2111.3	≤ 0.0001
Error	18322	505.181	0.0275724		
Total	18479	853.527			

Results for factor LCt

6.5.2 Session 6.3.B (Retest Results).

A linear repeated measures model being applied to investigate the effect of incorrect/correct questions on subsequent inspection times to questions. Another factor examined by this model is the effect of question type on subsequent inspection times. Where necessary post hoc tests are included to investigate a significant p value in the general analysis of variance table.

Table 6.4. ANOVA Analysis for Repeated Measures - Session 6.3.B

DESIGN

Dependent variables

Factors

No Modifications

RESULTS

General Results

18480 total cases

ANOVA

Analysis of Variance For **LTimeSB**
No Selector

Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Const	1	1219.9	1219.9	46037	≤ 0.0001
QTy	3	172.86	57.62	2174.5	≤ 0.0001
Prn	153	122.838	0.802864	30.299	≤ 0.0001
LCB	1	9.34717	9.34717	352.75	≤ 0.0001
Error	18322	485.582	0.0264983		
Total	18479	796.944			

Results for factor QTy

6.5.3 Results of Study 6.3

Repeated measures means that the same subject is being measured multiple times. Thus, in the analysis explicit account must be taken of the subject. This is done in the ANOVA by putting the PrN (Person) into the model; this is a factor that accounts for the average effect for each person. Such an analysis is called a repeated measures design. From Table 6.3 it is noted that the statistically significant variables on the repeated measures analysis are LCt (Lag Correct Time) and QTy (Question Type). Thus, whether a subject gets the last question correct has an effect on the next question response time. Similarly the Question Type has an impact on average response time.

In terms of the size of this effect, since the response is the log of time, the coefficient represents the proportionate delay (gain) for a given factor. To get this, one needs to exponentiate the coefficient (and for a confidence interval the coefficient +- 1.96*standard error). For example, for IP exp(-0.1294) = 0.88, i.e. 88% with the 95%CI of (0.875,0.882).

So, if you had the last question wrong, the time takes 20% longer - specifically it takes 121% of the average - thus 1.21 +/ - 95%CI (1.19,1.23). For greater detail, consult Appendix 13-6.7.G, Tables 6.7.G.1 and 6.7.G.2.

6.5.4 Conclusions of Study 6.3

This analysis clearly demonstrates that feedback from answering test items incorrectly significantly delays subsequent test item responses.

6.6 Correlation between means of test items times Session One (Test) versus Session Two (Retest).

Further results are available in Appendices 9-6.2.B, 10-6.3.C and 11-6.4.D.

Table 6.5 Test Retest Correlations between mean inspection times of Vq, Vp, Ip and Iq

		Session B	Session B	Session B	Session B
		Vq	Vp	Ip	Iq
Session A	Vq	0.67			
Session A	Vp		0.37		
Session A	Ip			0.42	
Session A	Iq				0.62

Table 6.5 shows the highish correlation between test items in text (word) form and a lowish correlation between test items in picture form.

6.6.1 Reliability of VItrap when non-English speakers were removed.

The test-retest correlation for Vq/Iq was 0.33 which indicates that the reliability improved to 0.33 from 0.32 when non-native English were removed from the test sample which indicates the removal of confusion due to lack of knowledge of English. This also indicates the importance of sampling from a homogenous group.

6.6.2 Reliability for VItrap when those who got 6 questions or more incorrect were removed.

Reliability for VItrap in Session A (Test) and Session B (Retest) for those in the sample who have 95% or more of the questions correct in Session A and Session B. This analysis involved the removal of five individuals who got six questions or more incorrect. The maximum number of questions responded to incorrectly was six out of one hundred and twenty questions. The test-retest correlation for Vq/Iq was 0.32, which indicates that the reliability did not improve from 0.32 when those giving an excess of incorrect answers were removed from the test

sample, which is expected when only removing five people from a large sample of 154. There was no obvious need to remove any other outliers from any of the homogenous samples.

6.7 Summary of work completed by researchers working on the assessment of cognitive style.

Table 6.6: Summarizing related work on Riding's test of the V-I dimension

Authors	Test	Sample Size n	Test – retest interval	No. of Test Items N	Test–retest reliability r	Internal reliability	Ratio VQ/IQ	Ratio ***
Redmond et al. (2002)	CSA	51	14 days		- 0.19	Not possible		
Parkinson et al. (2004c)	CSA	27	23 months		0.36	Not possible		
Vitrap Test	Expt 1	154	4 weeks	60	0.32		*	
Vitrap test	Expt 1	154	4 weeks	120	0.31	0.887- S. A 0.886 – S. B		*
Vitrap test	Expt 2	35??	6 weeks	60	0.31		*	
Rezaei & Katz (2004)	CSA	73	1 week	73	0.35	Not possible		
Rezaei & Katz (2004)	CSA	36	1 month		0.30	Not possible		
Peterson et al. (2003a)	**CSA	50	8.5		0.20			
Expt 1		50	8.5		0.30			
Expt 2		100	8.5		0.09			
Extended **CSA		50	8.5		0.27			
Extended **CSA - A and **CSA - B		50	8.5 days		0.17	0.025		
VICS Test (2005a)	Expt 1	50	1 week	232	0.56	0.720		*
VICS Test v2 (2005a)	Expt 2	100	1 week	232	0.55	0.885		*
Peterson et al. (2003a)	**CSA	50	8.5		0.20			

Table 6.6.1: Summarizing related work on Riding's test of the V-I dimension.

* indicating presence

**CSA Peterson's approximate model of CSA.

*** Ratio – median ratio (Peterson) - verbaliser test items (word and image forms)/imager test items (word and image test items)

6.8 Discussion

As mentioned previously the Spearman Browne formula was used to calculate the required number of test items to improve the reliability of a proposed measure of the V-I dimension. It indicated that a very large number of test items was required to improve reliability of the CSA. This proved impractical due to

the probability of user fatigue; similarly this problem would arise with VItrap (Vid. Kline (1993)).

These results show that the test retest reliability of VItrap (r = 0.31 - 0.32) is an improvement on Riding's result (r = -0.19 - 0.36) for the V-I dimension over the same test retest interval (Parkinson, Mullally & Redmond (2005)). These results indicate that the V-I dimension can be assessed with greater stability with VItrap than Riding's CSA. Having identified and correcting the shortcomings of Riding's assessment of the V-I dimension, the retest values still fell short of Kline's r = 0.7 criterion. The test retest values increase marginally when they were corrected for non-native English speakers. The test retest values also marginally increase if those with 95%, or more, test items correct are corrected for in either Session A or B.

The internal reliabilities for both sessions were within acceptable limits. This is verified when Session B results are examined. The results clearly demonstrate that there was no learning affect between sessions for each of the test sections. There was a practice affect however for each of the four sections Session A versus Session B.

Redmond, Mullally & Parkinson (2002) reported on the possible negative consequence of getting a test item incorrect on the next test item to be answered in the CSA. The results indicate that getting feedback about getting a test item incorrect does adversely affect the inspection time recorded for the next question in Session A and B.

The other ratios did not provide any great insight as to how better the Verbaliser-Imager dimension could be assessed. The pertinent issue is what do the other ratios, particularly combinations of word and image forms, actually measure anyhow?

Attempting to explain why the test retest was not more reliable is a difficult task. Lohman (1999) has referred to the use of a ratio as not being particularly useful since it may give rise to inaccuracies due to the compounding of errors both in the numerator and denominator.

Peterson et al. (2005a) report on the development of a new test of the V-I dimension - VICS. This test assessed the V-I dimension using both picture and word forms. The test contained two hundred and thirty two test stimuli, half of which were designed to assess the Verbaliser dimension, the other half the Imager dimension. The test objects were selected from Rossion and Poutois' (2001) coloured images set (Peterson (2003)). The test items were presented in a pseudo-random order contrasting with lack of randomization in VItrap. The testee had three options when responding to the Verbal and Imagery sections when completing the VICS.

Peterson et al. (2005a) calculate the V-I ratio in the following manner: "this study used a method similar to the CSA, namely a ratio of each participant's median reaction time (RT) on the verbal section overall with each participant's median (RT) on the imagery test section overall (verbal / imagery ratio)". In other words the ratio obtained included (verbal (picture and word forms)/ imager (picture and word forms)). There is no justification presented for calculating the ratio in this manner, nor indeed have the authors provided any basis for doing so with reference to the current neuropsychological literature. Peterson didn't report on the verbal word form/ imagery word form ratio.

The reliability values for VItrap (Vid. Tables 6.1 & 6.2) show that there was a slightly higher ratio for the average inspection time for the word forms versus the pictorial forms.

Some concerns relating to the design of Peterson's test may be that it has a very short test retest interval, and the fact that testees were asked to imagine whether an object was bigger than another when in actual fact it looked quite similar in the image form. This is potentially confusing and would have been alleviated by the inclusion of a scale or a more realistic scaling of the image.

Peterson (2005a) does not indicate for whom this test might be intended but children would find the Imagery section of the test (picture form) quite confusing. As presented in her paper (2005a) one might think that the screw was bigger than a toaster; this would probably be confusing to a child. There may also be too

many test items, which would give rise to fatigue due to test length. Particularly so if this test were intended for sample populations such as school children as has typically been used by Riding in his studies.

Peterson (2005a) reports on a second experiment, which examines a slightly different form of the test. Her test has been altered so it is therefore a different test. The test retest interval is different with a sample size of n = 100. Eleven changes were made to this newer version. The test retest with this updated version is r = 0.55 and the internal consistency increased significantly from 0.720 to 0.885. The authors of this test state, "that in a total sample of 150 participants, the VICS test is internally consistent and moderately reliable".

In reality, in her test, the objects in both image and word forms are used to assess both the Verbal and Imagery dimensions of the V-I dimension, the crucial difference being that the test items used to distinguish between a Verbaliser and Imager were different. Peterson (2005a) summarizes the results as follows: "internal consistency of the crucial verbal/ imager was high and the verbal/imagery ratios at retest on the VICS was stable across about a week". The sample size was n = 50, r = 0.564, and the internal consistency at Session One was r = 0.720.

Peterson has successfully shown that it is possible to design a somewhat reliable test using word and picture test items. However her test may be limited due to user fatigue resulting from a large number of test items.

One of the results of this thesis demonstrates that image forms are processed a lot faster than the word forms. This, while not being the focal point of this thesis, is an interesting result in itself (Appendix 9-6.2.B). This has reported upon by Fraisse (1968), Paivio & Begg (1974), Potter & Fulconar (1975) and Smith & Magee (1980).

Another interesting finding is that the word forms of the Verbaliser and Imager test items are more stable over the test retest interval when compared with the image forms, which are supposedly evaluating these dimensions also.

However the results presented here in this thesis demonstrate that both ratios are practically equivalent. This suggests that perhaps test item is what is central to the assessment of the V-I dimension. These are issues, which need to be tested with larger sample sizes in order to investigate whether these findings can be replicated.

In assessing the design of this test as an enhancement of Riding's assessment of the V-I dimension the following question may be considered: could the design of the test have been improved upon? The selection of the test items included in the VItrap could have been more statistically based. This would have resulted in perhaps better choice of test item from the test item bank. If a normalized set of items were available in digitized image forms similar to those used by Peterson et al. (2005a), this may have leant greater robustness to the test. An interesting suggestion would have been to use different test items to assess both of the continuums. These results also reveal the effect of a negative response on subsequent test items inspection time, so should a version be designed which omits feedback?

One of the limitations of this study is that the results are not applicable to the population as a whole. The VItrap was tested on Computer Science information and Engineering students. The sample was reduced, by removing the data of non-native English speakers, and when the test retest reliabilities were calculated for both groups there was an improvement in the test retest reliability from 0.32 to 0.33. Another possible limitation is perhaps that those of a particular cognitive style may be attracted to the Computer Science and Engineering disciplines.

What is lacking, in this specific area, is a benchmark test against which a new test such as VItrap can be compared.

While there has been renewed interest in this long established area, all these attempts have failed to assess the V-I dimension in a satisfactory manner. The critical issue now becomes the question as to whether there are other alternative paradigms, to self-reporting questionnaires and/or inspection time, worth investigating?

6.9 Improving The Test

The VItrap could be improved by using less ambiguous and sharper images; perhaps a more rigorous selection of test items using a statistical approach would help and also perhaps a neurophysiological approach could help in selecting test items. Better images would have given a better incremental improvement from the picture test items.

Peterson (2005a) has produced a V-I test VICS, which has a test-retest reliability of 0.56. However this test is about four times longer than Riding's so it unlikely to be of practical use due to testee fatigue. However VICS shows that it possible using a inspection time test to approach Kline's desired figure of 0.7.

6.10 Conclusions

In this Chapter, a test (VItrap) was described which was an improvement upon Riding's V-I. Riding's test used 48 test items, while VItrap uses 60, to test the V-I dimension. (The picture images did not contribute anything useful to the VItrap test so therefore, the rest of the study concentrates solely on Vq/Iq ratio and ignores the 60 picture test items). The test-retest reliability estimated using two different samples is much more stable than Riding's V-I. However the estimated test-retest reliability of 0.32 falls far short of Kline's recommended value of $r = 0.7$. It would be necessary to improve VItrap in ways as suggested above.

Riding's comment that it is necessary to look at other approaches other than reaction times should be explored (Riding (2003)). Chapter 7 will explore the validity of the VItrap, assuming that the VItrap reliability may be improved dramatically in the near future.

Chapter 7:

Validation of the VItrap

7.0 Introduction

The approach used to validate this new test of cognitive style, VItrap, is partly influenced by Riding's rationale for the validity of his CSA. Riding believes that the validity of a new test of cognitive style is best tested by investigating whether the test is independent of personality, independent of intelligence, related to observable behaviour and has an underlying neurophysiological basis (Riding (2000)). Ideally the validity of a new test of cognitive style should be investigated before implementation in the human computer interface so that individual differences can be accommodated. In practice proving the validity may be an iterative process building up more evidence of the test's validity over time.

This Chapter therefore examines:

a. whether VItrap is independent of IQ as assessed by Raven's Matrices APM (II) and Mill Hill Inventory
b. if VItrap is independent of Personality measures such as Eysenck's EPI and Speilberger's STAI
c. whether VItrap can be used to satisfactorily predict the learning style classification as assessed by Felder's Index of Learning Styles (ILS) i.e. either Verbal Learner or Visual Learner (Felder & Soloman (2005), Felder & Spurlin (2005)) and Sadler-Smith's Instructional Preference Questionnaire (Sadler-Smith & Riding (1999).

Riding did not consider Concurrent Validity as a necessary condition of test validation or External Convergent Validity of the CSA. The last study in this Chapter examines the Concurrent Validity of VItrap by correlating it with Mark's VVIQ.

7.1 What is meant by the validation of a psychometric test?

It is important in assessing the validity of a test of cognitive style to consider the psychometric properties of tools that are being used to assess this validation. As has been discussed in Chapter 5, there are dangers in blindly accepting the psychometric properties of test instruments based on the work of others. In

assessing the validity of VItrap, it was felt necessary to review the validity and reliability of each tool that would be used in validating VItrap. Psychometric tools assessed in this Chapter are: Raven's Matrices (Court & Raven (1975?)), the Mill Hill Inventory (Court & Raven (1975?)), State Trait Anxiety Inventory (Spielberger (1977)), Felder's ILS (Felder & Soloman (2005)), Mark's VVIQ (Marks (1989)) and Sadler-Smith's Learning Preference Questionnaire (Sadler-Smith & Riding (1999)).

In this Chapter VItrap's independence of intelligence is examined by correlating it with Raven's Matrices. Raven's Matrices were chosen because they are well-established intelligence tests. Raven's Matrices have been known to be valid and reliable and also measure fluid intelligence (Fluid Intelligence is the basic reasoning capacity of the brain, dependent upon (though not, of course, entirely) its neurological structure. It was thus expected to have, and was shown to have, a considerable genetic determination (Cattell (1982)). The VItrap's independence of intelligence is also examined using the The Mill Hill Inventory, also used as a test of IQ, as it too has been shown to be reliable and valid. The Mill Hill test measures crystallised intelligence as assessed by language comprehension, something a little different from Raven's Matrices (Crystallised Intelligence represents the investment of fluid intelligence in the skills valued in a culture. In most Western cultures, at least, education is essentially verbal and thus we would expect crystallised intelligence to be predominantly verbal (Cattell (1971)).

VItrap is also examined in this Chapter with regard to its association with personality using Speilberger's State Trait Anxiety Inventory (STAI), which assesses the Anxiety facet of personality (Spielberger (1977)). STAI has been shown to be independent of Intelligence and it has been extensively validated. VItrap is also correlated with Eysenck's EPI which assesses two dimensions of personality (Eysenck & Eysenck (1975)).

Curry's Model is an attempt to unify different style models into one three-layer model (Curry (1983)). The model is somewhat dated and in light of the Coffield Report (2004) the validity and reliability of the model probably require more extensive

investigation. The VItrap test would be placed in Curry's three layered style model in the innermost layer (cognitive personality). This layer deals with cognition-centred and personality-based models. Riding placed his CSA, as did Peterson her VICS model (Peterson (2005a)), in this layer. As VItrap assesses a similar dimension i.e. the V - I continuum this appears to be the appropriate spot.

Whether VItrap influences an individual's observable behaviour is assessed by attempting to correlate it with measures of learning preference, such as Sadler-Smith's Instructional Preference Questionnaire. This psychometric instrument could arguably be placed in Curry's outer layer (as discussed in Chapter 3).

The VItrap is next correlated with VVIQ in order to examine its external convergent validity i.e. VVIQ is used as a benchmark test for VItrap in the absence of any more suitable test. The purpose of the VVIQ is to assess the self-reporting vividness capability of participants.

Riding (2000) does not include any reference to reliability of any kind when describing the necessary conditions for style test validity. It is this thesis author's contention that a test can be reliable and not valid but never valid and not reliable. Riding has not referred to his perspective on the relationship between test validity and reliability until very recently and this view has been disputed (Coffield (2004), Parkinson, Mullally and Redmond (2004c)).

Riding (2000) has stated that a test of cognitive style should be independent of intelligence and independent of personality. Additionally he said that cognitive style must relate to observable behavior. In testing his CSA he checked whether his CSA was independent of intelligence (it was), personality (generally it was, but with some caveats, discussed later), related to observable behaviour (construct validity), (in general he found it was, but his findings are questionable (Vid. Chapter 5)).

This approach of Riding to establishing the validity of a test of cognitive style had some merit and was worth pursuing. His approach was limited, however, as it didn't attempt to establish

the external divergence or convergence of the test relative to any other recognised test. However he may have been hampered because there were no other suitable tests available particularly for the V - I dimension. Perhaps the only suitable test available at the time might have been Richardson's VVQ (Richardson (1977)). However, Riding did not publish any correlations of other tests with the CSA.

More recently, Mayer and Massa (2003) found there was little correlation between Riding's CSA and Richardson's VVQ (Vid. Chapter 5).

In this Chapter the VItrap is examined as to its independence of intelligence and personality. The correlations between VItrap and some other psychometric tests were measured: VVIQ, Felder's LSI, and Sadler-Smith's Learning Preference questionnaire.

7.2 Cognitive Style and Intelligence: a Brief Review

This section presents a brief overview of the relationship between some cognitive style and intelligence. This is necessary as it should provide some empirical evidence demonstrating that some cognitive styles are reported as being independent of intelligence (as would be expected since, by definition, cognitive style should be independent of intelligence as discussed in Chapter 3). In this light the following cognitive styles were considered worthy of review; a review which should lead to a justification of whether VItrap is independent of intelligence. The Group Embedded Figures Test (GEFT), KAI and the CSA are examined with respect to their independence of intelligence to serve as examples of cognitive styles that might be considered to be independent of intelligence.

Goldstein and Blackman (1978) have raised concerns about the independence of Witkin's field independence / dependence cognitive style from intelligence. They report a correlation of about $r = 0.4$. McKenna (1984) also maintained that the Geft was not a measure of cognitive style as it correlated with ability tests.

Some other researchers also argue that the construct of field dependence is valid but the difficulty lies with the Geft itself as a

psychometric instrument (Parkinson & Redmond (2001), Goldstein & Blackman (1978) as was discussed in Chapter 3.

Kirton's KAI is a measure of the Adaptor – Innovator continuum (Kirton (1982)). Kirton's KAI has been shown to be independent of intelligence (Clapp (1993), Kirton (2003)).

Riding, in two studies demonstrates the independence of the CSA of intelligence (Riding and Pearson (1994) (Riding & Agrell (1997)). Their findings demonstrated that the CSA didn't correlate significantly with the different subtests of the British Ability Scales. In the Riding & Pearson study the sample size was quite robust (n = 119), the population being 12 to 13 year old school children. In another study Riding and Agrell (1997) found no significant relationship between the CSA and the Canadian test of cognitive skills. The sample size in this case consisted of 205 14 to 16 year old school attending students. Peterson (2003) demonstrated the independence of her test of cognitive style against another measure of cognitive ability, the Elkstrom cognitive battery (Elkstrom (1976)), using sample sizes of approximately 100 University students; (this was discussed previously in Chapter 3).

Riding suggested in his CSA Manual the possibility of using Raven's Matrices and the Mill Hill Inventory as measures of intelligence (Riding (1991c)). This suggestion is followed up in this thesis both for the CSA and VItrap and the results are given in Appendix 7-5.2.B for the CSA and later in this Chapter for VItrap.

7.2.1 Study 7.1 - VItrap vs. Intelligence as Estimated by Raven's Matrices APM2 and also the Mill Hill Inventory

To investigate the relationship between VItrap and Intelligence as measured by Raven's Matrices APM 2.

Description of Sample

A sample of 154 students, mainly Computer science and Engineering, completed these tests.

How the test was administered and scored

The students completed these tests in their own time in a quiet room. Raven's Advanced Progressive Matrices, along with the

Mill Hill Inventory, are well-established tests of intelligence. The test-retest of both of these tests has been thoroughly investigated (Court & Raven (1975)). The test-retest reliabilities for the APM is about 0.95 and for the Mill Hill Inventory range from 0.87 to 0.90. The APM 2 consists of thirty-eight test items and was scored automatically using carbonized sheets. The multiple-choice form of the Mill Hill Inventory was used; each test item required the testee to tick a synonym for the seventy six words presented. Again the responses were scored automatically using carbonized sheets.

7.2.2 Results of Study 7.1

Table 7.1 Pearson Product correlation for VItrap Session A (Test) and Session B (Retest) with Raven's APM 2

	Raven's APM2	Mill Hill
VItrap Session A	0.038	-0.029
VItrap Session B	0.005	0.034

The results, given in Table 7.1, demonstrate that the new test of the V-I dimension is independent of intelligence as assessed by APM 2, with $r = 0.038$ and $r = 0.005$.

These results clearly show that VItrap is independent of intelligence as measured by the Mill Hill inventory. The correlation coefficient varies between $r = 0.034$ and $r = -0.029$.

7.2.3 Conclusions of Study 7.1

These results indicate that Vitrap was independent of intelligence as estimated by Raven's Matrices APM2 and also by the Mill Hill Inventory, both measures of intelligence. The results of Study 7.1 are presented in Parkinson, Redmond & Mullally (2006) in greater detail.

7.3 Study 7.2 Is VItrap Independent of Personality as Assessed by Eysenck's EPI and Spielberger's STAI?

7.3.1 Study 7.2.1 To investigate whether VItrap is independent of the EPI (Eysenck & Eysenck (1975)).

Riding has stated that the Cognitive Style is independent of personality; therefore it is necessary to demonstrate that the VItrap is independent of Eysenck's EPI.

Apparatus

A Pearson Product moment was calculated using Datadesk software.

Method

The EPI was administered to 154 Computer Science and Engineering students. They completed this test in approximately three minutes. Any student who scored greater than three on the L (Lie) scale rendered the test invalid; subsequently these were omitted from the calculation of correlation (n = 154 with 24 being omitted due to L scale correction). This questionnaire measures three personality variables: extraversion (E), neuroticism (N) and psychoticism (P). There is also an L scale i.e. a scale to screen out those giving socially desirable responses. Kline (2000) reports test-retest reliabilities varying from 0.7 to 0.9 for the three scales (E, N, P). These values are highly satisfactory. Kline also reports internal consistencies ranging from "0.7 to 0.8; the exception being P for normal females" which is slightly below this. He also reports, that "the validity of these scales is the best supported of any personality measure".

7.3.2 Results (Shown in Table 7.2)

Table 7.2 VItrap vs. Personality as Assessed by EPI and STAI

	E (Extraversion)	N (Neuroticism)	State	Trait
VItrap Session A	-0.051	-0.126	-0.109	-0.040
VItrap Session B	0.026	-0.029	0.088	-0.032

Comment: The E and N labels, denote the Extraversion and Neurotic Scales on Eysenck's EPI. The VQ/IQSA denotes VItrap Session A scores similarly VQ/QSB denotes VItrap Session B scores. A Pearson Product moment correlation was calculated

and no correlations of note are evident in this table between VItrap Session A or B with the E and P scales of Eysenck's EPI.

7.3.3 Study 7.2.2 To investigate whether VItrap is independent of measures of personality as assessed by the STAI

Introduction

The State Trait Anxiety inventory was developed by Charles Spielberger (Spielberger (1977). Since then there have been over 2000 studies reported in the literature. This self-reporting questionnaire evaluates an individual's current state of anxiety and general state of anxiety. Anxiety as defined by the Penguin Dictionary of Psychology (Reber & Reber (2001)) is "a vague unpleasant emotional state with qualities of apprehension, dread, distress and uneasiness. It can also be used to "describe relatively stable individual differences in anxiety proneness as a personality trait" (Spielberg (1977)). In this study the independence of VITRAP with respect to the STAI is investigated.

Sample

The sample initially comprised one hundred and fifty five students mainly from the disciplines of computer science and engineering. The sample included a student from psychology and one from microbiology. They were paid to participate in this study. Some students were omitted from the analysis, as they had not fully completed the STAI.

Materials

VItrap

All students completed the VItrap as previously described.

The State Trait Anxiety Inventory

The STAI is a self-evaluatative questionnaire. It consists of two forms. The first form STAI Y-1 comprises twenty questions. This set of questions evaluates how the individual feels currently. It can also be used to assess how an individual felt at a particular moment in time, given a particular scenario real or imaginary. The characteristic feelings assessed by the STAI are "apprehension, tension, nervousness and worry" (Speilberger (1977)).

The second form STAI Y-2 is printed on the back of form 1. It too contains twenty questions; these assess the individual's general state of anxiety. State anxiety "may also indicate individual differences in the frequency and intensity with which anxiety states have been manifest in the past " (Speilberger (1977)). Alternatively it evaluates "relatively stable individual differences in anxiety proneness" (Speilberger (1977)). Each of the twenty test items on each form is scored from 1 to 4. However for anxiety absent test items the scoring is reversed on both State and Trait scales. The final score for each of the forms S and T are calculated by adding each of the individual test item scores. The scores for each of these scales can range from twenty to eighty.

The STAI's psychometric properties have been thoroughly investigated and reported upon in the Manual, which accompanies the test instrument. Internal reliabilities range from $r = 0.9$ to 0.97 for both the Trait and State scales. The test retest reliabilities for the Trait scale ranges from $r = 0.73$ to $r = 0.84$ and for the S scale $r = 0.16$ to $r = 0.62$. Speilberger (1977) reports, that "relatively low stability coefficients were expected for the State scale because a valid measure of the state anxiety should reflect the unique situational factors at the time of testing". Both the State and Trait scales are also correlated at $r = 0.6$ (Speilberger (1977)).

The validity of the STAI has been well documented and a summary of this related research is presented on pages 34 - 44 of the test manual. This research describes studies that provide evidence as to the convergent and discriminant validity of the STAI. This review also indicates that state and trait anxiety are largely independent of scholastic achievement. Spielberger states, that "it would appear that the STAI scales are essentially unrelated to aptitude and achievement for college students" (Speilberger (1977)).

Administration

The STAI was given to 154 computer science and engineering students in a manner outlined in the test manual. As recommended the STAI-1 was administered before STAI Y-2. The students have an unlimited amount of time in which to complete this test; the average time taken by University students is of the order of 10 minutes.

7.3.4 Results & Conclusion of Studies 7.2.1 & 7.2.2

Table 7.3 VItrap vs. Personality as Assessed by EPI and STAI

	E (Extraversion)	N (Neuroticism)	State	Trait
VItrap Session A	-0.051	-0.126	-0.109	-0.040
VItrap Session B	0.026	-0.029	0.088	-0.032

The VITRAP is independent of Eysenck's measure of personality given the poor correlations in Table 7.2.

These results indicate that the VITRAP does not correlate with either State nor indeed Trait Anxiety as assessed by the STAI. It is also interesting to note that the correlation between STAI – Y1 and STAI – Y2 is r = 0.588 in this study and these values agree with typical values reported in the literature.

In some cases where individuals didn't complete the questionnaire they were not included in the statistical analysis. The STAI, while being a self-report questionnaire, was reported as being highly reliable. It is arguably deficient as it lacks a Lie scale. An individual may not like to acknowledge anxiety and as such this dishonesty is difficult to detect. This type of response may indeed be symptomatic of a higher degree of anxiety experienced by the individual when compared with that recorded by the STAI. These results are also discussed in Parkinson, Mullally and Redmond (2006) in greater detail.

7.3.5 Overall Section Conclusions on Personality & the VItrap

A real question of significance is whether cognitive style is correlated with personality. The VItrap was not correlated with either of two personality measures, the EPI or the STAI. What can be concluded is that VItrap measure of the V-I dimension is independent of these personality measures. Peterson also demonstrated that her new tests of cognitive style were independent of personality (Peterson et al. (2005b)).

However there is evidence that this may not always be the case. Riding and Dyer (1980) examined the relationship between the Verbaliser - Imager cognitive style assessed using the verbal - imagery code test and personality. Personality was assessed using the Junior Eysenck Personality Inventory. The sample used

was two hundred and fourteen children whose ages ranged between twelve and twelve point nine years of age.

They based their hypothesis that Imagers were Introverts and Verbalisers, Extraverts on earlier studies completed by Riding and Wicks (1978).

Riding and Dyer (1980) report, that "Extraversion was found to be significantly negatively correlated with the Verbaliser - Imagery ratio". They observed that there was "a significant interaction between extraversion and passage type in their effect on recall". They claim that their findings were "consistent with the interpretation that Extraverts are Verbalisers and Introverts are Imagers".

In a later study, using a different means of assessing the Verbaliser - Imager ratio. Riding and Wigley (1997), state that "cognitive style and personality sources are not the same since the correlations between them are approximately zero". This study used a different sample to that used in the previous study described above (n = 340 college students, 148 male, 192 female) and different measures of personality and cognitive style. Cognitive style was measured using the Cognitive Styles Analysis while Personality was measured using a variety of instruments. These personality measures were: the EPQ-R Short Scale, the IVE scale (Eysenck & Eysenck (1991)) and State Trait Anxiety Inventory (Speilberger (1977)).

Isaksen, Lauer & Wilson (2003) designed a study to examine the relationship between Personality type and Cognitive style. They assessed personality type using the Myers Briggs Type Indicator (Briggs-Myers (1990)) while cognitive style was assessed by Kirton's Adaptor - Innovator Inventory. Their study had a sample size of 1,483 participants from around the United States. They found that there was a correlation (r = 0.55) between an Innovative as measured by the KAI and Intuitive as measured by the MBTI. They also report that there was a strong relationship between an Innovative cognitive style and a Perceiving psychological type (r = 0.44). Based on the results, they state that "that there does seem to be some conceptual overlap between measures of psychological type and cognitive style, specifically between the MBTI and KAI and that these relationships held

when separating the sample on the basis of gender". Sternberg (2005) believes that it depends on whether one is coming from a personality or cognitive centrist tradition.

In assessing some of the literature in the area as discussed, and having communicated with some authoritaive authors in the field of cognitive styles, it appears that the link between cognitive style and personality is equivocal (Armstrong (2005), Sternberg (2005)).

7.4 Study 7.3 VItrap vs. Sadler-Smith's Instructional Preference Questionnaire to Affect User Behaviour

7.4.1 Introduction

This investigation is continued using Sadler-Smith's Instructional Preference Questionnaire as it was appropriate to see if his study could be replicated with respect to VItrap. This psychometric test has not been widely reported upon in the literature and evidence as to its reliability is scant. The Coffield Report (2004) however does raise concerns about the validity of Riding and Sadler-Smith's study and asserts "It is a step too far to move from this finding to the recommendation that students should be given what they prefer". He also cites Atkinson's research (Atkinson (1998)) which found that holistic students who were taught by teachers using a collaborative approach obtained poorer grades than any other group. This instrument's internal reliability was calculated and the factor analysis was calculated. Again participants' VItrap scores were correlated with their scores on this psychometric instrument.

Sadler–Smith and Riding (1999) discuss their findings on the relationship between cognitive style and instructional preferences. In their study the sample consisted of 240 business students.

They reported on the design of a questionnaire that assesses an individual's preference for assessment method and instructional methods and media. They presented a hypothesis based on Curry's model (Curry (1983)) and cognitive style. They state that "the present study will argue that the innermost layer (cognitive style), is relatively fixed (Riding 1991) and affects the outer layer (instructional preferences)".

The paper discusses the factor analysis, which was approached in the following manner: principal component analysis, followed by an orthogonal (Varimax rotation). The authors, based on this factor analysis, divided instructional media into two groups "(a) print – based media (handouts, workbooks, text books, journal articles)" and (b) nonprint based media (overhead transparencies, slides, video tapes)". The questionnaire items that composed the assessment method preference category of the questionnaire were similarly analyzed. Resulting from this analysis, assessment was divided into two main categories, formal and informal assessment methods. Formal methods included: examinations, tests and essay questions. Informal methods included multiple-choice questions, short answer type questions, and multiple-choice questions. The internal reliability for the questionnaire is not presented.

Their findings show that students' "overall preferences were for dependent methods using print based media" and assessment by "informal methods". Their results also provide evidence that cognitive style e.g. Wholist – Analytic dimension was related to instructional preference. Wholists, with the exception of female Verbalisers, "expressed the strongest preferences for informal methods by comparison with formal methods".

The authors also report that their results revealed a V - I by Gender effect for informal methods. The authors argue that their findings also "lend support to the onion model put forward by Curry, i.e. that style is an underlying dimension which affects the "outer skin of the onion" (instructional model).

Peterson (2003), examined student instructional preferences using this questionnaire. The objective of this study was to investigate whether individuals as assessed by the VICS would demonstrate an interaction for instructional preferences by cognitive style. This study did not find any significant relationship between cognitive style and preferences. It was reported that the factor analysis obtained in this study was quite similar to that reported by Smith and Riding (1999).

Materials

Sadler-Smith's Instructional Preference Questionnaire.

Method

One hundred and fifty four students were administered Sadler-Smith's and Riding's instructional Questionnaire. They were then administered the VItrap. The internal reliability of the questionnaire was also calculated. It was felt that the factor analysis had already been well substantiated (Sadler-Smith & Riding (1999) and Peterson (2003)).

Aim:

Using Sadler-Smith & Riding's (1999) Instructional Preference Questionnaire and Curry's model, it was hypothesized that cognitive style as assessed by the VItrap would affect the individual's instructional preference within the three different sections.

Table 7.4 Principal Component Analysis of Sadler-Smith's Instructional Preference Questionnaire

Rotated Component Matrix[a]

	Component		
	1	2	3
Lecture	-.182	.009	.644
Tutorial	-.026	.096	.579
Roleplay	.739	-.043	-.379
ODFLearn	-.209	.260	-.598
Grwork	.756	.154	.169
CAL	-.032	.916	-.006
B.games	.500	.684	-.019

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

Table 7.5: This shows a multiple regression model for cognitive style and the different facets of instructional methods section of Sadler-Smith's Instructional Preference Inventory

Dependent variable is: **Cognitive style**
No Selector
130 total cases of which 14 are missing
R squared = 11.4% R squared (adjusted) = 5.7%
s = 0.1492 with 116 - 8 = 108 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	0.309551	7	0.0442216	1.99
Residual	2.40447	108	0.0222636	

Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	0.937005	0.08379	11.2	≤ 0.0001
Game based...	-0.00363975	0.01685	-0.216	0.8294
Computer a...	0.0356299	0.0158	2.25	0.0262
Group work	-0.00799577	0.01325	-0.603	0.5476
Open/distan...	0.0228773	0.01347	1.7	0.0924
Role play	-0.00716786	0.01524	-0.47	0.6390
Tutorial	-0.0181996	0.01335	-1.36	0.1757
Lecture	0.0200557	0.01411	1.42	0.1582

Table 7.6 Factor Analysis for the Media Section of Sadler-Smith's Instructional Preference Inventory

Rotated Component Matrix[a]

	Component		
	1	2	3
Handouts	.041	.784	.074
Workbks	.036	.800	.075
Textbooks	-.135	.301	.703
Journals	.133	-.070	.864
OHT	.669	.309	.261
Slides	.851	.003	.021
Vtapes	.747	-.079	-.145

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 4 iterations.

Table 7.7: Instructional Media Section of Smith's Instructional Preference Questionnaire

Dependent variable is: **Cognitive Style Ratio**
No Selector
130 total cases of which 6 are missing
R squared = 5.4% R squared (adjusted) = -0.3%
s = 0.1595 with 124 - 8 = 116 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	0.168675	7	0.0240965	0.947
Residual	2.95097	116	0.0254394	

Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	1.03064	0.07065	14.6	≤ 0.0001
Video tapes	0.00366476	0.01495	0.245	0.8068
Slides	-0.0173591	0.0179	-0.97	0.3342
Overheads/t...	0.000672295	0.01806	0.0372	0.9704
Journals/art...	-0.021735	0.01571	-1.38	0.1691
Textbooks	0.0121727	0.01598	0.762	0.4478
Workbooks	0.00911292	0.01707	0.534	0.5944
Handouts	0.0226265	0.01635	1.38	0.1690

This shows a multiple regression model for cognitive style and the different facets of the Instructional Media section of Sadler-Smith's Instructional Preference Questionnaire.

7.4.2 Results of Study 7.3

Table 7.8 Factor Analysis for Different Forms of Assessment

Rotated Component Matrix[a]

	Component		
	1	2	3
Examination	.962	-.023	.139
Gassign	-.128	.089	.674
InAssign	.962	-.023	.139
Tests	.598	.131	-.297
EssayQ	.121	-.269	.512
MchoiceQ	.042	.805	.126
ShortQ	.032	.820	-.163
Cstyleratio	.048	.025	.534

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 4 iterations.

Table 7.9: A multiple regression model for cognitive style and the different facets of assessment section of Sadler-Smith's Instructional Preference Questionnaire.

Dependent variable is: **Cognitive styles**
No Selector
129 total cases of which 8 are missing
R squared = 1.7% R squared (adjusted) = -3.5%
s = 0.1622 with 121 - 7 = 114 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	0.0504116	6	0.00840194	0.319
Residual	2.99942	114	0.0263107	

Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	1.01756	0.08188	12.4	≤ 0.0001
Short answ...	-0.00398535	0.01857	-0.215	0.8304
Multiple que...	-0.00246668	0.01604	-0.154	0.8781
Essay quest...	-0.000152641	0.01213	-0.0126	0.9900
Individual a...	-0.0121428	0.01553	-0.782	0.4360
Tests	0.0142459	0.01348	1.06	0.2929
Group assig...	0.00776683	0.0127	0.611	0.5421
Examinations	0	0	•	•

7.4.3 Internal Reliability of Sadler-Smith's Instructional Preference Questionnaire:

Table 7.10 Internal Reliability of Intructional Preference Questionnaire

Reliability Statistics

Cronbach's Alpha	N of Items
.615	21

This is somewhat disappointingly low for an Internal reliability measure. Kline (2000) cites Guildford (1956) and Nunnally (1978) who have stated that certainly "alphas should never drop below 0.7". These results are also presented in (Redmond, Parkinson & Mullally (2006)) in greater detail.

7.4.4 Conclusions of Study 7.3

Each of the three sections in the Sadler-Smith's Instructional Preferences were subjected to a principal components analysis. The criterion for factor extraction was those factors that had an eigenvalue greater than one. The factors were then extracted and subjected to a varimax rotation.

The results of this factor analysis compare well with Sadler-Smith's finding's for the first two of sections in his Instructional Preference Questionnaire. However three factors were extracted in the Assessment method preferences in contrast with Sadler-Smith's factor analysis where they extracted only two.
The internal reliability of the questionnaire was calculated as $r = 0.615$. This had not been reported before. There was no test retest reliabilities reported for this questionnaire in the literature.

The objective of the study was to investigate the relationship between the participants' cognitive styles and their instructional and assessment preferences. This approach was seen as a means of validating this new test of style, VItrap.

Study 7.3 was based on the assumption that cognitive style as measured by VItrap would be placed in the innermost layer of Curry's model. Therefore this innermost layer would affect the Instructional Preferences layer as envisaged by the Curry model. Sadler-Smith's Instructional Preference Questionnaire assessed the individuals' preferences for instructional media, instructional method and assessment method. While Study 7.3 used a different statistical analysis method to Sadler-Smith, no significant results emerged.

The results of the multiple regressions in Tables 7.5, 7.7 and 7.9 demonstrate that, in the case of instructional preferences for learning media and assessment, no significant results were reported. When the preferences for instructional methods were examined only one facet was found to be significant, computer assisted learning ($p < 0.05$). This was an unexpected result that possibly needs further investigation.

These data do not support the Curry model. As style is located in the innermost core of Curry's model it would be expected to affect the outermost layer i.e. instructional preferences. There

are some possible reasons to suggest why this may be the case: a) the moderate reliability of Vitrap, b) the low internal reliability of Sadler-Smith's Questionnaire, c) the concern that the factor analysis was not replicated for the sample used in this study and d) the sample type used in this study was different from the sample type used by Sadler-Smith & Riding in their study.

These results do not provide any evidence that the V-I dimension affected individuals' preferences as assessed by Sadler-Smith's Instructional Preference Questionnaire. The sample size was not divided up by cognitive style grouping as in Sadler-Smith's study. Sadler-Smith did not report any relationship between the V-I dimension and instructional preferences. (He did find that there was a V-I by Gender interaction). It is probable that the students would not have been accustomed to the wide range of methods of instruction and assessment as assessed in this questionnaire, being undergraduate Computer Science and Engineering students.

7.5 Study 7.4 Is There a Relationship between VItrap and Felder & Silverman's ILS?

7.5.1 Introduction to Study 7.4

The aim of this study is to examine whteher an individual's cognitive style as assessed by VItrap affects his/her learning style classification as assessed by Felder & Silverman's Index of Learning Style (ILS) (Felder & Silverman (1988)). The Felder & Silverman ILS could arguably be placed in the second layer of Curry's model as the ILS classifies individuals according to learning style.

7.5.2 Description of the Index of Learning Styles

The questionnaire was developed by Felder and Silverman (1988) and consists of 44 test items. These items assess four distinct scales and classify students as having preferences for one of two categories along four dimensions. This instrument is available in pen and pencil form or it can be completed online. Its primary aim was to assist the design of third level engineering courses to accommodate individual's differences as assessed using this instrument. Felder and Spurlin (2005) explain this in greater detail "to provide guidance to instructors on the diversity of learning styles within their classes and to help them

design instruction that addresses the learning needs of all their students". The second use of the ILS is to "give individuals insights into their possible learning strengths and weaknesses".

Felder and Spurlin (2005) describe the four dimensions, with the first dimension being: "Sensing (concrete thinker, practical, orientated toward facts and procedures) or Intuitive (abstract thinker, innovative, orientated toward theories and underlying meanings)".

The second category describes the Visual – Verbal dimension. This was added to the later version of this measure (Felder and Spurlin (2005)). Felder and Spurlin (2005) describe this category as: Visual (prefers visual representation of presented material, such as pictures, diagrams and flow charts) or Verbal (prefers written and spoken explanations).

The third category describes the Active to Reflective dimension. This continuum is described as "Active (learn by trying things out, enjoying working in groups) or Reflective (learn by thinking things through, prefer working alone, or with a single familiar partner)"

The final category describes the Sequential to Global dimension. They define this by someone being characterized as being sequential (linear thinking process, learn in small incremental steps) or global (holistic thinking process, learns in large leaps). The authors report that the ILS had been validated and its reliability investigated. Test retest reliabilities range from $r = 0.511$ to $r = 0.870$, for the different scales and the Chronbach's alpha exceeds a value of 0.5 for all but one of the subscales. The Sensing – Intuitive dimension of the ILS correlated with the Sequential and Global dimension: the Pearson coefficient is reported as being from $r = 0.32$ to $r = 0.48$.

The authors refer to studies, namely Zywno (2003) and Livesay (2002), which have provided empirical evidence which supports the validity and reliability of this measure.

Materials: Felder's Index of Learning Styles (Felder & Silverman's (1988))

Method: The students filled in Felder and Soloman's Index of Learning Styles (Felder & Soloman (2005).

7.5.3 Results & Conclusions of Study 7.4 on Felder & Silverman's ILS

The Pearson Product - Moment Correlation for Vitrap Session A (Test) was -0.145 and for VItrap Session B (Retest) was -0.209 when correlated with Felder & Soloman's Visualizer - Verbalizer Preference Scale (Felder & Soloman (2005)). Summary statistics for this Study are reported in Appendix 16-7.2.B. These results indicate that the individual's results on the V - I dimension (via VItrap) do not affect his/ her's categorisation as a Visual - Verbal learner as assessed by the ILS. A possible explanation is that the VItrap and the ILS are measuring two very different constructs, namely cognitive style and learning style respectively.

7.6 Study 7.5 Investigating the Convergent Validity of VItrap

To assess the external convergent validity of a new test (Vitrap) of the Verbaliser-Imager dimension it was compared to a long established test – the Vividness of Visual Imagery Questionnaire (VVIQ).

The VVIQ test is questionnaire-based, has a five point Likert scale for each of the sixteen questions presented in the test and estimates how vivid a respondent's imaging capabilities are. It contains four subsections each consisting of a scenario followed by four questions; the testees are asked how easily they can visualize particular aspects of the described scenario.

The VVIQ's reliability has been the subject of mixed reviews in the literature to date. Some high reliability values have been reported along with some low ones. It was felt that the VVIQ might serve as a reasonable benchmark against which to validate the Imager end of the Verbaliser-Imager continuum.

A sample of one hundred and seventy computer science and engineering students completed the VVIQ test and Vitrap ("Test"). The test was investigated as to its internal consistency. Cronbach's alpha was found to be greater than 0.7. A hundred and fifty four of the original cohort of computer science and engineering students were administered the VVIQ and Vitrap about four weeks later ("Retest"). A Pearson Product-Moment

Correlation was calculated for both Sessions A and B for each of the tests and negligible correlation was found between the two measures for Session A and Session B. The test-retest correlation for the VVIQ was estimated to be $r = 0.22$ after approximately four weeks which is much lower than reported in the literature.

7.6.1 Introduction to the VVIQ

The objective was to examine whether there was a meaningful correlation between individuals' scores on the VVIQ and Vitrap, a new test of Verbaliser – Imager dimension (Redmond, Parkinson and Mullally (2004b), (Parkinson, Redmond and Mullally (2005)). This was an attempt to establish the external convergent validity of a new test of the Verbaliser-Imager dimension.

The VVIQ (Marks (1973, 1989)) was developed in 1973 and consists of sixteen test items. A copy of the test is given in Appendix 14-7.1.A. This test is based on a subset of Betts' (1909) questionnaire. It requires that the respondents generate images depending on the different cues presented in each of the four scenarios in the test. They are then asked to rate the cues using a five-point scale. A low score represents a high introspective imagery rating and a high score, a low imager rating. A copy of the test is given in Appendix 15-7.1.A.

The validity of this test has been investigated by a number of researchers (McKelvie (1990), Marks (1989)). Moran (1993) reports that the test-retest correlation of the VVIQ ranges from r = 0.67 to r = 0.87. Parkinson et al. (2005) reported the VVIQ as being of low reliability in a test-retest situation with a sample size of n = 154 of mainly Computer Science and Engineering university students.

Chara and Verplank (1986) report that their research indicated lack of evidence to support the validity of the VVIQ. Moran (1993) in his review of various tests of imagery reports that the VVIQ "is probably the best researched, and most popular, test of imagery currently available". Moran (1993) quotes McKelvie (1990) as saying "part of the difficulty here, as with most imagery tests, is that there is no single criterion ready for the VVIQ to predict".

McKelvie (1995) reviewed approximately two hundred and fifty studies, which concluded that the reliability of this measure was "adequate but not overwhelmingly impressive". The definitions of "imagery" and "vividness" are ambiguous. The concept of vividness as envisaged by various researchers, and its ambiguity in definition, has also added difficulty in attempts to validate this instrument. What meaning do we associate with vividness? More importantly what is the perception of vividness by those individuals that complete the VVIQ?

Benjafield (1997) disagrees with McKelvie's (1995) definition, which insists that different dictionary definitions emphasise vividness "as the degree to which it resembles normal perceptual experience (is lifelike)". Benjafield's definition taken from the Oxford English dictionary is quite different, "vivid" meaning "full of life"; vigorous, active or energetic ... lively or brisk". He refers to the fact that the "Japanese word corresponding to the English word image has several different meanings depending on the context and so the specific meaning being used by an experimenter needs to be spelt out".

7.6.2 Methods

A sample of one hundred and seventy computer science and engineering students completed the VVIQ test and Vitrap ("Test" – Session A). The test was investigated as to its internal consistency using Cronbach's alpha which was found to greater than 0.7. A hundred and fifty four of the original cohort of computer science and engineering students were administered the VVIQ and Vitrap about four weeks later ("Retest" – Session B).

Results Section

The data gathered were analysed using DataDesk v6.0 (DataDesk 2001) and SPSS v.12 (SPSS 2002).

7.6.3 The Test-Retest Reliability of the VVIQ

Table 7.11: Pearson Product-Moment Correlation between the test-retest correlation for the VVIQ .

	VVIQ - Session A	VVIQ - Session B
VVIQ - Session A	1.00	
VVIQ - Session B	0.223	1.00

7.6.4 The Internal Reliability of the VVIQ

Table 7.12: This table presents an internal Cronbach's alpha calculated for the VVIQ Session A values.

Cronbach's Alpha	Number of Items N
0.896	16

Similar values for internal consistency were again obtained four weeks later when the test was taken by the same cohort of students (Parkinson, Mullally & Redmond (2004b)).

7.6.5 The Test-Retest Correlation between Vitrap and VVIQ

Table 7.13: Pearson Product-Moment Correlation between the test-retest correlation for the VVIQ and Vitrap Sessions A and B

	Vitrap – Session A	Vitrap – Session B	VVIQ – Session A
Vitrap – Session A	1.000		
Vitrap – Session B	0.326	1.000	
VVIQ – Session A	-0.005	0.061	1.000

A Pearson Product-Moment Correlation was calculated (Table 7.13) for both Sessions A and B for each of the tests and negligible correlation was found between the two measures for Session A and Session B. This table shows the correlation between the VVIQ and VItrap. The correlation between these two measures is very low indicating little or no association between these two measures.

7.6.6 Factor Analysis

Tables 7.14 and & 7.15 give the factor analysis for the VVIQ.

Table 7.14: Total Variance Explained – Principal Component Analysis - VVIQ

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.354	39.715	39.715	6.354	39.715	39.715	3.069	19.181	19.181
2	1.569	9.806	49.521	1.569	9.806	49.521	2.851	17.817	36.998
3	1.167	7.295	56.816	1.167	7.295	56.816	2.222	13.886	50.884
4	1.086	6.788	63.605	1.086	6.788	63.605	2.035	12.721	63.605
5	.850	5.315	68.920						
6	.698	4.362	73.282						
7	.644	4.025	77.307						
8	.569	3.558	80.865						
9	.538	3.363	84.228						
10	.460	2.877	87.105						
11	.437	2.731	89.836						
12	.407	2.543	92.379						
13	.343	2.145	94.525						
14	.326	2.035	96.560						
15	.312	1.948	98.508						
16	.239	1.492	100.000						

Extraction Method: Principal Component Analysis.

Table 7.15: Kaiser Normalization – Principal Component Analysis

Rotated Component Matrix[a]

	Component			
	1	2	3	4
VAR00002	.770	.195	.233	.065
VAR00003	.825	.176	.159	.180
VAR00004	.737	.239	.103	.005
VAR00005	.665	.127	.069	.356
VAR00006	.183	.367	.450	.237
VAR00007	.142	.305	.391	.495
VAR00008	.164	.232	.085	.809
VAR00009	.191	.234	.150	.779
VAR00010	.374	-.069	.741	.218
VAR00011	.066	.170	.721	.202
VAR00012	.219	.298	.682	-.122
VAR00013	.576	.135	.331	.162
VAR00014	.202	.715	.219	.151
VAR00015	.293	.730	.148	.221
VAR00016	.143	.812	.162	.108
VAR00017	.149	.726	.056	.311

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
[a]. Rotation converged in 5 iterations.

Analysis – VVIQ

The factor analysis in Tables 7.14 ad 7.15 shows that there four principal components corresponding to the four scenarios in the VVIQ.

7.6.7 Conclusions on Study 7.5 on VVIQ

The results indicate that the test-retest correlation value of the VVIQ at r = 0.22 is a lot lower than reported in the literature for this test over a four week test-retest period. McKelvie (1995) reported a range of values much higher than this.

The internal reliability of the VVIQ was estimated, using Cronbach's alpha, as r = 0.90. The VVIQ was also factor analyzed, and four factors (corresponding to the scenarios) emerged explaining most of the variance.

The VVIQ is a self-evaluatory questionnaire and may be limited as a result. Individuals may have the tendency to rate themselves as being clear at seeing "in the mind's eye" as opposed to seeing unclearly. It is plausible that, the more confident an individual is, the more likely he/she would be to rate themselves highly. This has also been commented on by McKelvie (1994). The questions in the VVIQ appear to us to be value laden and definitely not value neutral. The terminology of "first place", "second place" is biasing the expected result.

The questionnaire may also be dependent on long term memory and, perhaps, is limited given the scenarios which it presents to the individual. The VVIQ does not provide accurate definition of the term "Vividness" so individuals taking the test may have varied in their interpretation of this term. It would have been an interesting dimension to this study if some of the sample had been interviewed beforehand, as to their own understanding of the terms imagery and vividness. In this respect greater instruction should probably be given to the students prior to the test being taken.

The purpose of using this psychometric instrument was to assess the external convergent validity of VItrap. One could hypothesize that there should be a correlation between individuals' self-assessments of their imaging tendencies using the VVIQ and whether they were deemed to be Imagers or not as a result of using the VItrap. Implicit in this argument is an understanding that a low score on the VVIQ indicates that an individual is a good Imager (i.e. the individual is good at assessing his/her imaging ability).

The question may be raised as to why there was in fact no association between these two measures (i.e. Imagery as estimated by VVIQ and Imager as assessed by the Vitrap when it might have been expected that there should have been such an association. It is plausible to suggest that the VVIQ is measuring an entirely different Imagery construct to the Vitrap. An additional contributory factor may be the low test-retest reliabilities of this questionnaire. One might argue that an individual's ability to imagine may be relatively stable and that the test-retest reliability should have been considerably higher in this case. Also the estimated reliability of the VItrap is low at about 0.33 and this may also have affected the lack of correlation between the VVIQ and VItrap.

It is always necessary to attempt to assess the external convergent validity of a new test such as VItrap. Since the VVIQ was an easily available measure with a good reported reliability in the first instance, it was useful to perform this exercise. The VVIQ assesses visual imagery but does not assess other forms of imagery such as spatial imagery. The instructions at the beginning of the questionnaire are deficient in that they are

prejudicial against a positive response.

This study highlights the difficulty of finding a measure comparable with the VItrap in order to assess its external convergent validity. A more detailed account of this study may be found in Parkinson, Mullally, Moore & Redmond (2005) and Parkinson, Mullally & Redmond (2004b)).

7.7 Comment on Reliability and Validity

Ideally for any psychometric test, the product of the Reliability (estimated from 0 to 1) times the Validity (estimated also from 0 to 1) is what needs to be maximised (Parkinson, Mullally & Redmond (2004c)).

In general, difficult as it may be, it is possible to come up with a number from 0 to 1 for Reliability. (There are several pertinent estimates of Reliability, in particular Internal Reliability and Test Retest Reliability. Test Retest Reliability is the more applicable of the two because the construct of cognitive style is understood in this instance to be stable.

However, Validity is much more difficult. It is generally difficult to even get agreement as to which type of Validity measure should be used. It is important to appreciate that a test can be reliable and not valid but never valid and not reliable.

Ideally it is desired that the product of reliability times validity (both estimated in the range from zero to one) should be maximized i.e.
max r*v
where r (reliability) and v (validity) have a range of 0 to 1.

7.8 Overall Chapter Conclusions

- VItrap is independent of two measures of intelligence: Raven's APM2 and the Mill Hill Inventory

- VItrap is independent of personality as assessed by Eysenk's EPI and Spielberger's STAI

- The VItrap did not correlate with the VVIQ (a benchmark measure of convergent (concurrent) validity

- The VItrap did not correlate with Sadler-Smith's Instructional

Preference Questionnaire

In this Chapter VItrap has been shown to be independent of Intelligence and Personality. VItrap has not been shown to predict observable behaviour as examined by correlating it with Sadler-Smith's Instructional Preference Questionnaire. While Riding & Sadler-Smith reported significant interactions for the CSA with this measure, Peterson's (2003) results support the VItrap findings in relation to Sadler-Smith's Instructional Questionnaire. Both Riding and Peterson (2003) did not empirically demonstrate the external convergent validity of either of their V-I style measures.

To conclude, the test-retest reliability value of the VItrap needs to be significantly increased before the above conclusions can definitely be held to hold and be valid.

In Chapter 8, the neurophysiological basis of the test will be examined similarly to what had been done by Riding, Glass, Butler and Pleydell-Pearce (1997) for Riding's CSA. This will be approached using a method a little different to that used in Riding's study. The Neural Navigator is used to record participants' EEGs while doing the VItrap in a slightly modified form (to that used in the other studies in this thesis).

In Chapter 9, computer applications are designed which integrated VItrap as part of the interface. The VItrap was used in these cases to determine the appropriate Human Computer Interface for an individual. The goal in these cases was to explore whether the VItrap matched interface would give optimal performance. Individuals were also assesses against a mismatched interfaces. This approach also examines the construct validity of the test.

Chapter 8:
A Neurophysiological Approach to Validating VItrap

8.0 Introduction

In the previous chapters the validity of VItrap was investigated. This Chapter examines the neurophysiological basis of this test. Studies by Riding, Glass, Butler & Pleydell-Pearce (1997) and Glass & Riding (1999) concluded that cognitive style had a neurophysiological basis using electroencephalography (EEG). Riding (2000) has commented on the importance of establising a neurophysiological basis for cognitive style.

This Chapter presents an overview of how the EEG is generated, a review of the literature on hemisphericity, and finally one study - Study 8.1. Study 8.1 (n = 19) examines whether it was possible to replicate the findings of Riding, Glass, Butler and Pleydell-Pearce (1997), which had broadly established that cognitive styles had "an observable physiological basis".

8.1 Brief overview of the EEG.

In 1929, Hans Berger (Berger (1932, 2006)) was reported to have been the first to record brain electrical activity in humans. This electrical activity that he recorded, from the cerebral cortex, was called an electroencephalogram (EEG).

The EEG signal is generated by special neurons known as pyramidal cells in an area on the neuron known as the post synaptic dendrites. Neurons communicate by means of action potentials (electrical impulses), which move down the axon resulting in the release of a neurotransmitter. The function of the neurotransmitter is either to increase or decrease the permeability of the cell membrane. These signals are only a few microvolts and they need to be amplified by a factor of a million so that they can be viewed on a computer screen.

Neurons communicate with each other by means of electrical impulses known as action potentials. When a neuron receives a stimulus of ample strength (which can be the sum of multiple signals) and which exceeds that neuron's resting threshold an electrical current moves down along the length of its axon.

When a neuron is at rest the inside of the axon is charged negatively with respect to the outside. In other words the chemical and electrical gradients are at equilibrium. The neuron has a resting membrane potential of approximately -70 microvolts on the inside.

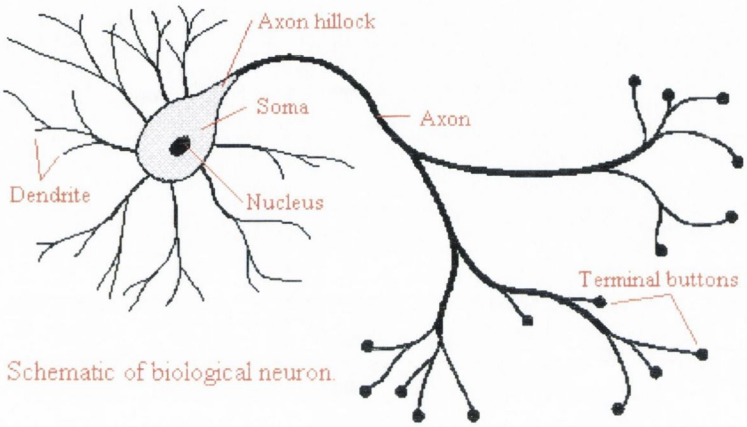


Fig 8.1 Schematic Diagram of a Biological Neuron

A neuron fires an action potential when it reaches a certain electrical threshold. This can be best explained by the domino effect. The incoming stimuli are summed and unless they reach a certain potential nothing occurs. When the threshold for the neuron has been reached, the neuron fires the electrical impulse or action potential.

When the neuron fires, the action potential is transmitted along the axon of the neuron. The neuron goes through a series of ion channel opening and closings along various sections of the axon. In somes cases axons are coated with a myelin insulation sheet which helps with the conduction of this electrical impulse (action potential).

When the neuron receives a stimulus it is transmitted along the axon. The transmission of this impulse causes a reversal of charge to occur at specific points along the axon as the impulse moves along that particular section of the axon. When the impulse has progressed onto the next section of the axon, the charge is restored and the impulse moves down along through the rest of the axon in this manner. Finally, this electrical impulse reaches the end of the axon and arrives at the synapse.

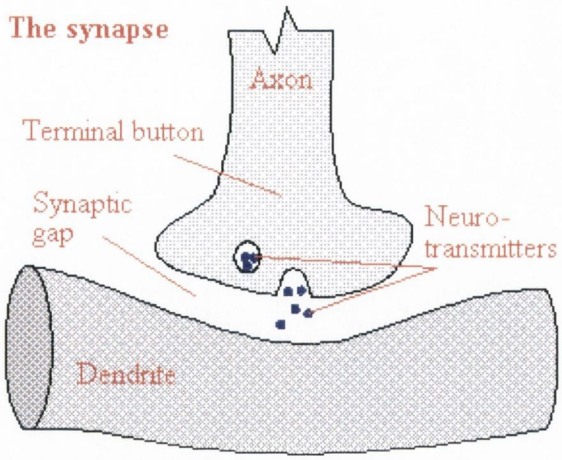


Fig. 8.2 The Synaptic Gap

The electrical impulse is unable to bridge the gap between one neuron and another, hence the role of a neurotransmitter chemical. The synapse is a small gap, which exists between one neuron and another and is normally found between dendrites of one neuron and the axon terminals of another. The electrical impulses are unable to cross these synapses and they stimulate vesicles in the synaptic knob to release a neurotransmitter chemical.

This chemical diffuses across the synaptic gap that is the space between the neuron that is sending the impulse (the presynaptic) and the neuron that is receiving the impulse (the post-synaptic). This impulse (action potential) is regenerated when the neurotransmitter recombines with special receptors on the post-synaptic neuron.

The neurotransmitter can inhibit or excite the post-synaptic neuron. In other words it either stimulates the synaptic knob into firing a pulse or inhibits it from firing. Neuro transmitters, which are excitatory produce excitatory postsynaptic potentials (ESPSs) and similarly those which are inhibitory, produce inhibitory post-synaptic potentials (ISPSs).

The combination of these two types of post-synaptic potentials (ESPS and IPSPS) causes currents to flow within and about each of the neurons creating a potential field, which is recorded on the surface of the cortex. The EEG records the electrical activity, which is generated by pyramidal cortical neurons (Sterman (2002)).

The EEG when recorded during sleep or daily activity is generally rhythmical, the pyramidal neurons being influenced by oscillating input in the thalamocortical circuits. Rowan and Tolunsky (2003) state that the underlying mechanisms that explain this rhythmical oscillatory process are not fully understood. They briefly refer to two main processes understood to be responsible: a) the interaction between the thalamus (acting as a pacemaker) and the cortex. The second explanation is based on the functional properties of large neuronal networks in the cortex that have an intrinsic capacity for rhythmicity.

Sterman (2002) provide a greater in-depth analysis that explains the generation and rhythmicity of the EEG.

8.1.1 Hemisphericity of the Brain

Much research has concentrated upon the hemispheric specialization of language and spatial ability. The quest has been to investigate what areas in the brain are responsible for these two capabilities. The phenomenon of hemispheric specialization has also been referred to as functional hemispheric asymmetry, lateralization and localization. The study of split-brain patients and those suffering from lesions in both hemispheres has provided many opportunities for researchers to study brain hemisphericity.

This review of the literature has been influenced by the following reviews (Martin (1998)), (Hellige (2001)) (Gazzaniga (2002)), (Banich (2004)). Traditionally the left hemisphere is associated with language, and with rational and analytic thought. The right hemisphere is allied to visuo-spatial, holistic, intuitive and emotional functionality. This view presupposes that these two hemispheres are entities and function independently of each other. Sperry (1964) in his research involving split-brain patients popularized this view. His work focused on patients that had the corpus callosum severed, resulting in the partial separation of the left and right hemispheres.

Galin and Ornstein (1972) report that "... different cognitive functions depend ... on either left or the right hemisphere". They refer to literature, which supported the assertion that left hemispheric brain damage affected language processing and analytic tasks whereas damage to the right hemisphere disrupted

"spatial relations and gestalt, synthetic tasks". They report, that "ratios of average power (1 - 35 Hz) in the contralateral (homologous) leads T4/T5 and P4/P5 were computed. This ratio (right over left) was greater in verbal tasks than in spatial tasks".

Central to this issue is whether one can reasonably assume the right brain as being creative and the left brain as being strictly logical. Is this view a little too naïve? Hellige (2001) argues, that "this view is far too simplistic" and that the left brain and right brain are components of an extensive cortical and sub-cortical cognitive layer.

8.1.2 Hemisphericity and the lower animal kingdom

Rogers (1980) reports a preference for the left foot amongst parrots. Collins (1985) reports on paw preference in mice. Bradshaw (1993) reviews motor asymmetries in the lower animal kingdom and attempts to relate this to human handedness. Hellige (1990) in a review of hemispheric asymmetry in non-humans concludes, that "the ubiquity of behavioral and biological asymmetries indicates that brain asymmetry is neither unique to humans nor completely dependent on the development of language".

Hatta and Koike (1991) report that the left hand of monkeys seems to be preferred for reaching and the right for manipulation of objects. Indeed this finding is somewhat analogous to what right-handed humans do when writing. The left hand is used perhaps to steady the paper or book but the finer psychomotor skills of writing are the domain of the right hand. Kinsbourne (1995) provides some evidence of hemispheric asymmetry in the lower animal kingdom. He reports that song birds have left hemisphere control of song.

Gazzaniga, Ivry and Mangun (2002) state that "the human species exhibits differences in the function of the two hemispheres" and "hemispheric specialization is not surely a human characteristic".

8.1.3 Language hemispheric specialisation

Language ability involves many facets, for example, comprehension, semantic and phonetic analysis, humour, metaphors, and finally speech production. One's spatial ability

involves functions like map reading, mental rotation and pattern matching. Can it be assumed that only the left hemisphere and right hemisphere are specifically hemispherically asymmetrical for language and spatial ability?

Bottini et al. (1994) investigated the role of the right hemisphere in the interpretation of figurative aspects of language, using positron emission tomography activation as a technique and stated "we conclude that the interpretation of language involves widespread distributed systems bilaterally with the right hemisphere having a special role in the appreciation of metaphors".

Binder et al. (1997) reported increased left activation for phonetic and semantic analysis of words. Hellige (2001) in his review of the literature on behavioural asymmetry for language states that "the left hemisphere seems dominant for the production of overt speech ... the perception of phonetic information ... aspects of semantic analysis". Interestingly he reports that "the right hemisphere seems dominant for certain other aspects of language ... use of pragmatic language and intonation and prosody to communicate emotional tone".

Martin (1998) presents a brief review of research in this area. He reports that Peterson (1988) found increased activation in the left posterior cortex during passive listening of words. He also reported increased left activation for phonetic and semantic analysis of words. He reported increased activity when listening to words over Broca's and Wernicke's area. This was confirmed by Zatorre et al. (1992).

Martin (1998) also states that "However although there is left dominance for language and speech the right hemisphere does play some part in language processing." Hellige (2001) states "there is no doubt that the left hemisphere is typically dominant for a number of important aspects of language, overt speech, phonetic decoding, syntactic and semantic processing." However he also concludes his review of the language and behavioural asymmetries in humans by stating: "Consequently we must be very cautious in identifying one hemisphere as dominant for something as complex and multi-facetted as language. The

reality is likely to be that the hemispheres play complementary roles for language."

The literature reviewed above has shown a movement away from left hemisphere processing for language. It is now accepted that some language processing occurs in the right hemisphere.

8.1.4 Visuo-spatial processing, visual imagery and hemisphericity

For some time, it had been assumed that the right hemisphere processed images. Morgan et al. (1971) report greater activation over the right hemisphere during a spatial task amongst subjects. However Givens et al. (1979) reported, based on two different studies, which they undertook, that "no evidence for lateralization of different cognitive functions was found in the EEG". A number of studies provide evidence to suggest that visuospatial ability is characteristic of the right hemisphere (De Renzi (1982)).

A review of the literature presented by Ehrlichman and Barrett (1983) reported little confirmation of right hemispheric dominance for Visual Imagery. In their article they questioned the assumption that the right hemisphere plays a very unique role in Mental Imagery. Their review of the literature considered brain damaged and normal subjects. This review of the literature effectively refuted their initial hypothesis that there was right hemispheric specialization for Mental Imagery. The categorization included the following: "Failures (studies which failed to find a result if the hypothesis were true); Contradictions (studies which present results which were apparently incompatible with the hypothesis); Indeterminate ("studies which present results that support the hypothesis but could be interpreted as supporting a bilateral representation of Imagery").

Ehrlichman and Barrett (1983) concluded that none of the studies that they reviewed "can be described as unequivocally supporting the hypothesis of right hemisphere specialization for Mental Imagery and some appear inconsistent with such a formulation". They also contend that "because a large number of reports in fall into the indeterminate category it might be argued that well designed experiments specifically intended to study differential hemispheric involvement in Mental Imagery have

been extremely infrequent, and thus no conclusions should be drawn from the existing literature". These authors also argue that because an "assumption that Imagery is a privileged function of the right hemisphere" that research has been restricted "on the differential use of the hemispheres". They state that wider research could turn out to be quite advantageous in areas like spatial ability and metaphor for example.

Farah (1984) poses the question "Is Visual Imagery a lateralized function of the brain"? Also she wondered which of the two hemispheres was specialized for imaginal thought. This study concluded, that "only the left hemisphere performed on the Imagery task". Another study by Farah (1986) suggests that the left hemisphere is critical for the generation of mental images.

Sergent (1990), cited in Hellige's (2003) review, concludes that there is little evidence to support Farah's claims and suggests that instead "that both hemispheres contribute to image generation". Hellige, in his review, also states that "whatever is the eventual resolution of the issue regarding hemispheric asymmetry for image generation, it seems very unlikely that either hemisphere completely lacks the ability to generate visual images".

Visual Imagery can be thought of as being the generation of a mental image without a physical stimulus. Kosslyn (1987) states that "Visual Imagery is usually identified as producing the conscious experience of seeing with the mind's eye rather than with real ones". He further develops this definition of Visual Imagery by explaining it in terms of "the experience of seeing in the absence of the appropriate sensory input".

Kosslyn (1988) also argues, based on studies with split brain and normal subjects, that two different types of processes are used when images were being formed. He states "some of the processes used to arrange parts are more effective in the left cerebral hemisphere and some are more effective in the right cerebral hemisphere; the notion that mental images are the product of right hemisphere activity is an over simplification". Kosslyn et al. (1995a) have reported that imagery is not a single mental process but "involves a host of processes working together".

They state that "images can be regarded as a particular way in which information is stored and these representations are processed in specific ways". They also add, that "the system in which images are processed can be conceptualised in terms of four major abilities". These are (a) Image inspection (b) Image generation (c) Image transformation and (d) Image retention.

Kosslyn et al. (1995b) argue in another paper that there are two different processes that can be used to arrange them.
The first process uses stored descriptions to arrange parts, more efficiently performed by the left hemisphere". The second process is characteristic of the right hemisphere, which uses stored memories of precise distances and positions.

Martin (1998) in the review of recent literature in this area concludes that "most of the invasive and non-invasive studies ... ascribe a greater role for the right hemisphere in visuo-spatial processing". Critically he states, at the end of the review, "that lateralization is not absolute: one hemisphere is often able to subsume the function of the hemisphere dominant for that function even if this ability is at a rudimentary level".

Springer and Deutch (2003), referring to recent research on Visuospatial processing and lateralization, suggest that "contemporary research suggests that most hemispheric differences are one of degree that is both differ in the relative efficiency of individual processes rather than overall capabilities".

Hellige (2001) in an extensive review of the area states "that there is a long way to go before we have a complete cognitive neuroscience of Visual Imagery ".

Gazzinaga, Ivry and Mangum (2002) describe the current perspectives on hemispheric lateralization as "modern conceptualization of hemispheric specialization; namely the two hemispheres may work in concert to perform a task even though their contributions may vary". In other words each hemisphere contributes to the task in a specialized manner. The authors state interestingly that "language is not the exclusive domain of the left hemisphere and the right hemisphere also might contribute", and "the left hemisphere does not defer to the right hemisphere

on visuospatial tasks but processes this information in a different way". These authors, based on these perspectives on hemispheric specialization, conclude, "that much of what we learn from clinical tests of hemispheric specialization tells us more about our tasks rather than the computations performed by each hemisphere". Excluding speech production, "they conclude that each hemisphere has competence in every cognitive domain".

These authors also make a very significant comment when they state "recent work such as the frequency hypothesis has moved lateriaity research towards a more computational account of hemispheric specialization". They also add, "that theoretical advances are taking the field away from the popular interpretations of cognitive style and providing a scientific basis for these robust processes".

8.2 The Work of Riding and his Co-workers using EEG

The background to this study of the neurophysiological basis of the VItrap is a review of Riding and his co-researchers' papers on their work and cognitive style. Riding and his co-workers believe that hemisphericity of the brain and alpha suppression are important underlying factors in establishing the neurophysiological basis for his CSA. Riding and his co-researchers published three papers in the 1990's, which discussed cognitive style and the EEG.

Riding, Glass and Douglas (1993) discussed the following idea and questions:
"It is predicted that subjects who are Verbalisers would show marked EEG alpha asymmetry (resting and active) with suppression over the left hemisphere. Those in the Imager group should correspondingly show the most asymmetries over the right hemisphere, with right hemisphere EEG alpha suppression";

"Is performance of a Verbaliser correlated with greater 'active' or 'resting' alpha asymmetry? That is, is alpha rhythm suppression greater over the left hemisphere (for example, the posterior temporal [T5] region) than the right hemisphere? The same question can be asked of the correlation of Imagers with asymmetrical right hemisphere alpha suppression";

"A related question concerning right hemisphere function has been posed, and answered, using alpha asymmetries by Fürst (1976). He found that those subjects whose speed of thinking was greatest in solving three-dimensional puzzles tended to show the greatest relative asymmetrical EEG alpha suppression over the right hemisphere".

Riding, Glass & Douglas (1993) concluded that the proposed strategy would have the following benefits. It would:

(1) have implications for the refinement of models of cerebral localisation and lateralisation;
(2) validate cognitive styles in terms of cerebral function;
(3) clarify work on hemispheric specialisation by removing the confounding effects of individual differences in cognitive style.

Riding, Glass, Butler and Pleydell-Pearce (1997) further developed this model relating cognitive style to task hemispheric asymmetry and alpha suppression. Simply put: "the suppression of the alpha rhythm is an indication of local cortical activation". They refer to the work of Williamson (1997) who also provides support for this assertion. They apply this argument to the examination of lateral electrodes on both left and right hemispheres. In other words for the Verbaliser there would be greater left hemispheric suppression at T5 as opposed to T6 , F7 and F8.

In this paper they also refer to the early work of Bogen (1969) who supported the theory of hemispheric specialization. Riding, Glass, Butler and Pleydell-Pearce (1997) also mention that "hemispheric specialization has long being associated with the left hemispheric being the location of verbal function" while also acknowledging "that the evidence has not always been that clear" (Bisiach and Berti (1990), Langhinrichsen & Tucker (1990) and Riding, Glass, Butler and Pleydell-Pearce (1997)).

Riding, Glass, Butler and Pleydell-Pearce (1997) also refer to the work of Cohen (1982) who proposed, "that individual differences in left/right hemispheric specialization may be due to cognitive style differences". However Cohen (1982) as quoted by Riding, Glass, Butler and Pleydell-Pearce (1997) notes that it was

necessary to provide independent evidence of this phenomenon. This was the underlying objective of Riding and his co-researchers in the study discussed below in their (1997) paper and further reported upon in their (1999) publication.
In their first experimental work paper Riding, Glass, Butler and Pleydell-Pearce (1997) described individual differences (cognitive style) with respect to the EEG. These differences in cognitive style were interpreted in the context of one frequency range: the alpha band. This study is discussed in more detail later.

In Glass and Riding (1999), the population used was exactly the same as that used in the (1997) study. This particular paper reports on further analysis based on the data previously collected.

The authors' objective was to examine "whether the cognitive style effects found in the alpha band range are also present in some of the other frequency bands".

The authors reported that in the "paramedial cluster Verbalisers have greater right power than Imagers for all bands except alpha" and that "overall power was greater on the right for Imagers than Verbalisers frontally and the converse occipitally". In the paramedial cluster of electrodes, there was a significant interaction of Verbal-Imagery style with hemisphere and frequency ($f = 3.90$, df 2.05, $p = 0.034$, epsilon = 0.41). The authors report that "Verbalisers exhibited greater left than right preponderance of delta, beta2 and gamma and this asymmetry was significant for Imagers". There was also a "significant interaction between region hemisphere and the Verbal-Imagery style". Based on the results of this study presented above the authors state, that "this must be regarded as a hemispheric difference between Imagers and Verbalisers" but they also mention "that it cannot be readily explained ... because the neurological significance of power over all frequencies is not understood" (Glass & Riding (1999)).

This brief review of the literature raises some interesting issues. It would appear that where Imagery and Language is processed is far from definite, as discussed above. There is now evidence

that suggests that Imagery processing is not solely characteristic of the right hemisphere.

8.2.1 Basis for the Riding, Glass, Butler and Pleydell-Pearce (1997) study

Riding, Glass and Douglas (1993) present a review of the literature that served to provide a basis for their 1997 study (Riding, Glass, Butler and Pleydell-Pearce (1997)). This basis is briefly summarised below.

As has been mentioned above in this Chapter, hemispheric specialisation ascribes certain attributes to the left and right hemispheres. Chiefly, verbal functioning is assigned to the left hemisphere and imagery function to the right hemisphere.

The review already presented in this Chapter concludes that hemispheric specialisation is far from being clear cut. However, Riding, Glass & Douglas (1993) state, that "where studies have paid no difference (sic) to individual differences in cognitive style it is not surprising that results regarding lateralisation and localisation have often been unclear". Riding, Glass, Butler and Pleydell-Pearce (1997) argue, that "the methods of task-related EEG alpha (or beta) asymmetries can be used to investigate the relationship between cerebral location and cognitive style dimensions."

The basic assumption, that is fundamental to the interpretation of task related asymmetries, "is that when the alpha rhythm is suppressed the cerebral cortex underlying the EEG scalp recording electrode is more active". On the other hand, when the cerebral cortex at that point is at rest there is more alpha rhythm present. Essentially what this means is that "suppression of the alpha rhythm is an indication of local cortical activation"; on the other hand "when the alpha rhythm is more abundant, the subject is mentally at rest and the underlying cortex is assumed to be relatively inactive" (Riding, Glass & Douglas (1993), Riding, Glass, Butler and Pleydell-Pearce (1997)).

The theoretical basis for this assumption is explained by Anderson & Anderson (1968) and Riding, Glass & Douglas (1993). Riding, Glass and Douglas explain this in the following manner: "a network of synchronously active cortical neurons

signalling electrical responses are carrying information which is redundant when their electrical activity is synchronous. They also state that "these neurons are processing less information than the same network of neurons acting independently of the other neurons in the network; their electrical activity is then desynchronised". When there is desynchronised alpha activity there is high frequency and low amplitude while the reverse true for synchronous alpha activity when the subject is mentally at rest.

They use the example of the stock exchange to metaphorically explain this: when various members of the stock exchange are involved in different transactions there is a low hum (low amplitude, high frequency, equivalent to desychronised activity) in contrast to a scenario where all members of the stock exchange are singing together (no work being done, high amplitude, low frequency, synchronous activity) (Riding, Glass & Douglas (1993)).

Riding, Glass and Douglas (1993) and Riding, Glass, Butler and Pleydell-Pearce (1997) refer to empirical research that has supported the view that the alpha rhythm has been suppressed when individuals have been engaged cognitively relative to when they are in a restful state (Berger (1932), Glass (1964), Osborne and Gale (1976)).

Riding, Glass, Butler and Pleydell-Pearce (1997) argue that, based on the definition of cognitive style, "individuals will, where possible, represent and process information in their habitual modes", and also while "verbal information will result in predominantly left hemispheric activity in the Verbaliser, it may also produce right hemispheric activity on Imagers where the material allows for this". By implication the authors argue "Verbalisers ... will translate pictorial information into words or semantic representations, imagers will represent information into mental pictures whenever possible". Consequently it would be expected that for Verbaliser greater suppression would occur over T5 (left posterior cortex - Wernicke's area) than at T6, and the reverse for Imagers". The aim of their 1997 study was to provide comparison of EEG features in different regions with an individual's location on the Wholist -Analytic and Verbal - Imagery dimensions.

8.2.2 Method of Riding and Glass

Riding, Glass, Butler and Pleydell-Pearce (1997) provided each individual with 40 computer-presented word targeting trials. These were tasks differing in the amount of words presented to the individual per second. The individual was required to press a button "whenever a noun from the superordinate categories fruit or vegetable was displayed (e.g. apple or carrot). Accuracy was not monitored in this experiment. Each participant's EEG was recorded while this task was being done. When these word trials had been completed, all participants completed the CSA, categorised according to cognitive style.

8.2.3 Results of Riding and Glass

Riding, Glass, Butler and Pleydell-Pearce (1997) report that at the locations Fz, Cz and Pz (midline), there was a significant interaction between the Verbal- Imagery style, Location and Task in their effect on alpha power ($F = 2.05$, $df = 8.88$, $p = 0.049$). There were no significant interactions between the left posterior or antero-posterior locations with the Verbaliser - Imager dimension. However there were significant interactions between the Verbaliser - Imager dimension, Lateralisation and Task in their effect on the output for the following pairs -

Pairs	F value	Degrees of freedom	Significance level
T3 & T4	3.43	4.44	0.016
T5 & T6	2.69	4.44	0.043
T3 & T5	2.59	4.44	0.049
T4 & T6	NS	NS	NS

Their sample consisted of 15 right-handed volunteer adults, six females and nine males. The subjects completed a word categorization task while their EEG was being recorded, and then they did the CSA. The sample was then divided into Verbaliser and Imagers, Wholists and Analytics based on the scores obtained using the CSA. Eye movement artifact was controlled for but there was no suggestion that time of day recording was adjusted for.

They reported for the Verbaliser - Imager style dimension, as measured by the CSA, that "there was a style - hemisphere effect with Verbalisers having relatively more alpha suppression on the

left posterior temporal location T5 compared to right T6 and Imagers having the reverse". They report for the midline cluster at location Fz, Cz and Pz that there was a "significant interaction between the Verbal - Imager style, location, and task in their effect on alpha power output". They also reported significant interactions between the Verbal-Imager dimension, lateralization and task for these pairs of electrodes, T3/T4, T5/T6 and T4/ T6.

These findings supported the hypothesis as presented by Riding, Glass & Douglas (1993).

8.2.4 Conclusions of Riding, Glass, Butler and Pleydell-Pearce (1997)

Riding, Glass, Butler and Pleydell-Pearce (1997) stated, that "the intention of the pilot investigation was to establish that cognitive style could be related to localised/lateralised brain activity as indicated by EEG output they conclude that "this relationship has been confirmed and that further study would be justified. They also state, that "the Verbal - Imagery style shows a hint of a lateralisation effect as was hypothesised by Riding, Glass & Douglas (1993).

8.3 Study 8.1 Study of the Neurophysiological Basis for the VItrap

8.3.1 Introduction to Study 8.1

In a similar fashion to the Riding, Glass, Butler and Pleydell-Pearce (1997) reviewed above, the new test of the Verbaliser - Imager dimension (VItrap) was investigated as to whether it has an observable physiological basis. Riding has commented previously on the importance of establishing the validity of a construct in this manner (Riding (2000) in Riding and Rayner (2000)).

The aim of the study is to investigate whether the Verbaliser-Imager dimension interacts with task, region and hemisphere based on a) the theory of alpha suppression and b) hemispheric lateralisation. Based on Riding's et al.'s results (1997, 1999) and with a better test of the V-I dimension VItrap, it might be expected that these results could be broadly replicated.

8.3.2 Method



The EEG was recorded using the Neural Navigator (Sterman 2002). The 10/20 International configuration was used. The data were analyzed using Skil neurological software. Prior to the files being processed the Skil software was used to remove artifacts.

8.3.3 Filter Settings used in the Neuronavigator:

Each participant was recorded while engaged in three specific tasks. These were: Eyes Closed, Task 1 and Task 2. Where the recordings failed for an individual task, the participants were requested to repeat the exercise again on another occasion.

8.3.4 Description of method

The author spent two weeks training on EEG acquisition and use of analysis software under the guidance of Professors Barry Sterman and Denise E. Maikowicz at the Institute for Achievement of Human Potential, Wyndmor, Pennsylvania, USA in August 2002. Later, in 2003, the author also spent a week at a SABA 2 Workshop (Saba = Society for Advanced Brain Analysis) on the island of Santa Catalina, California, USA trying out the new version of the Neuronavigator equipment and using Skil (Sterman Kaiser Imaging Laboratory) EEG analysis software.

8.3.5 Preparation of Participants for EEG recording

Before each of the participants' EEGs was recorded they were asked a few personal details. These details included questions as

to gender, reported handedness, date of birth, and whether they had been in accidents previously and suffered head injury.

The participants also answered questions relating to the taking of medication. These particular questions were used to screen the sample used in the study. Medication, head injury and other factors such as ADHD can affect the type of EEG recorded so therefore the sample had to be controlled for these variables as far as was possible.

Possibly the most important requirement is that the participant be relaxed and comfortable prior to the EEG being recorded. This was achieved by explaining exactly what each of the various stages entailed prior to the stage being initiated. It was important to answer any questions that the individual had, and to assure them of confidentially. Their records were coded so as to guarantee anonymity. Prior to the recording being taken, for the purpose of reassurance, the subject was asked to examine some pieces of the apparatus used, for example: blunt needle and syringe, and blunted dowel in particular. This was imperative as it had the desired effect of making them less stressed.

Finally before the recordings were taken the participant had the opportunity of monitoring the effect of creating their own artifact on an unrecorded EEG segments. This served two objectives: it illustrated to the subject how an artifact could be caused by a muscle movement or eye blink. It also created a greater sense of participant involvement in the process.

The recordings were taken under the following conditions: the room was darkened, the subject was seated comfortably on a high backed chair; there was little or no noise present. During each of the recordings the EEG was monitored by the Author using the computer screen. At the same time a close watch was kept on the subject for signs of nervousness or other factors which might underlie artifact production for instance due to teeth grinding. It was felt that this might have been useful when the files were being processed for artifact removal as such information could provide clues as to how the artifact was generated initially. At the beginning of each task the individual was reassured and any questions were answered. This created a good subject /clinician rapport and also facilitated better

recording conditions. The subjects were advised to wash their hair up to 24 hours previously but not to apply hair conditioner (because it increases conductivity of the hair and reduces the voltage recorded).

8.3.6 Fitting the Electrode Cap

When fitting the electrode cap on the subject two measurements were taken. The circumference of the head was measured using a measuring tape and the appropriate size of electrode hat was then selected. The next stage was positioning the electrode hat correctly on the head. The distance was measured from the naison to the inion. Ten percent of this distance was then calculated. The forehead was marked with an India pencil at a point 10 per cent from the bridge of the nose (nasion). This acted as a guide to placing the cap symmetrically on the head.

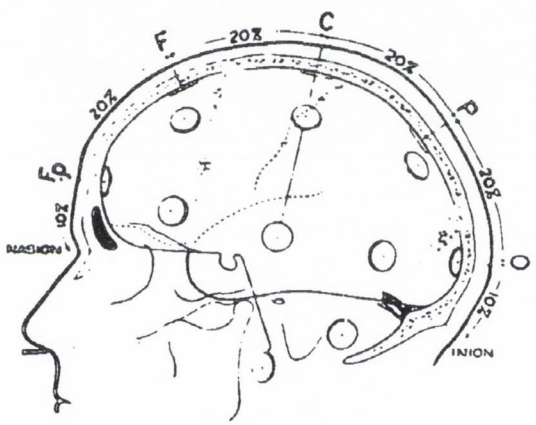


Fig. 8.1. Side elevation of electrode placement positions; note inion and nasion (adapted from Rowan & Tolunsky (2003))

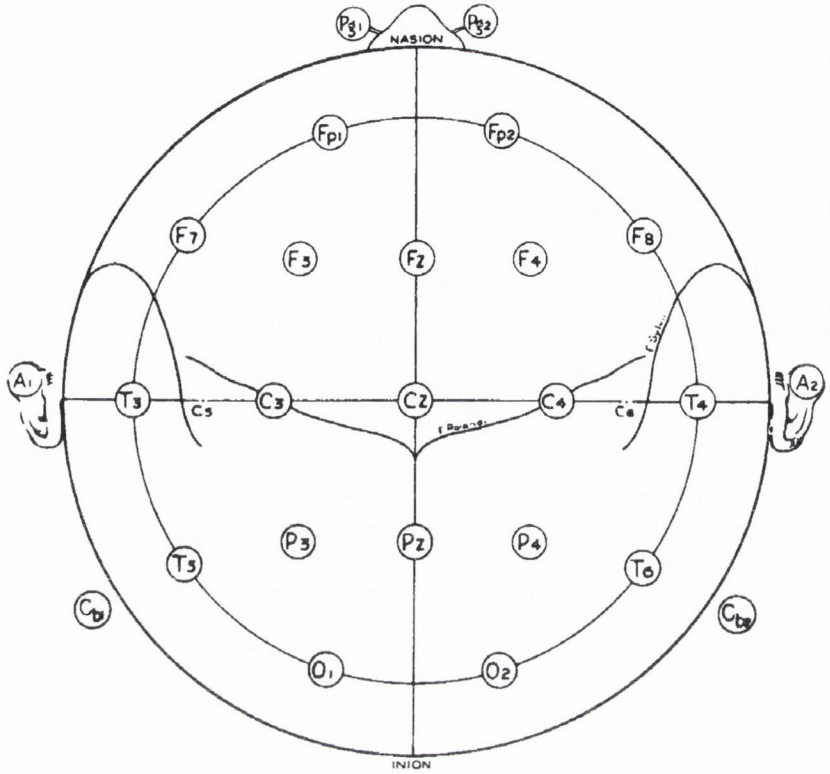


Fig. 8.2 Plan view of electrode placement positions (adapted from Rowan & Tolunsky (2003))

8.3.7 Impedances

High impedance results if there is a not a good enough contact between the scalp electrode interface. Typical electrode impedances should be less than five thousand ohms.

8.3.8 Earlobes

The earlobes were cleaned thoroughly with an ethanol-saturated Qtip. This involved applying thirty to forty strokes. When this was completed Nuprep was applied to the earlobes in much the same manner. Nuprep is an abrasive gel used to remove cells on the skin's outer layer.

At this point the ear clip electrodes were filled with a Ten20 conductive paste and were placed on the earlobes. A small piece of cotton wool was placed at the back of the ear clip to prevent the metal electrode from making contact with the skin and arteries of the neck. (Should this happen the EEG would record a pulse artifact).

The electrode cap conformed to the 10-20 electrode designation. The electrode cap was then placed symmetrically on the head. Sponges were placed under cap at positions FP1 and FP2. The electrode cap was checked to ensure that it was a good fit.

Nuprep was then applied using a blunted dowel through each of the twenty electrode openings. This was repeated until contact was achieved with the scalp in each case.

At this stage electrode gel was applied via a syringe with a blunt needle through each of the twenty electrode openings. The syringe was moved in a clockwise direction until contact was again made with the scalp.

When this process was occurring the impedances for each of the twenty electrode points were being monitored using the NeuroNavigator software. When impedances fell below 5 K ohms for all the electrodes the recording was started. This was depicted in the NeuroNavigator software by the impedance bars changing from red to yellow (less than 10 kOhms) and finally to green (less than 5 kOhms).

When the EEG had been successfully recorded and acquired using the NeuroNavigator acquisition software it was analyzed using the SKIL software. This package automatically performed artifact rejection whose accuracy was checked manually by Sterman and Malkowicz (2002). At this point the EEG was checked for closed head injury and ADHD. If there were any obvious pathologies evident in the topographical plots, these records were removed from the sample (the original sample size of 26 dwindled to 19 and then 12 depending on the study). It was necessary to discount one set of recordings due to closed head injury. (See Appendix 16-8.1.A for sample figures, topographic maps and spectral plots).

8.3.9 Sample

The sample initially consisted of twenty-six participants (male = 12, female = 14), mixed in age range. Some of the samples were mature university students who volunteered (and subsequently were given a twenty five euro book token). The other participants in the study varied in background, some being University lecturers and others from the business sector. This

sample was screened with respect to self-report for head injury, medication and right-handedness. The sample size that was then suitable for use in this analysis was n = 19. Three recordings were taken for each individual.

The basis for selecting right-handed individuals was based on some of the studies reviewed by Hellige (1990), which indicated that "left- handers are more variable than right-handers in both degree and direction of hemispheric asymmetry".

Studies which Hellige (1990) reviews in respect of gender differences and asymmetry lateralisation indicates that "evidence ... seems restricted to verbal tasks, and even here the differences may have to do with sex related differences in preferred strategy than with hard wired differences in hemispheric asymmetry". Hellige (1990) also suggests that there is not enough evidence to support the view that in females there is less hemispheric specialisation than in males. Hellige (2001), in a further review, mentions that there are gender differences in the size of the corpus callosum, but states "it remains to be determined how the size of different regions of the corpus callosum relates to functional asymmetry but there is some indication that the nature of the relationship may be different for males and females". It is possible that studies using functional magnetic resonance imaging may clarify this issue in time.

8.3.10 Task One: Eyes Closed - Ec

The subjects sat with their eyes closed for six minutes to eight minutes. During this task they were asked to relax. It is difficult to get subjects to relax entirely and it can be quite impossible to determine what the effect of thinking has on this baseline recording if they do not.

When the recording was completed the files were saved and imported into the Skil analysis software for further analysis. The EEG data for each recording were corrected automatically by the package for time of day.

8.3.11 Task Two

The subjects sat in front of a laptop computer. Thirty questions (word form "Type") from VItrap were presented. These were answered by the subject, using a mouse as the input device,

which required minimal movement by the subject thus reducing muscle artifact. The test questions were presented centrally on the screen so as to keep eye movement to a minimum and reduce left or right field effect.

8.3.12 Task Three

The subject sat in front of a laptop computer. Thirty questions (word form "Shape") from VItrap were presented. These were answered by the subject, using the mouse. The test questions were presented centrally on the screen.

8.3.13 Statistical Analysis of Data

The EEG records were automatically screened for artifacts during the analysis with the Skil software. When each record had been checked for artifact, it was then further processed by the Skil package to remove any artifacts. This package then calculated the mean spectral power values for each of the electrode positions for each task. These values were then imported into an Excel package for further analysis in SPSS using a specially written macro. These data were then analyzed using a repeated measures analysis of variance design using SPSS.

Each individual had three data files, one for each recording. (Eyes Closed, Task 1 and Task 2). The electrode positions were arranged into three clusters - midline, paramedial and lateral. These groupings of electrodes for the purpose of the analysis covered the occipital, frontal, central, and temporal regions on the cortex for each hemisphere.

8.3.14 Calculation of an Approximate Cognitive Style Ratio

Each participant completed each of the three tasks already mentioned. The first one served as baseline measurement. As the EEG recording was being taken, its trace on the computers VDU was being monitored by the Author.

In the case of equipment recording malfunction, these recordings were repeated. The EEG trace was time based so that it was relatively easy to calculate the total time taken for each of the two tasks (Tasks 1 and 2) completed by each of the participants while their EEG was being recorded. The approximate cognitive style ratio was calculated by dividing the total time taken to answer the word form type questions divided by the total time

taken to answer the shape questions. It is important to remember that Task 1 consisted only of type questions presented in word form and Task 2 consisted of only shape questions. This contrasts with taking the VItrap under normal conditions where the word and image forms would be in two separate parts in the test, but presented in a fixed mixed order.

These type and shape questions were the same questions used in the VItrap test. While the participants were doing Tasks 1 to Task 2, no feedback was provided to them as to the correctness of their response to each of the test items. This was for a number of reasons: it had already been established that feedback on correctness or incorrectness of response to a test item affected to the subsequent test items. (Vid. Chapter 6 for further details). It was also felt that additional anxiety as a result of this may have had an adverse effect on the individual's recording.

The mean of the ratios were calculated and those below the mean were considered Verbalisers, those above the mean Imagers. The mean in this case was 1.04. This is an approximate means of calculating cognitive style and dividing the group into two different types for the purpose of this analysis. This approach was taken as it was similar to that used by Riding, Glass, Butler and Pleydell-Pearce (1997).

8.3.15 Data Processing & Hardware Calibration

The software, which accompanied the Neuronavigator, provided default settings for the analysis of the EEG data. The hardware was extremely sensitive and at some times some channels malfunctioned during the EEG acquisition phase of the experiment. Where this occurred the recording was deleted from the record file and the participant was asked to come back a second time for another recording. (This occurred on two occasions).

8.3.16 Interactions of interest

Since the variable of interest is cognitive style, the following interactions will therefore be of interest: cognitive style by lead, cognitive style by task, cognitive style by hemisphere in each of the clusters, lateral and paramedial. In the midline, lead by cognitive style, task by cognitive style is what is pertinent.

8.3.17 Statistical Analysis

An analysis of variance for the Verbaliser - Imager style (2 styles) with repeated measures on tasks (3 tasks) was performed on the alpha power output for each of the three locations midline, paramedial and lateral clusters. A repeated measures design was used and the Greenhouse - Geisser (Winer (1971) approach to significance testing was employed when relevant. In this instance, the uncorrected degrees of freedom, the corrected p value, and the epsilon value of the correction factor are reported. Sample screenshots of topographic maps are shown in Appendix 17-8.1.A.

8.4 Results of Study 8.1:

8.4.1 Midline cluster (Fz, Pz, Cz)

An analysis of variance of verbaliser dimension (2) with repeated measures on the anterior -posterior location (Pz, Cz, Fz) and task (3) was performed on the data:

Main effects

Lead df = 1.213, F = 14.214, p = 0.001 (significant)
Task df = 1.077, F = 43.816, p = 0.000

Lead by Cognitive style df = 1.213, F = 0.145, p = 0.755 (ns)
Task by Lead, df = 1.159, F = 17.611, p = 0.000 (significant)
Task by Cognitive Style, df = 1.077, F = 0.237, p = 0.650 (ns)

Lateral cluster (F8, T4, T6, F7, T3, T5)

An analysis of variance of verbaliser dimension (2) with repeated measures on the temporal lobes (3) and task (3) was performed on the data.

Significant Main Effects

Lead df = 1.195, F = 21.132, p = 0.000
Task df 1 = 1.176, F = 29.009, p = 0.000

Lead by Cognitive style df = 1.195, f = 0.125, p = 0.771 (ns)
Hemisphere by Cognitive style df = 1, f = 0.995, p = 0.330 (ns)
Task by Cognitive Style, df = 1.176, F = 0.423, p = 0.554 (ns)
(ns = non-significant).

Paramedial Cluster (Fp2, F4, C4, P4, O2, O1, P3, C3, F5, Fp1)
An analysis of variance of verbaliser dimension (2 dimensions) with repeated measures on the antero-posterior (5 pairs of leads) and task (3 tasks) was performed on the data.

Main effects:
Task df = 5, F = 6.716, p= 0.000
Frequency df = 5, F = 79.778, p = 0.000
Lead d f= 4, f = 4.429, p = 0 .019
Hemisphere df = 1, df = F =0.47, p = 0.832

Interactions
Task by Cognitive style df=2.938, F= 2.136, p= 0.109
Frequency by cognitive style , df 2.116, F= 1.236, p=0.304
Lead by cognitive style, df= 4, F= 1.589, p = 0.218
Hemisphere by cognitive style df= 1, F=0.009, p= 0.925.

8.4.2 Discussion of Study 8.1

These results clearly show that there were no significant interactions with respect to the variable of interest - cognitive style. The following interactions were of interest for the three clusters, lateral, paramedial and midline: cognitive style by lead; cognitive style by task; cognitive style by hemisphere (no midline). In the midline, lead by cognitive style and task by cognitive style were the interactions of interest.

What is interesting is the fact that these results contrast dramatically with those of Riding, Glass, Butler and Pleydell-Pearce (1997). While acknowledging that the experimental design was somewhat different in this study when compared with Riding, Glass, Butler and Pleydell-Pearce (1997) (the EEGs being recorded while subjects were actually doing the VItrap test items), it might have been expected that some significant interactions would have arisen in this instance based on the results reported by Riding, Glass, Butler and Pleydell-Pearce (1997) but this was not borne out. As a matter of fact, the question could be asked as to why Riding, Glass, Butler and Pleydell-Pearce (1997) did not do their study in this obvious manner? This is all the more puzzling as the VItrap test was more reliable than Riding's CSA; also the equipment and software used in this study (Sterman's qeeg Neuronavigator) was more up-to-date and this study had a greater sample size (n

= 19 vs. n = 15) when compared with Riding, Glass, Butler and Pleydell-Pearce (1997).

Additionally no significant interactions were reported for the Verbaliser - Imager dimension and Task type, providing little or no insight as to what happens when matching/mismatching between cognitive style and task occurs.

Why the contrast between the results in this study and that of Riding, Glass, Butler and Pleydell-Pearce (1997). Study 8.1 had a slightly larger (and therefore more powerful) sample size of participants than Riding's Study. Again, as has already been pointed out, the approach taken in the design of the experiment differed somewhat from that of Riding, Glass, Butler and Pleydell-Pearce (1997). This study (n = 19) was based on a study similar in size to Riding's (n = 15) in an attempt to replicate broadly his findings with respect to the new test of cognitive style. The time spent by each of the participants while doing the EEG for each of the tasks was recorded and the ratios were calculated based on these values. (The time taken for completion of the word form - Verbaliser test items was divided by the time taken for the completion of the word form Imager test items).

What was of interest was the fact that cognitive style as assessed was not significant in any of the interactions in this study. The sample size was commensurate with Riding's study with the main difference being the number of tasks being investigated in this study: (one task in Riding et al.'s study versus three tasks in this).

Riding, Glass, Butler and Pleydell-Pearce (1997) did not record any Eyes Closed conditions.

These data from Study 8.1 were also analysed using Eyes Closed as a baseline (These results are not reported in detail in this thesis since Riding et al. (1997) did not measure an Eyes Closed condition). Again there were no significant main effects for cognitive style. Task by cognitive style, frequency by cognitive style, lead by cognitive style, hemisphere by cognitive were all non-significant for each of the three clusters in each of the three regions (midline, lateral, paramedial).

This study did not establish any hemispheric differences between Imagers and Verbalisers, contrary to what was clearly demonstrated by Riding and his co-researchers, even accounting for differences in experimental design. Based on Riding's work it was assumed that there would have been some cognitive style interactions by task and hemisphere, cognitive style by task interactions, cognitive style by hemisphere interactions and cognitive style by lead interactions in this study particularly in the lateral cluster. It appears Riding, Glass, Butler and Pleydell-Pearce (1997) did not use the test items in his CSA in this neurophysiological study. Basically their study was not an adequate test of Riding's CSA.

The results of the study could be interpreted as indicating that VItrap has not been demonstrated to be valid enough from a neurophysiological approach.

More importantly no interactions were found despite the fact that this study had a greater sample size, had more modern equipment and was using a more reliable test (VItrap) of the V-I dimension than Riding's CSA.

How can the results of the study of Riding, Glass, Butler and Pleydell-Pearce (1997) then be interpreted and how confident can one be about their validity? The unreliability of the CSA's Verbaliser-Imager dimension and the design of their experiment is a factor in considering these reported results in the first place.

Is it the case that the results obtained in study 8.1 may be broadly indicative of the recent literature which suggests that there is not equivocal support for the theory of hemisphericity? Perhaps establishing test validity in this instance will only prove possible when more is known about the processing of verbal and imagery information using advanced neuroimaging techniques such as functional magnet resonance imaging.

Of additional interest is what happens to the Imager and Verbaliser when their EEG is being recorded in mixed task and matched conditions. Use of EEG as applied in this study was limited in its ability to pinpoint function.

Gevins et al. (1999) state, that "it is a topic of current research to determine how to combine EEG and fMRI data from the same subjects doing the same tasks".

8.5 Conclusions

Study 8.1 failed to replicate Riding, Glass, Butler and Pleydell-Pearce (1997).

The concept of hemisphericity has not been supported by the Study.

The findings from Study 8.1 may not necessarily imply that the new test of the V-I dimension has not got an underlying neurophysiological basis. The EEG approach appears not to be sensitive enough to pick up any interactions of interest, if any. However the Riding, Glass, Butler and Pleydell-Pearce (1997) approach has not proved useful.

These findings, not only in failing to replicate the Riding, Glass, Butler and Pleydell-Pearce (1997) study, contradict Riding's claim that cognitive style has a neurophysiological basis. This raises the question as to whether the V-I dimension of cognitive style, as assessed by VItrap, from a neurophysiological basis, fails to exist at all. Given the constraints of this particular study, this conjecture remains an open question. A more sensitive EEG approach might be more fruitful.

Consequently, the neurophysiological basis of the VItrap has not been tested and is still an open question.

As stated previously, it was necessary to investigate the reliability and validity of the new test of cognitive style, VItrap, before it could be used with confidence in the interface. It might take several iterations of improvements to get these values to satisfactory levels (if at all). Having done some iterations on reliability in Chapter 6 and validity in Chapters 7 and 8, Chapter 9 examines the accommodation of the V-I cognitive style by embedding the VItrap in the design of some prototype applications.

Chapter 9:
The design and implementation of prototypical applications incorporating VItrap

9.0 Introduction

The purpose of this thesis is to examine the accommodation of individual differences within the human computer interface using cognitive styles. As described in earlier chapters, a new test of cognitive style, Vitrap, was developed for the Verbaliser - Imager dimension (Note that the main interest in the V-I dimension for this thesis was in the extreme ends of the dimension i.e. Verbalisers and Imagers, and not on the Bimodals who it was felt could cope adequately for themselves anyway). Now that a new test has been developed, it is useful to try to incorporate it into some actual user environments - "the proof of the pudding". More formally these can be considered as tests of the construct validity of the Vitrap. Three user applications were originally conceived and considered: i. a biological identification key, ii. an electronic golf stroke saver/simulator and finally iii. a route navigational aid. The considerations of time and expertise available meant that the first option was not implemented. Screenshots of the golf stroke svaer applications can be found in Appendices 19-9.2.B and 20-9.3.C. The other two options were implemented and are discussed in this Chapter.

In Chapter 2, a number of different approaches were examined as to the different ways that flexible systems can be designed and from these it was felt that a mainly non-adaptive system was most appropriate. In Chapter 4, different systems that accommodated cognitive styles were also investigated in Studies 4.2 and 4.3. Various ways of accommodating Cognitive Styles were investigated in these Studies. These Studies ranged from accommodating Cognitive Style by means of cues (e.g. site maps) to the input of cognitive style test scores assessed separately other than by interaction with the interface. This evolutionary trend suggests that it is logical to embed a test of cognitive style directly in the interface. In the prototype applications in this Chapter the VItrap is so embedded. Shneiderman (1998) has commented on the need to accommodate Individual Differences

(such as Cognitive Styles) in the human computer interface but has not yet done so himself.

This chapter describes some prototype systems that accommodate cognitive style non-adaptively. The test of cognitive style was built to function as an adaptor to select the type of human computer interface to be presented to the user based on the test score. The applications were designed to be as realistic as possible even though they were prototypes. Two versions of the Route Navigator/ Simulator application were implemented: one was designed for use using a CD and the other was designed and implemented as a Web-based application. Two versions of the Golf Stroke Saver application were designed and implemented: one was CD-based and the other was designed for a Personal Digital Assistant but could also be run on the Web. All versions were designed by the author of this thesis and were implemented by students as part of final year project requirements closely supervised by the author. The implementations and implementors are given in Table 9.1:

Table 9.1 Applications designed to test cognitive style accommodation

Prototype	Mode	Application	Implementor(s)
Route Navigator 1	Cd-Rom	Authorware 6.5	Colm Moore Post-Grad
Route Navigator 2	Web	Dreamweaver, MySQL	David Fletcher IS5
Golf Stroke Saver 1	Cd-rom	Authorware 6.5	Ultan Murphy ICT4
Golf Stroke Saver 2	Web/ PDA	Java, Tomcat server, Xml	Piotr Kotwicki, Coran McSweeney, Brian Perse, Joseph Gleeson, Stephen Walsh

The cognitive style score for an individual testee was used in the initial run of each prototype to present the testee with the most matched interface and in the next run with the most mismatched interface. The matching hypothesis was discussed in Chapter 4.

Generally it is good practice to assess the usability of an interface. To that end, Shneiderman's Quis (Quis (2006)) was used in Study 9.1.

9.1 Study 9.1: Design of a Route Simulator in Authorware for a Cd-Rom

Aim: To investigate whether the incorporation of the new test of the V-I dimension improved an individual's performance as measured by a Route Simulator using the matching hypothesis.

Purpose of Study 9.1

One postulate is that the most preferred (most matched) route selected by the application using the candidates' cognitive style scores would result in the fastest route completion times for each of the cognitive style groupings (Verbaliser, Imager) when compared with less preferred routes. More formally, this can be phrased as follows:

H0: There is no statistically significant difference between the means of the Verbaliser and the Imager in their most preferred routes (Experiment 1).

H1: There is a statistically significant difference between the means of the Verbaliser and the Imager in their most preferred routes (Experiment 1).

However, there are other issues such as how Imagers (alternatively Verbalisers) perform in the four experiments of greater or lesser match of routes to their cognitive styles; the learning effect and the treatment effect.

9.1.1 Design and Description of the CD based route simulator

The Route Navigator was designed to show a series of screens which represented the progression of junctions while following a particular route through County Wicklow. At each junction a decision had to made as to how to proceed. Where a junction was incorrectly chosen a time penalty was incurred and a decison had to be remade so as to progress through the junction. Based on Riding's, and his co-workers', literature a matched interface for the Imager comprised an image of the junction and an image of a cue. For the Verbaliser a matched interface consisted of both the junction and cue, described textually.

The route chosen was located in County Wicklow and it comprised twenty junctions. The particular data of interest were these that pertained to each junction. Each junction was recorded using video and photographed using a digital camera. The junctions' data were recorded from inside a car so that the perspective was realistic in the design of the application.

The data were recorded using a Sony Digital Recorder, and a small Olympus digital camera. The route data were edited using

Photoshop Elements, Olympus Camedia Software and Microsoft Movie maker. The Simulator was designed and implemented in Authorware 6.5.
The objective was to simulate what direction decision a driver should make upon arrival at a junction. It was necessary to provide the user with a realistic view of the junction when approaching it.

This simulator consisted of four different versions of the one simulated route. The versions were:

(1) A Text (Verbaliser) version – the junction was described textually and the cues were presented textually for each junction.
(2) An Image (Pictorial) version – an image of the junction was presented and the cue was then presented as an image for that junction also.
(3) A Mixed Mode version – the junction was presented in textual form and the cue was presented in image format.
(4) Another mixed mode - the junction was presented in image format and the cue in textual format.

These versions conform to Table 9.2 below.

9.1.2 The Four Experiments

Parts of the VItrap test were incorporated into the Simulator, essentially the sections of the test which contain the textual forms of both Type and Shape questions. The result of the test, depending on how the individual was categorized, determined which of the interfaces of the Route Simulator was presented. In other words if the Vitrap classified the test taker as being an Imager, in Experiment 9.1.1 the testee (Image-based) was presented with an image-based junction and image-based cue. However if the testee was a Verbaliser, he/she was presented with a textual description of the junction and a textual cue.

Next in Experiment 9.1.2, the Imager was presented with a most mismatched condition i.e. a text-based junction description followed by a text-based cue description. A Verbaliser was given his/her most mismatched condition i.e. an image of the junction and an image-based cue.

In Experiment 9.1.3, the imager was given a partially matched/mismatched junction and cue i.e. the junction was

described textually and the cue was given as an image. The same interface was presented to the Verbaliser. In Experiment 9.1.4, another partially mismatched interface with the junction as image and the cue as text for both Imager and Verbaliser was used.

The preferred route's junctions were presented sequentially to the individual in the user's preferred mode. The junctions for the other three non-preferred routes were presented in random order for each non-preferred route in an attempt to reduce any learning effect. The data collected for each of the junctions were a) time taken to choose, and b) number of attempts. The time recorded for each junction was the time taken for the individual to eventually choose the right option.

Although, elsewhere in this thesis, the use of style labels based on the VItrap test was discouraged, here the above procedure, based on style labels, was used in the interest of efficiency. After analysing the VItrap data the following cutoff points/ranges were decided upon:

less than or equal to 0.965 is a Verbaliser
greater than or equal to 1.06 is an Imager
a Bimodal is greater than 0.965 and less than 1.06.

9.1.3 Authorware

The edited route data (digital pictures) were imported into Authorware (2002), where the designed Route Simulator was implemented. Authorware is a generic multimedia package, which facilitates easy design and was used to build this type of application.

The following are some sample screen shots from the application.

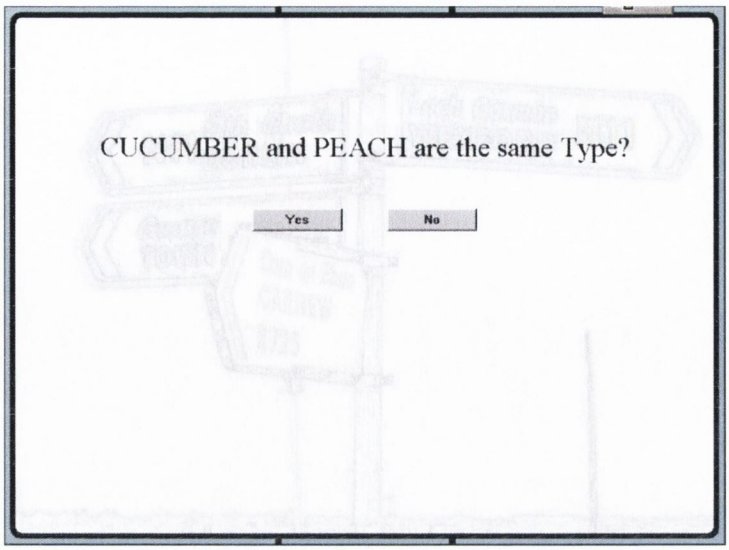


Figure 9.1.3.1: This is a screenshot from the word form of the Vitrap test, which was incorporated as part of the Route Simulator – multimedia format.

The screenshot is an example of a test item which assesses the Verbaliser aspect of the continuum.

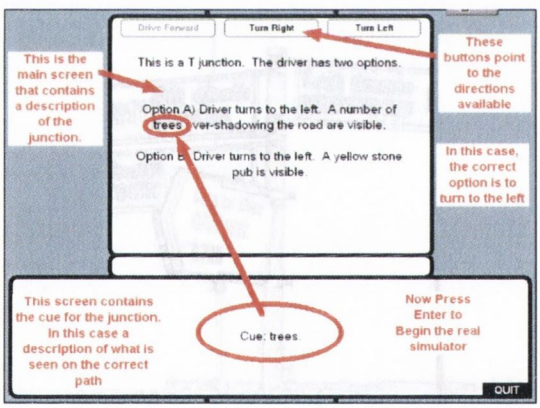


Figure 9.1.3.2: This is a screenshot from the tutorial for the Textual (Verbaliser mode).



Figure 9.1.3.3: A screen shot from a junction in the Imager mode. The cue to the junction is presented at the bottom of the interface.

The first figure illustrates the interface presented to the Verbaliser. Note the cues are presented in the bottom panel of the application. These are highlighted in red.

The second figure illustrates the interface as being compatible with the Imager. Again the cue is presented in the lower panel of the interface. The road sign icons indicate the possibilities presented at this junction.

9.1.4 Procedure:

One hundred and sixty eight mainly Computer Science and Engineering students completed the Route Simulator tests quietly and without any distractions. The four experimental procedures were contained on one cd-rom and were completed sequentially. Before starting the experimental routes the students completed the word forms of the Vitrap test and this determined each student's cognitive styles score. They also completed the Route Simulator tutorial, which provided the students with practice as to how to use the application. Any of the students' queries were answered at this point prior to commencing the application proper. At the end of the session the students evaluated the usability of the entire application using a shortened form of the QUIS. The application saved the data for each individual junction to a floppy disk for further statistical analysis. The experiment ran with no obvious difficulty.

9.1.5 Route Simulator Study 9.1: Analysis of data

The data from the floppy disks were input via an Excel spreadsheet. The data were analyzed using the statistical package R (2006) applying a repeated measures statistical model. The data were presented using boxplots initially. The marginal means for each of the twenty junctions were calculated for each of the four experiments. It was noticeable that there was a learning effect in each of the four marginal mean plots for each of the experiments noticeable for at least the first five or six junctions. A learning effect was discovered which was a confounding factor in testing the hypothesis (Vid. Fig 9.1.8). Since there is no randomisation of candidates to treatment, it is not possible to estimate the size of the learning effect. A design, which included randomisation, would have allowed for this to be calculated but would have needed a much larger sample size for effective calculations.

In order to compensate for learning effects, the first ten junctions were not taken into account in further analyses for each of the four experiments i.e. only the data from the next ten junctions were further analysed. An ANOVA analysis was then applied to this partial dataset using a Repeated Measures design. t-tests were applied as post-hoc tests and confidence intervals were also calculated.

9.1.6 Psychometric Tools.

Questionnaire for User Satisfaction (QUIS)

The usability of the interface was assessed using Shneiderman's QUIS developed in the University of Maryland (Quis (2006)). This is a questionnaire consisting of twelve sections. Each section evaluates an aspect of the usability of the software using a nine point Likert scale. It assessed aspects of the interface such as: overall reactions to the system, screen characteristics, learning and online tutorials. At the end of each section there was further opportunity to add a written comment if the respondent desired.

The QUIS assesses the users' satisfaction in relation to the design of the interface. Chin's comment is appropos: "Although a system is evaluated favorably on every performance measure, the system may not be used very much because of the user's dissatisfaction with the system and its interface " (Chin, Diehl and Norman (1988)).

The questionnaire was not administered in full, as some of the sections could not be meaningfully interpreted within the context of the application. The questionnaire has been reported as being reliable. Chin et al (1988) report Chronbach's alpha for it as being at 0.89. However in this instance, given that the version administered was attenuated the internal reliability coefficients were calculated instead. The internal alpha in this case was estimated at 0.975.

9.1.7 VItrap

This contained test items, which assessed the Verbaliser - Imager dimension in word form. A more detailed description of VItrap is presented in Chapter 6.

9.1.8 Study 9.1 Results (n = 168)

Four experiments were run; the details of them are given in Table 9.2

Table 9.2 Versions of the interface used by each Group in sequence (Experiment x uses Interface x)

Route	Navigator	CD Version			
Group	Cognitive Style Category	Experiment 1 Matched Junction/ Cue	Experiment 2 Most Mismatched Junction/ Cue	Experiment. 3 Mismatched Junction/ Cue	Experiment 4 Mismatched Junction/ Cue
1	Verbaliser	Text/ Text	Image/ Image	Text/ Image	Image/ Text
2	Bimodal	Text/ Image	Image/ Text	Image/ Image	Text/ Text
3	Imager	Image/ Image	Text/ Text	Text/ Image	Image/ Text

Study 9.1 Experiment 1

In Experiment 1, the groups were presented with their most preferred modes i.e. Imagers were presented with Junction and Cue in image form; Verbalisers were presented with Junction and Cue in text format.

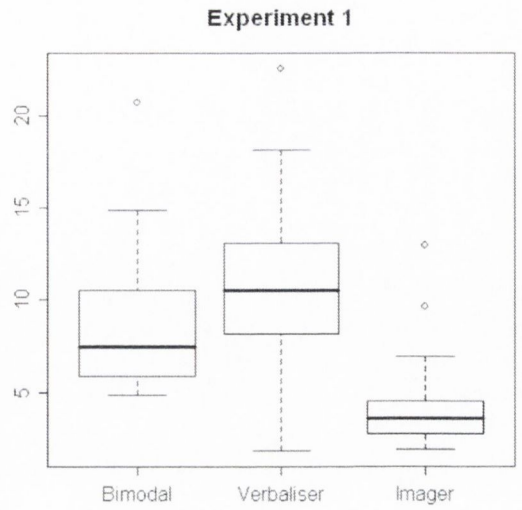


Figure 9.1.1.1 Study 9.1 Experiment 1: boxplots for Bimodal, Verbaliser and Imager

The route completion times (y-axis) are measured in seconds. The median route completion time for the Imager is much faster than the Verbaliser in matched conditions.

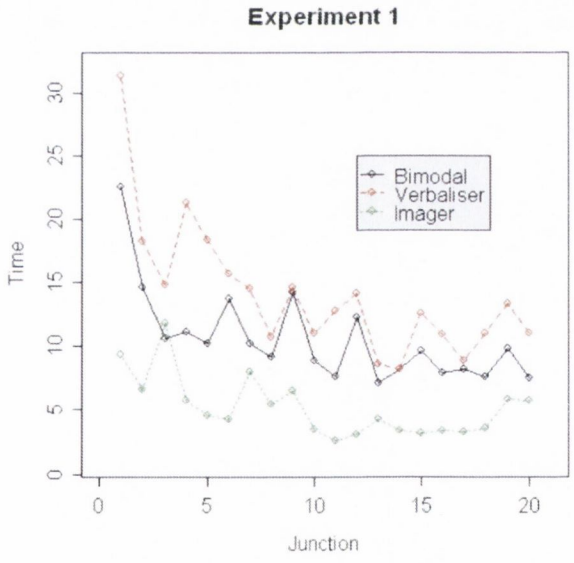


Figure 9.1.1.2: Study 9.1 Experiment 1: Marginalized means for each of the three groups at each of 20 junctions i.e. the mean route completion time for each of the three categories for each junction

Comment: The graph illustrates a learning effect over approximately the first ten junctions. To correct for this apparent learning effect, the data for the first ten junctions were removed

from the remaining analyses i.e. the last ten junctions only were used.

The statistical analysis for Study 9.1 Experiment 1 is given in Appendix 22-9.6.F. The pertinent t-test from this analysis is given here. An ANOVA model was applied initially and the t-tests serve as post-hoc tests.

********** Statistical Analysis Experiment 1 **********

Post Hoc t-tests Experiment 1

Welch Two Sample t-test for Group 1 (Verbaliser) and Group 3 (Imager)
data: J1S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]
t = 6.3387, df = 25.971, p-value = 1.039e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (3.210651 6.292531)
sample estimates:
mean in group 1 mean in group 3
8.579996 3.828405

********** End of Statistical Analysis **********

9.1.9 Study 9.1 Experiment 2

In this experiment the groups were presented with their least preferred modes i.e. Imagers were presented with Junction and Cue in text; Verbalisers were presented with Junction and Cue in image format.

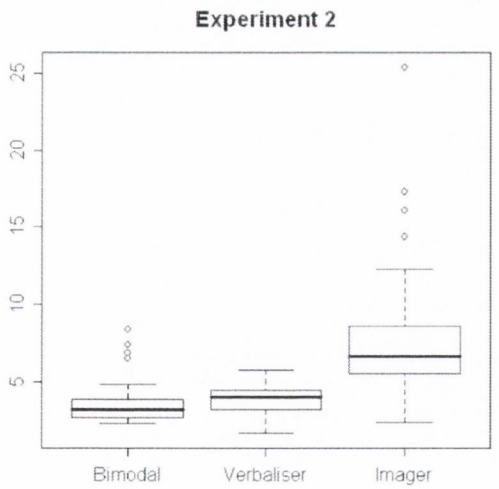


Fig. 9.1.2.1 Study 9.1 Experiment 2 - boxplots
Verbalisers perform better than Imagers in most mismatched conditions.

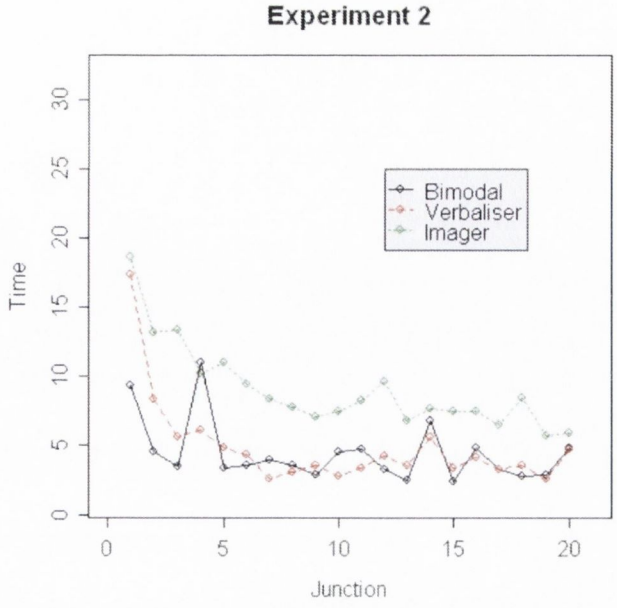


Fig. 9.1.2.2 Study 9.1 Experiment 2: Marginalized means for each of the three groups at each of 20 junctions

Results of Study 9.1 Experiment 2

********** Statistical Analysis Experiment 2 **********
Welch Two Sample t-test
data: J2S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]
t = -7.9739, df = 67.469, p-value = 2.567e-11
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-4.439797 -2.662257)
sample estimates:
mean in group 1 mean in group 3
3.839948 7.390975
********** End of Statistical Analysis **********

9.1.10 Study 9.1 Experiment 3

In Experiment 3 the groups were presented with their in-between preferred modes i.e. Imagers were presented with Junction as text and Cue as image; similarly Verbalisers were presented with Junction as text and Cue in image format.

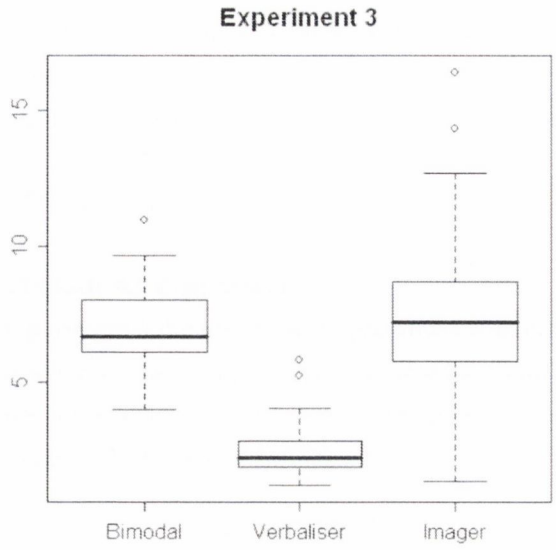


Fig. 9.1.3.1 Study 9.1 Experiment 3: boxplots
Verbalisers perform better than Imagers in partial mismatch - Text for junction, Image for cue.

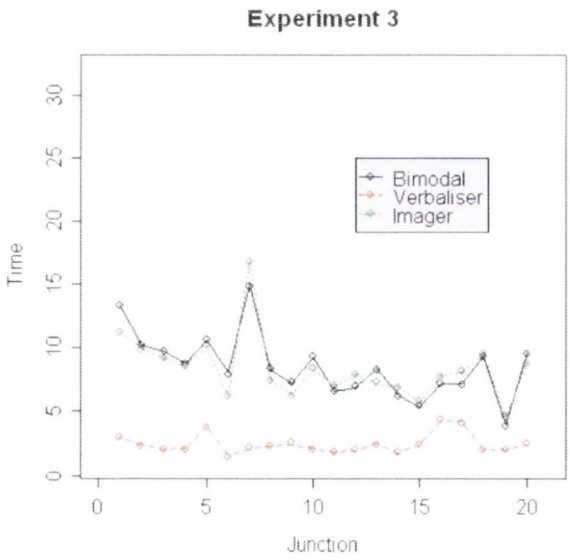


Fig. 9.1.3.2 Study 9.1 Experiment 3: Marginalized means for each of the three groups at each of 20 junctions

Results Study 9.1 Experiment 3

********** Statistical Analysis Experiment 3 **********
Welch Two Sample t-test
data: J3S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]
t = -0.6714, df = 46.994, p-value = 0.5053

alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-1.1006348 0.5498247)
sample estimates:
mean in group 1 mean in group 3
7.091644 7.367049
********** End of Statistical Analysis **********

9.1.11 Study 9.1 Experiment 4

In Experiment 4 the groups were presented with another in-between mode i.e. Imagers were presented with Junction in Image and Cue in text; Verbalisers were presented with Junction in Image and Cue in text format.

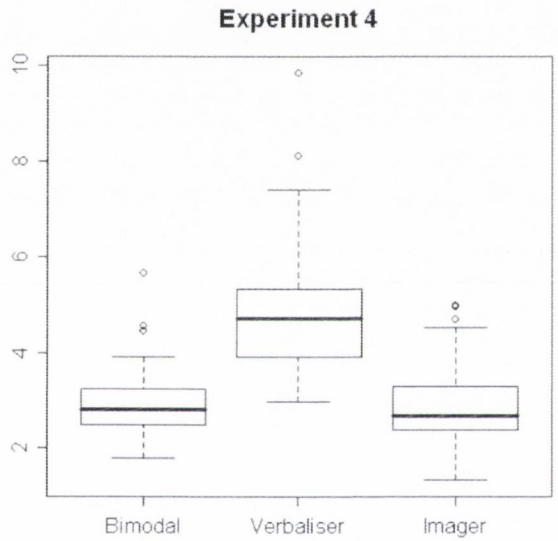


Fig. 9.1.4.1 Study 9.1 Experiment 4: boxplots

Imagers perform better than Verbalisers - Image for junction, Text for cue. However this apparent different is not statistically significant (See analysis in Appendix 22-9.6.F).

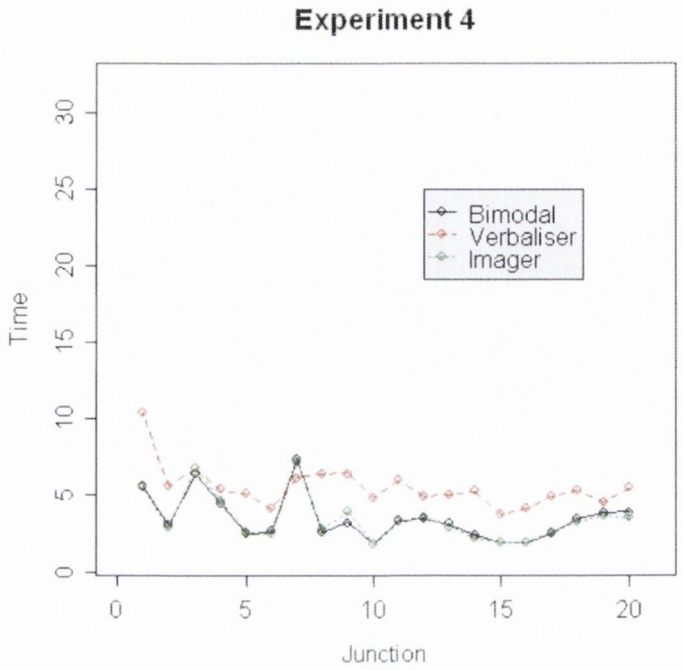


Fig. 9.1.4.1 Study 9.1 Experiment 4: Marginalized means for each of the three groups at each of 20 junctions

t-test Results Experiment 4

********** Statistical Analysis Experiment 4 **********
Welch Two Sample t-test
data: J4S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]
t = 0.7109, df = 32.558, p-value = 0.4822
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-0.2524230 0.5233843)
sample estimates:
mean in group 1 mean in group 3
3.013364 2.877883
********** End of Statistical Analysis **********

9.1.12 Combined Plot for all Four Experiments and Groups

To study the learning effect a bit more, all the plots for the four experiments were grouped into one plot (Fig.9.1.8). For the Verbaliser, there seems to be a consistent learning effect for all of the first three experiments (Junctions 1 to 41), but with a slight jump for Experiment 4 (Junctions 61 - 80). For the Imager, there seems to be a learning effect with each new experiment from 1 to 3 (i.e. at Junction 1, Junction 21 and Junction 41). However

Experiment 4 (junctions 61 - 80) appears to carry on the learning effect from Experiment 3.

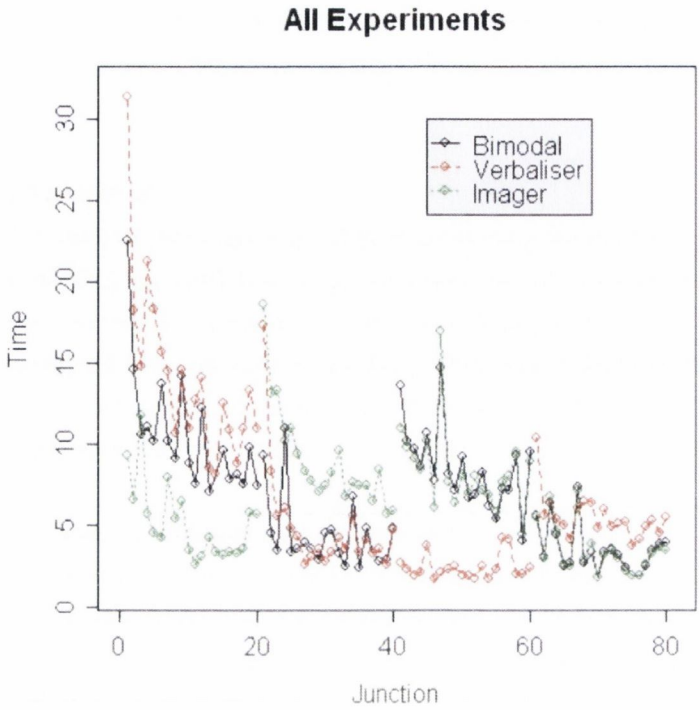


Fig 9.1.Overall Study 9.1 (n=168) All four experiments: Learning Effect

9.1.13. The Internal Reliabilty of a Truncated Quis

In this section the internal reliability of the QUIS is calculated based on Study 9.1 results. This was necessary as only some of the questions from the QUIS were used since some of the sections were not that relevant to the evaluation of this type of application. The internal reliability was calculated using SPSS.

Table 9.1.5: The internal reliability of sections used in the appended sections of the QUIS.

Cronbach's Alpha	Cronbach's Alpha based on Standardised Items	Items
0.975	0.975	N = 58

9.2 Study 9.2 (IS5 Group Route Simulator (n = 40))

9.2.1 Introduction - Study 9.2 CD-rom Study with IS5

A different sample of students from that originally involved in this study completed the Route simulation as part of the final

year course in HCI. Study 9.2 repeated Study 9.1 with a different sample of IS5 students (n = 40, mean age = 29)) and under different experimental conditions. The students in the sample completed the Route Navigator presented on CD at home under their own time. Otherwise the experiment was identical to those carried out in Study 9.1.

9.2.2 Method

The students were given a CD-Rom containing the Route Simulator to complete in their own time. They all then returned the Cd-Rom one week later along with a floppy disk which contained the route data for analysis. They were asked to set aside some time that would allow them to complete the application uninterrupted.

9.2.3 Results of Study 9.2 Experiment 1

In Experiment 1, the groups were presented with their most preferred modes i.e. Imagers were presented with Junction and Cue in image form; Verbalisers were presented with Junction and Cue in text format.

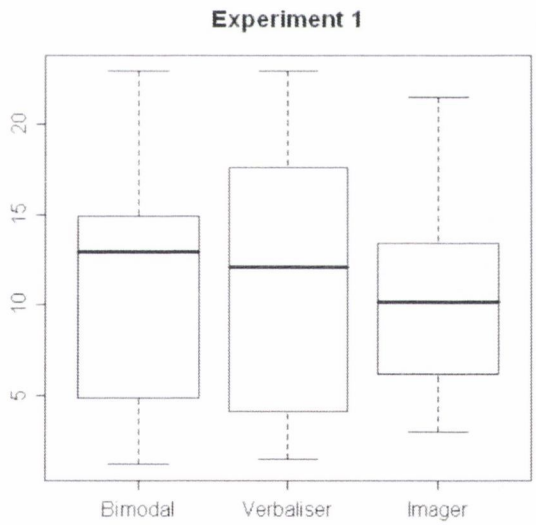


Fig. 9.2.1.1 Study 9.2 Experiment 1: boxplots

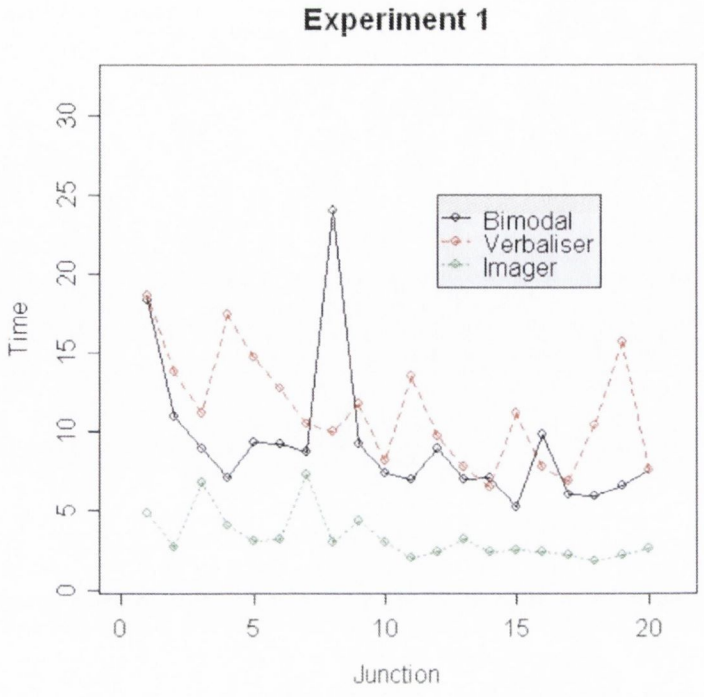


Fig. 9.2.1.2 Study 9.2 Experiment 1: Marginalized means for each of the three groups at each of 20 junctions

9.2.4 Study 9.2 Experiment 1 Analysis

********** Statistical Analysis Study 9.2 **********

Welch Two Sample t-test

data: J1S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]

t = 7.0348, df = 16.783, p-value = 2.156e-06

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: (3.316633 6.162269)

sample estimates:

mean in group 1 mean in group 3

7.118242 2.378791

********** End of Statistical Analysis **********

9.2.5 Study 9.2 Experiment 2

In this experiment the groups were presented with their least preferred modes i.e. Imagers were presented with Junction and Cue in text; Verbalisers were presented with Junction and Cue in image format.

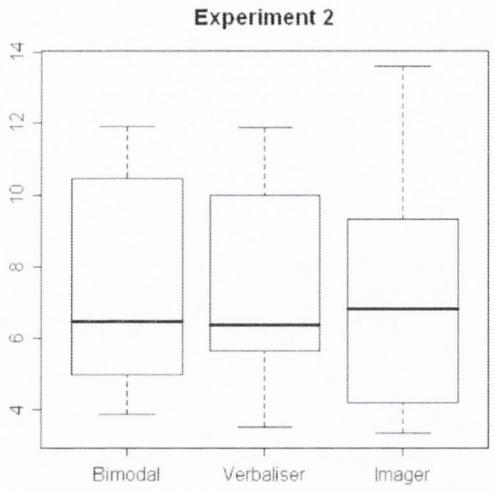


Fig 9.2.2.1 Study 9.2 Experiment 2: boxplots

In Experiment 9.2 Experiment 2 the Verbaliser had a median route completion time slightly lower than that of the imager in most mismatched conditions.

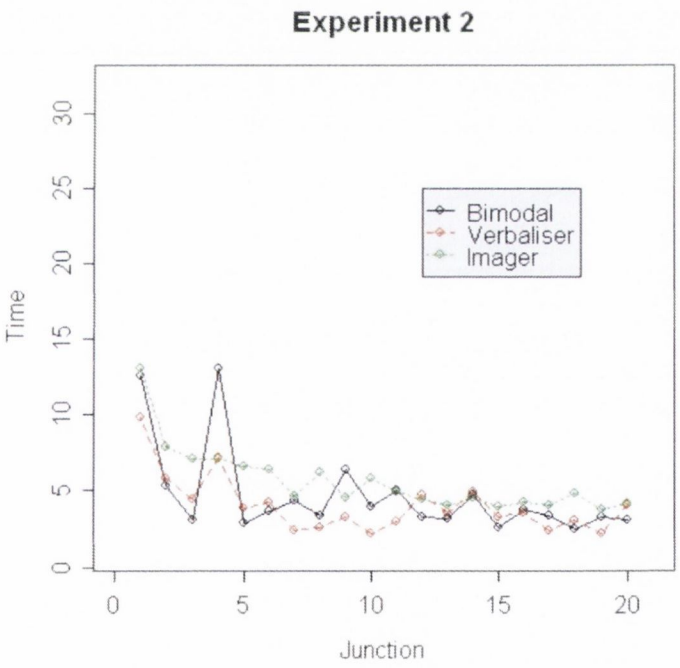


Fig. 9.2.2.2 Study 9.2 Experiment 2: Marginalized means for each of the three groups at each of 20 junctions

********** Statistical Analysis Study 9.2 **********

Welch Two Sample t-test

data: J2S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]

t = -0.8545, df = 13.722, p-value = 0.4075
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-2.841906 1.224828)
sample estimates:
mean in group 1 mean in group 3
3.488725 4.297264
********** End of Statistical Analysis **********

9.2.6 Study 9.2 Experiment 3

In Experiment 3 the groups were presented with their in-between preferred modes i.e. Imagers were presented with Junction as text and Cue as image; similarly Verbalisers were presented with Junction as text and Cue in image format.

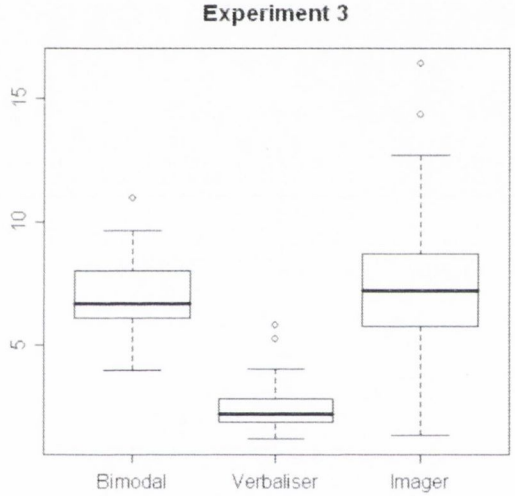


Fig. 9.2.3.1 Study 9.2 Experiment 3: Boxplots
Verbalisers appear to have done much better than Imagers in a partially mismatched situation (Junction = text, Cue = Image)

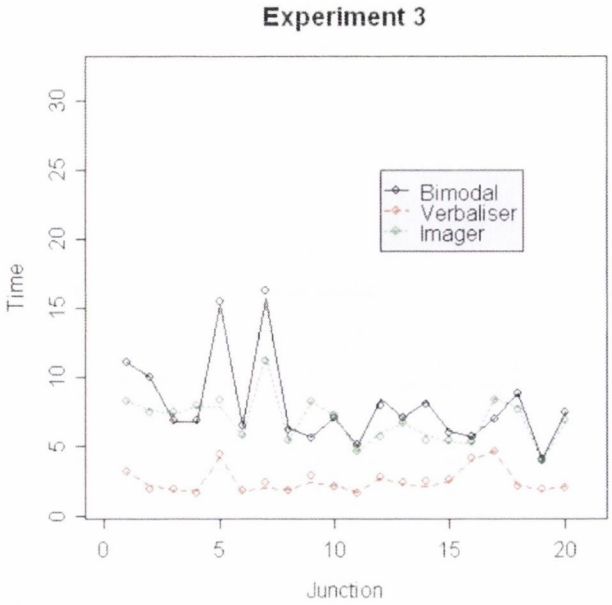


Fig. 9.2.3.2 Study 9.2 Experiment 3: Marginalized means for each of the three groups at each of 20 junctions

********** Statistical Analysis Experiment 3 **********
Welch Two Sample t-test
data: J3S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]
t = 0.4982, df = 14.425, p-value = 0.6258
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-2.453052 3.943017)
sample estimates:
mean in group 1 mean in group 3
6.834492 6.089509
********** End of Statistical Analysis **********

9.2.7 Study 9.2 Experiment 4

In Experiment 4 the groups were presented with another in-between mode i.e. Imagers were presented with Junction in Image and Cue in text; Verbalisers were presented with Junction in Image and Cue in text format.

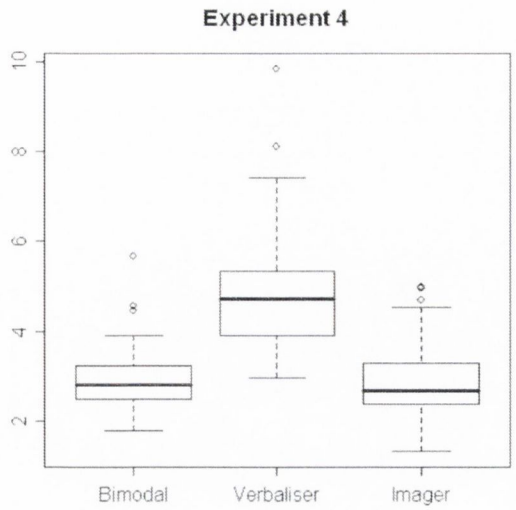


Fig. 9.2.4.1 Study 9.2 Experiment 4 boxplots

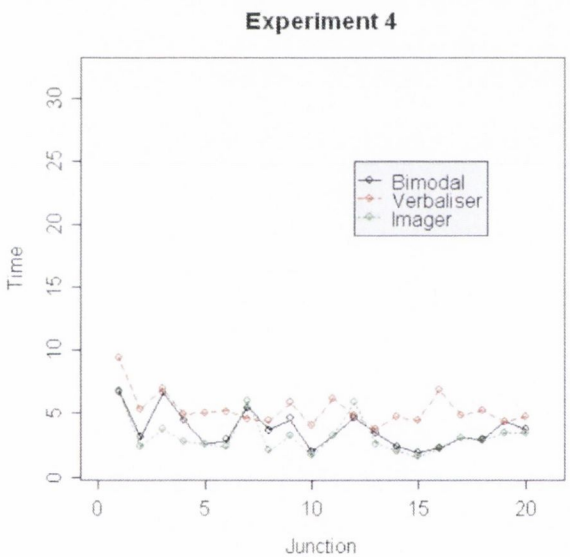


Fig. 9.2.4.2 Study 9.2 Experiment 4: Marginalized means for each of the three groups at each of 20 junctions

********** Statistical Analysis Experiment 4 **********

Welch Two Sample t-test

data: J4S[Group== 1 | Group== 3] by Group[Group== 1 | Group== 3]

t = 0.3858, df = 17.647, p-value = 0.7043

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: (-0.7428851 1.0765063)

sample estimates:

mean in group 1 mean in group 3

3.220883 3.054073

********** End of Statistical Analysis **********

Finally all results were plotted in one graph (Fig. 9.3)

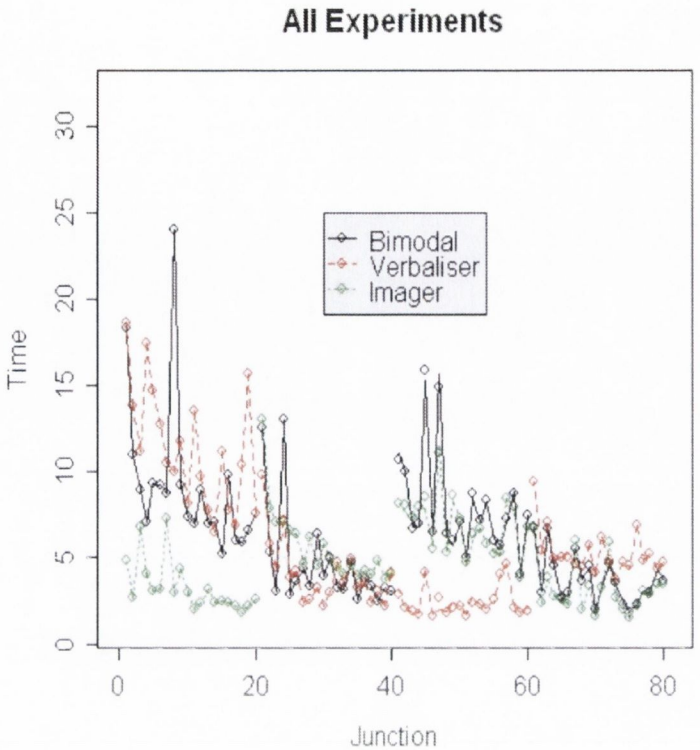


Fig 9.2.Overall Study 9.2: All four experiments: Learning effect

The pattern for the Verbaliser and Imager mimic the patterns previously obtained in Study 9.1. Further statistical analysis for Studies 9.1, 9.2 and 9.3 are available in Appendices 23-9.6.F, 24-9.6.F, 25-9.6.F, 26-9.6.F, 27-9.6.F and 28-9.6.F.

9.3 Study 9.3 Web-based Route Simulator (n = 20)

9.3.1 Introduction to Study 9.3

It was necessary to design and develop a Web-based application of the Route Simulator. This type of an application whilst being a prototype has many other potential applications. Initially it was hoped to design this application incorporating GIS technology. This was subsequently not possible for economic reasons since the various alternative Geographical Information Systems available were commercial products, which were prohibitively expensive. Instead the data collected for use in the design of the Cd-Rom were also used in the design of this application. Further screenshots of the Web Route Simulator can be found in Appendix 21-9.4.D.

9.3.2 Design of the Web-based Route Simulator.

A web-based version of this application was also developed using Dreamweaver, PHP and a MySQL database. The application ran on an APACHE server.

The design of the application was slightly different to that used in the Cd-Rom. As well as having three route formats an additional one was included. The final mode differed insofar as the junction was described textually and the junction cue was an image. The junctions in all the web-based routes were presented in sequence not randomized, as was the case in the Cd-Rom application.

9.3.3 Method

Prior to the application being run on the server it was tested on a number of occasions as to how it would perform under load.

Twenty students Information System Students participated in this experiment. These students were drawn from a different sample of students who had not previously participated in any part of this study. The students logged on initially and completed the cognitive styles test, then registered to use the application.

The students completed the application without any difficulty. The server recorded all of the data in the underlying MySQL database successfully. The server was able to cater for the load and didn't crash.

9.3.4 Design of the Application

The design and implementation of this application was achieved using Dreamweaver to design the website, and open source client and server technologies (Tomcat, XML) to implement the application.

These web pages were created in Dreamweaver by means of templates. The use of templates in developing this application facilitated consistency and easy of development.

Underpinning the development of this application was the client server model. This is a well-known protocol. Simply put, the browser sends a request to a web server for a resource; the web

server locates the resource and sends it back to the client. The client sends a request to the server, which in turns sends a request via PHP scripts to the MySQL server to the MySQL database installed on it. When the request has been analyzed the appropriate information is sent via the initial server and then finally back to the client. An example of a request would be one for a web page; in this instance the web page would be displayed on the client's web browser.

As already mentioned above PHP was used as a server side scripting language and had the additional benefit of being able to be integrated easily with MySQL and Dreamweaver. The web server used to develop the application was Apache. Like MySQL and PHP it was open source and was compatible with these also. The site was hosted on an Apache server, located at the Computer Science Dept at Trinity College.

9.3.5 Screenshots of the application

Some sample screenshots of the application are given below. Other can be found in Appendix 20-9.4.D

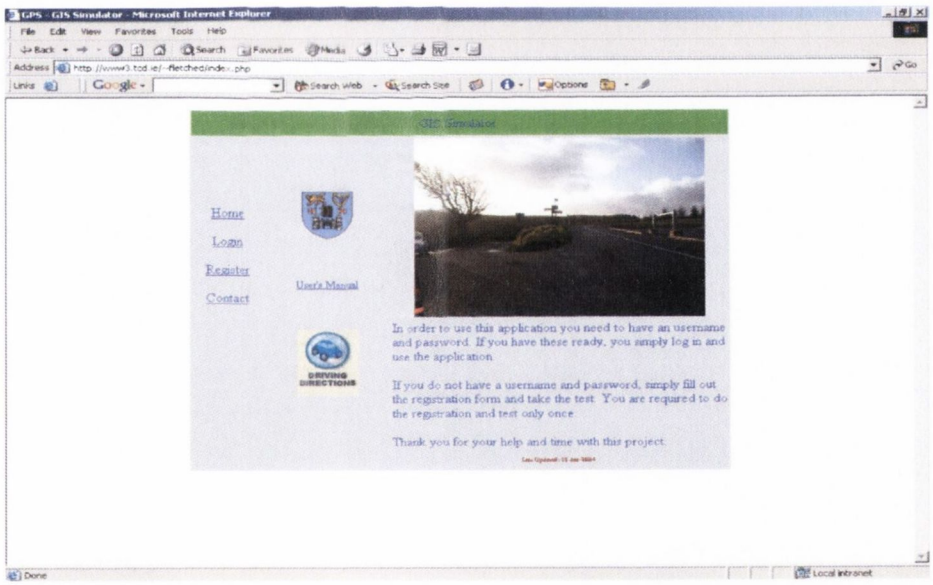


Fig. 9.3.5.1 Sample screenshot of the Route Navigator - Introductory Screen

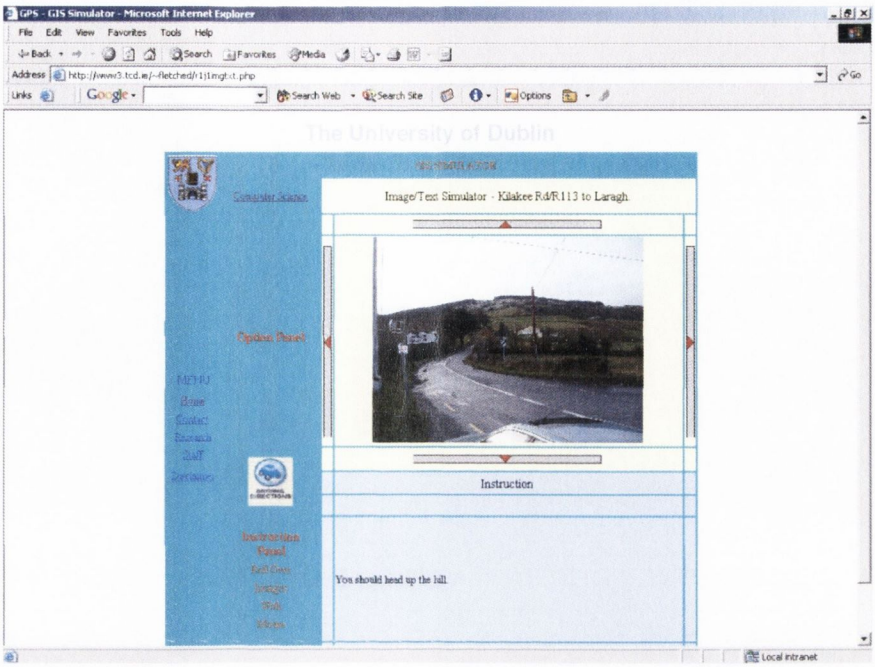


Fig. 9.3.5.2 Sample screenshot of the Route Navigator - Image junction, Verbal cue

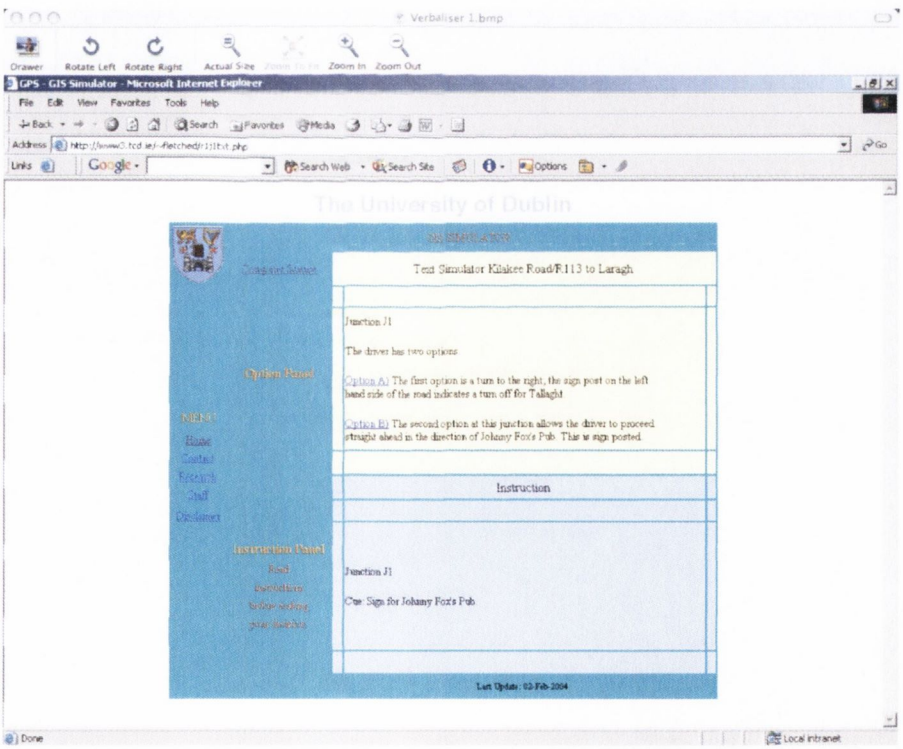


Fig. 9.3.5.3 Sample screenshot of the Route Navigator - Image junction, Verbal cue

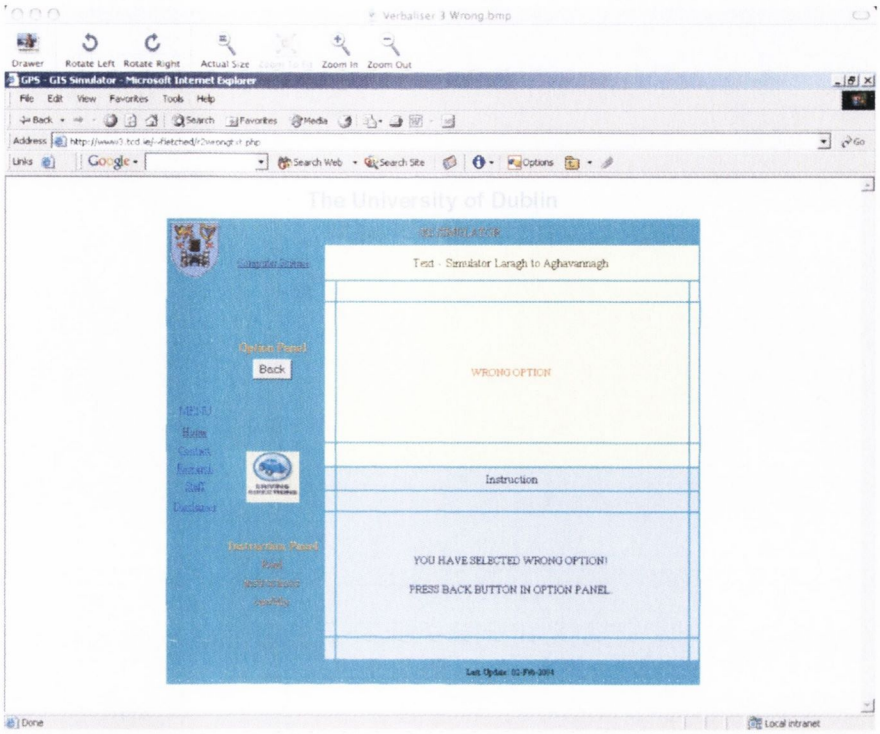


Fig. 9.3.5.4 Sample screenshot of the Route Navigator - Wrong junction - feedback

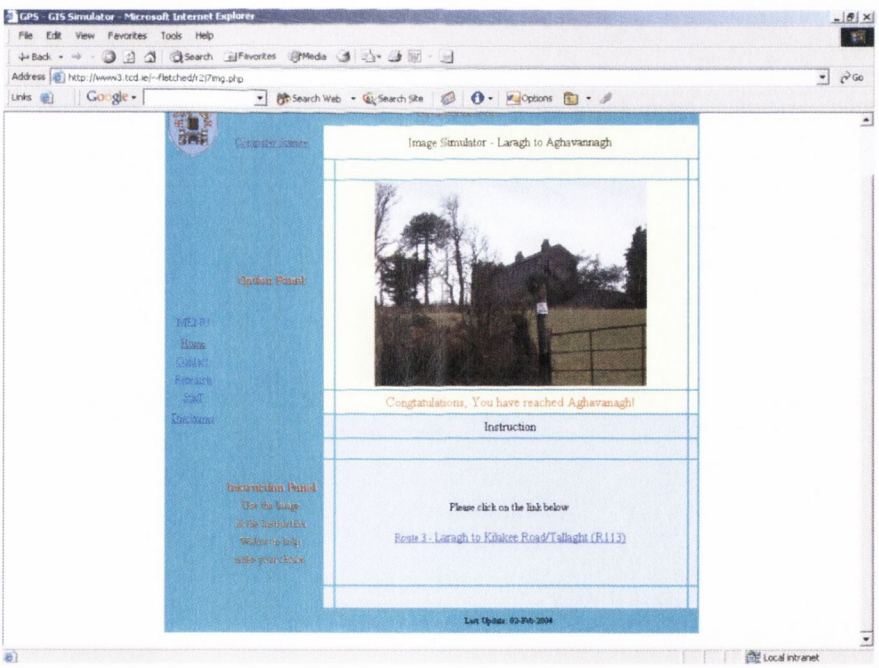


Fig. 9.3.5.5 Sample screenshot of the Route Simulator - Image junction, Verbal cue, Route completed

9.3.6 Introduction - Study 9.3 Web-based Route Simulator

A different sample of students from that originally involved in this study completed the Route Simulation as part of the final

year course in HCI. Study 9.3 repeated Studies 9.1 and 9.2 with a different sample of IS5 students (n = 20, mean age = 29)) and under different experimental conditions. The students in the sample completed the Route Simulator presented on a Web-based environment at the TCD computer laboratories.

9.3.7 Results of Study 9.3 Web version of Route Simulator

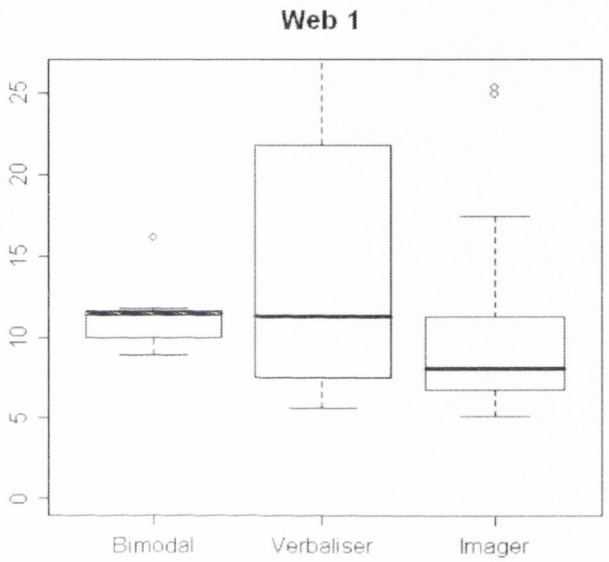


Fig. 9.3.1.1 Web - Experiment 1 Analysis: Boxplot - 3 Groups - Most Matched Condition Junctions 10 - 20

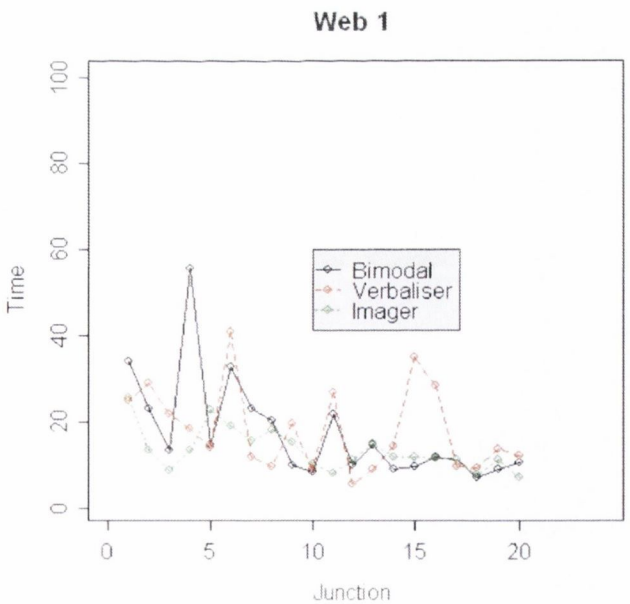


Fig. 9.3.1.2 Web - Experiment 1 Analysis: Plots of Marginalised Means - 3 Groups - Most Matched Condition Junctions 10 - 20

********** Statistical Analysis**********

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor(Group)	2	329.4	164.7	2.0601	0.1407
Residuals	40	3197.4	79.9		

Verbaliser = Group 2 and Imager = Group 3

Welch Two Sample t-test - Most Matched

data: T1S[Group == 2 | Group == 3] by Group[Group == 2 | Group == 3]

t = 1.809, df = 25.164, p-value = 0.08242

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: (-0.8034276 12.4367610)

sample estimates:

mean in group 2 mean in group 3

16.33889 10.52222

********** End of Statistical Analysis**********

9.3.8 Study 9.3 Web - Experiment 2

In this experiment the groups were presented with their least preferred modes i.e. Imagers were presented with Junction and Cue in text; Verbalisers were presented with Junction and Cue in image format.

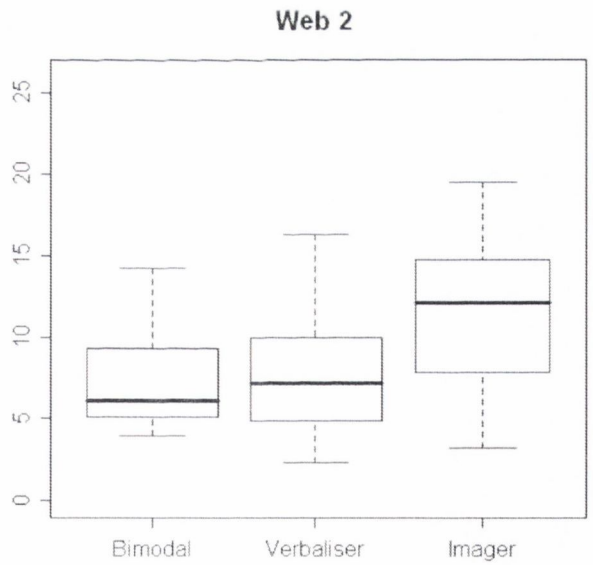


Fig. 9.3.2.1 Web - Experiment 2 Analysis: Boxplot - 3 Groups - Most Mismatched Condition Junctions 10 - 20

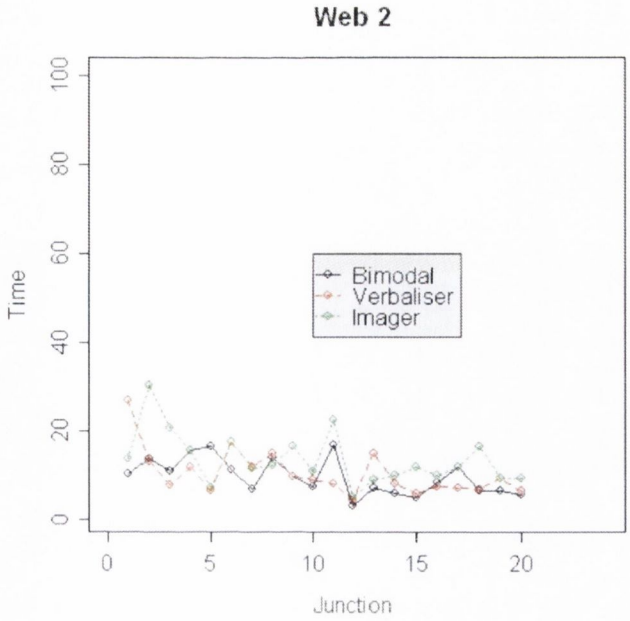


Fig. 9.3.2.2 Web - Experiment 2 Analysis: Plots of Marginalised Means: 3 Groups - Most Mismatched Condition Junctions 10 - 20

********** Statistical Analysis**********

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor(Group)	2	141.77	70.88	4.2969	0.02040 *
Residuals	40	659.86	16.50		

Welch Two Sample t-test - Most Mismatched

data: T2S[Group == 2 | Group == 3] by Group[Group == 2 | Group == 3]

t = -2.6588, df = 33.603, p-value = 0.01192

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: (-6.4018382 -0.8537173)

sample estimates:

mean in group 2 mean in group 3

7.766667 11.394444

********** End of Statistical Analysis**********

9.3.9 Study 9.3 Web - Experiment 3

In Experiment 3 the groups were presented with another in-between mode i.e. Imagers were presented with Junction in Image and Cue in text; Verbalisers were presented with Junction in Image and Cue in text format.

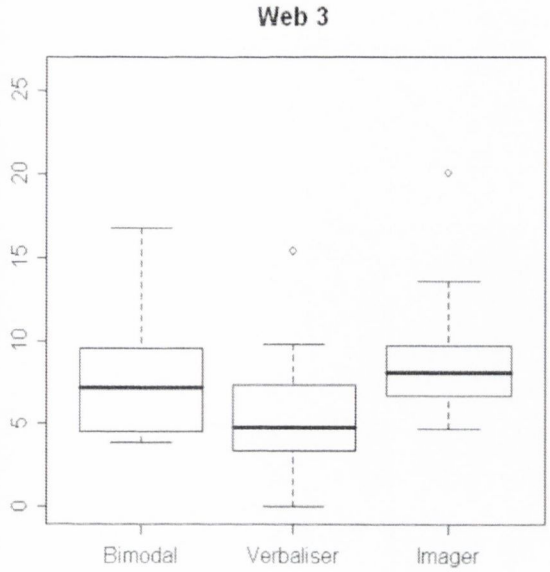


Fig. 9.3.3.1 Web - Experiment 3 analysis: Boxplot - 3 Groups - Partially Matched Condition Junctions 10 - 20

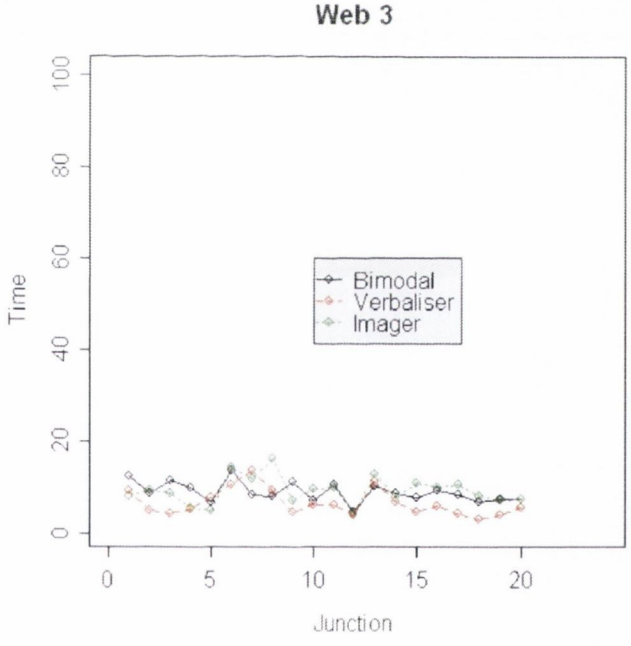


Fig. 9.3.3.2 Web - Experiment 3 Analysis: Plots of Marginalised Means - 3 Groups - Partially Matched Condition Junctions 10 - 20

********** Statistical Analysis**********

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
factor(Group)	2	100.90	50.45	3.6499	0.03499	*
Residuals	40	552.91	13.82			

Welch Two Sample t-test - Partially Mismatched

data: T3S[Group== 2 | Group== 3] by Group[Grou == 2 | Group== 3]

t = -2.7696, df = 34, p-value = 0.009027

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: (-5.6829392 -0.8726164)

sample estimates:

mean in group 2 mean in group 3

5.527778 8.805556

********** End of Statistical Analysis**********

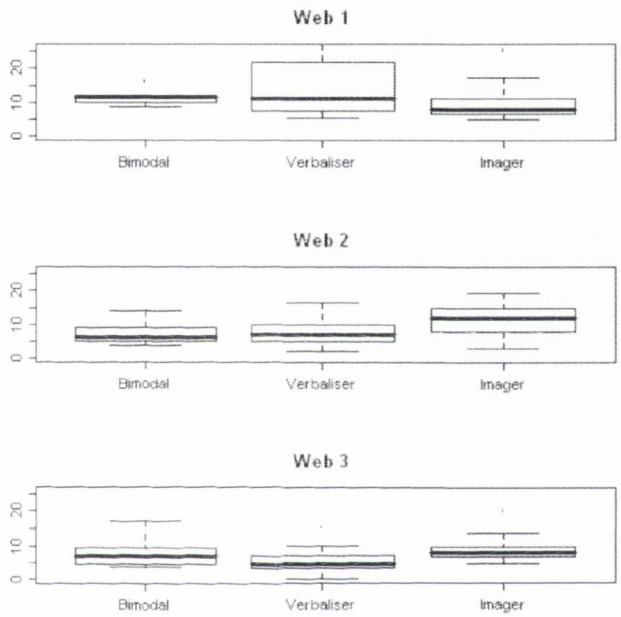


Fig. 9.3.Overall.A Web -Analysis Boxplot - All 3 Groups: - Most Matched Condition Most Mismatched and Partially Matched Junctions 10 - 20

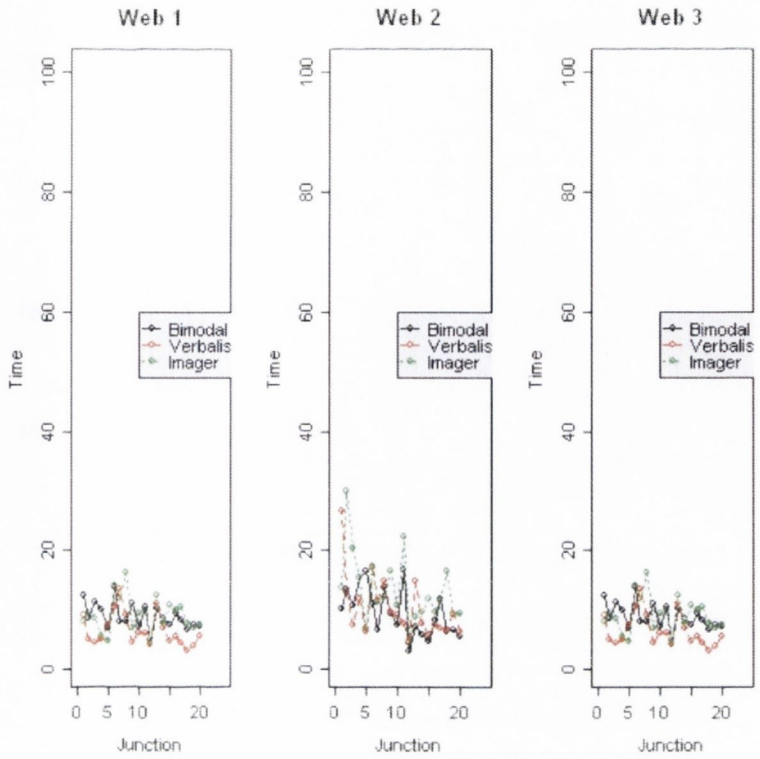


Fig. 9.3.Overall.B Web Analysis Plots of Marginalised Means: - All 3 Groups - Junctions 10 - 20

9.4 Study 9.4 and Study 9.5

As stated already in the Introduction to this Chapter, three user applications were originally conceived and considered: i. a biological identification key, ii. an electronic golf stroke saver/simulator and finally iii. a route navigational aid. The considerations of time and expertise available meant that the first option was not implemented.

The two golf applications, which were implemented by two student groups, (designed by the Author) were implemented. One was web-based and the other was Cd-rom based. They worked well but proved difficult to evaluate. Consequently they were considered not particularly pertinent to the purpose of this thesis. Appendices 9.3.C and 9.4.D provide some screenshots of these applications.

9.5 Discussion Pertaining to Studies 9.1, 9.2 and 9.3

There were few alternatives available when considering the design of the Route Simulator. One of the objectives was to prove

the concept that Cognitive Style could be used as an adaptor to generate a matching HCI to optimise user performance.

Initially a Geographical Information System was considered. This has been defined as being "a computer system for capturing, managing, integrating, manipulating, analyzing, and displaying data which is spatially referenced to the earth" McDonnell & Kemp (1995). The use of GIS facilitates the easy manipulation of spatial information. There are numerous possible applications of GIS systems, mapping the human body for instance, maintaining spatial databases, mapping population distributions useful in the context of business applications. A GIS system would also have been applicable in the case of the Route Simulator application. In this case spatial information could have been referenced using latitude and longitude and represented using raster or vector models.

Different layers in a GIS application store varying levels of geographical information. In route simulators such layers might have included railway and road networks, vegetation demarcations, property boundaries and different categories of urban settlements. The layers would then have been re-integrated to form a specific route simulator for the greater Co. Wicklow area. The geo database containing the vector model data was not complete so therefore the application could not be developed in this manner. It was also too costly to update the database.

Another way to design this type of application is to use GPS. However this was not realistic due to the prohibitive cost.

The only solution was to collect data along the routes, which pertained to the junctions. It was necessary to photograph each junction extensively and in some cases the quality of the images was hampered due to weather conditions. It was also hazardous as the task demanded stopping at very narrow busy junctions for some times up to ten minutes. It was also necessary to describe each of the twenty junctions textually and in a consistent manner. At times this proved difficult. Generally most junctions along the routes chosen did not provide much variety in change of direction. Usually the option was to drive straight through; this may have also limited the design of the application.

Initially every effort was made to insure that the design of these applications were as realistic as possible. The design of a Route Simulator such as this proved to be much more difficult than had been previously envisaged.

The statistical analysis revealed a very significant learning effect despite the randomization of the junctions in Presentations 2 and 3 to each individual in the multimedia CD.

As the application had been designed to present to the individual their optimal route mode based on their cognitive style score there wasn't scope for randomization.

9.6 Quis Analysis - Usability Results

Shneiderman's QUIS analysis was used to examine the usability of the Route Simulator used in Study 9.1.The subsequent analysis of the Quis results are given in Appendix 18-9.1.A. Pertinent comments on the Quis analysis are given next. The remaining results on the individuals' overall evaluation of the system are presented in Appendix 29-9.7.G.

The users' reactions to the system in general were presented by the histograms in Figures 9.1.12.1 and 9.1.12.38 in Appendix 18-9.1.A. The average score for each of the six scales is approximately five and a half. This score indicates an average satisfaction rating amongst the users of the Route Simulator system. A more detailed analysis is presented in Appendix 18-9.1.A.

There is some variation in the sample sizes that responded to the different questions in the QUIS. When an individual responded to a question by ticking the N/A option, these responses were coded as being zeros and were excluded from the analysis.

The results of the Quis can be grouped into the following categories: (a) General reaction to the system (b) Ease of interface interpretation (c) Ease of use (d) Evaluation of the tutorial (e) and (e) Quality of photographs.

9.6.1 (a) General reaction to the system

It can be concluded (Fig. 9.11.1 to 9.11.6 in Appendix 18-9.1.A) that there was an average degree of satisfaction expressed about

this system with a mean for overall satisfaction with the system being 5.1 (out of 9).

9.6.2 (b) Ease of interface interpretation

The participants found that the text was quite easy to read and that the screen layouts were helpful. These findings are important as they indicate that the users did not experience much difficulty in interpreting how to use the various sets of interfaces in each of the experiments. They also found that the amount of information that was displayed on the interfaces to be quite adequate with a mean of 6.08 (Fig. 9.1.11.11 in Appendix 18-9.1.A).

9.6.3 (c) Ease of use

When the various histograms are interpreted under this heading, use of the system and specific aspects of doing the task, the means ranged from 6.1 to 7.23. What can be assumed from this is that the tasks i.e. use of the Route Navigator didn't present problems to those using the System. This is confirmed by a mean of 7.19 being recorded by the question: "Is learning to use the system difficult to easy enough?" (Fig. 9.1.11.33 in Appendix 18-9.1.A).

9.6.4 (d) Evaluation of the tutorial

The results also show that the tutorial, which accompanied the Route Simulator was considered helpful with a mean of 6.43. These findings are critical as they suggest that there were no problems with the tutorial and that there were no problems with the usability of the Route Navigator (Fig. 9.1.11.23 to 9.1.11.27 in Appendix 18-9.1.A).

9.6.5 (e) Quality of photographs

The results demonstrate that the quality of the photographs used in the simulator were below average (average ranging from 4.5 to 5.4). Of most concern, was the question concerning picture quality (fuzzy to clear) with a mean of 4.5 being recorded. These results suggest that the images require updating with greater clarity (Fig. 9.1.11.34, Fig. 9.1.11.35 in Appendix 18-9.1.A).

The use of the Quis in this study was somewhat limited as it was cumbersome to administer. Initially it was envisaged that after experiment in Study 9.1 the Quis would be completed by the

sample. This would have provided an interesting dimension to the study as each experiment could be compared and analysed by cognitive style. However this proved impractical but instead the students completed the Quis after completing all experiments in Study 9.1.

9.6.6 Discussion of the Quis as an Evaluatory Tool

However the questionnaire as a means of evaluating the usability of the questionnaire is limited. The questionnaire provided the users with the opportunity to indicate a n/a response. In the analysis these responses were categorized as being missing values. In many instances this reduced the sample number, which responded to individual questions on this questionnaire. When the questionnaire was re examined for users comments where the n/a response was availed of, no written comments were made which explained the basis of these decisions. Generally the QUIS, as an index of user satisfaction, was not found to be that effective in assessing the usability of an application such as this.

The Quis in its correct form was difficult to use in a meaningful form in the study. It proved to be too cumbersome to be administered at regular intervals during the course of a half-hour experiment. What is ideally required is a shortened computerised form of this questionnaire, which would facilitate rapid evaluation of interfaces in a short time interval.

There is some variation in the sample sizes that responded to the different questions in the QUIS. When an individual responded to a question by ticking the N/A option, these responses were coded as being zeros and were excluded from the analysis. The analysis also indicates that the participants in the experiment understood the instructions and found the application easy to use. In general the application was rigid and inflexible but this was to be expected given its experimental nature. There was general satisfaction expressed at clarity and font screen size and logical presentation of screens and instruction. Dissatisfaction was expressed about the clarity of some of the images. One of the disadvantages of the web based Route Simulator was the fact that some of the sample were able to log out and didn't complete all three routes.

The design of the Route Simulators could be improved upon; routes and junctions with less predictability might be considered on the next occasion. The still images need to be of greater clarity. It would have been very worthwhile to develop this application using a Geographical Information System as was initially envisaged.

9.7 Discussion of Matching and the V-I Dimension as Discussed by Riding and his Co-Workers

The only studies on the matching hypothesis and the V-I dimension appear to be those carried out by Riding and his co-workers over the last 25 years or so.

Studies by Riding and Ashmore (1980) found that Imagers learned better from pictorial presentation and Verbalisers from written material. Riding and Calvey concluded that Imagers did best on recall when material was highly descriptive as opposed to when the text was unfamiliar and acoustically complex. They reported that this was reversed in the case of the Verbaliser. Riding, Buckle, Thompson and Hagger (1989) suggest that "trainees continually taught with material and methods which do not match their cognitive style will therefore find difficulty in learning. In another study Riding and Douglas (1993) reported that the "test plus text condition favoured the Verbaliser and the text plus picture the Imager". Riding and Sadler-Smith (1992) suggest that "the implications for instruction is that teachers and trainers should be aware of differences in cognitive style and should attempt to accommodate these into their instructional programmes". Riding and Rayner (1995) recommend the design of a theoretical adaptive learning system that provides support for the Verbaliser and Imager by presenting material textually and pictorially respectively.

Riding and Sadler-Smith (1997) argue for the accommodation of cognitive styles in training and the design of training materials. They favour an adaptive approach and outline three ways that cognitive style could be accommodated. One of these, "the adaptive approach", is based on the matching concept for Verbalisers and Imagers. It might have been expected based on the literature above that the matching hypothesis would have favoured both Imager and Verbaliser in Studies 9.1 and 9.2.

The empirical studies above by Riding and his co-workers, and the conclusions based on them, are limited by virtue of the unreliability of the CSA (Parkinson et al (2004)).

Similarly, although the new test of the V-I dimension, VItrap, was shown to be more reliable than Riding's CSA (33% versus about about 20% at best) it would also have been expected to support the matching hypothesis.

These Studies, 9.1.and 9.2, are perhaps the first which have investigated cognitive style and the matching / mismatching hypothesis using this type of time-dependent application. The Route Navigator was a time dependent application similar to real time applications such as mission critical interfaces. The results of Study 9.1 have been replicated by those of Study 9.2 and can be considered to be reliable. These results raise some interesting issues: how should an interface be customised for Imagers and Verbalisers? Matching appears to suit Imagers and mismatching appears to suit Verbalisers. Another issue is whether context (e.g. learning in a classroom vs. traversing a route) is relevant in determining when mismatching or matching should be favoured? These results raise further questions about how Imagers and Verbalisers process images and textual material in time critical applications. Finally the definitions of Verbaliser and Imager need to be re-examined and perhaps deepened.

The results from the web application Study 9.3 were also supportive of these findings.

9.8 Limitations of these Studies

There was no randomisation of participants to each of the experiments in each of the three studies. However randomisation would entail a much larger sample size, greater expense and size for what was an exploratory study, and perhaps another experiment.

There was difficulty in designing a Route Simulator, with more complex decisions at each junction. This was probably because it accurately reflects the chosen terrain i.e. County Wicklow. There were also Health and Safety considerations in collecting data in more populous terrain e.g. County Wexford.

The VItrap, with a test-retest reliability of about 33% was a constraint.

9.9 Overall Conclusions

The overall conclusions for this Chapter may be grouped as follows:

Conclusion 1: In matched conditions there was a statistically significant group effect. There was a statistically significant difference betwen Verbaliser and Imager groups based on the use of t-tests (as post-hoc tests). However it is worth noting that matching favoured the Imagers (Imagers had the lowest mean relative to Verbalisers in Experiment 1 (most matched) and relative to the Imagers' performance in Experiment 2 (Most mismatched).

Conclusion 2: In mismatched conditions there was also a statistically significant group effect. There was a statistically significant difference between the Verbaliser and Imager groups based on the use of t-tests as post-hoc tests. In this instance these conditions favoured the Verbaliser with the Verbaliser having a lower mean than the Imager in Experiment 2 (most mismatched). The Verbaliser group also had a lower mean in Experiment 2 relative to the Verbaliser mean in Experiment 1.

Conclusion 3: In Experiments 3 and 4 (partially mismatched) there was no statistically significant difference between the Imager and Verbaliser groups.

Conclusion 4: These results were replicated with different and smaller sample groups (n = 40 and n = 20). The results from this study were constrained by the relative unreliability of the VItrap and the limitations in the design of the Route Simulator.

Conclusion 5: It appears from the results that Matching favours the Imager and Mismatching favours the Verbaliser. Note that these results are not what was expected. In the design both Verbalisers and Imagers were their matched conditions. It appears now that Imagers should have been given images (as they were) but Verbalisers should also have been given images (as they were not). Consequently it appears that images alone will satisfy Imagers (most matched) and Verbalisers (most

mismatched) in these time-dependent applications. This has obvious implications for Human-Computer Interface design for Verbalisers and Imagers. These comments must be interpreted with respect to moderate level of reliability of the VItrap.

Conclusion 6: The analysis from the Quis also indicates that the participants in the experiment understood the instructions and found the application easy to use. In general the application was rigid and inflexible but this was to be expected given its experimental nature. There was general satisfaction expressed at clarity and font screen size and logical presentation of screens and instruction. Some dissatisfaction was expressed about the clarity of some of the images.

Conclusion 7: Based on these prototypes, the results would suggest that it is sufficient to design for Imagers only in these types of navigational systems. This appears to be a strong conclusion which needs further investigation.

The applications designed as Route Simulators enabled construct validity of the VItrap to be tested and indirectly, the testing of the matching/ mismatching hypothesis. The design of these applications was limited by the learning effect and the fact that the majority of junctions in the application generally required that the individual progress straight through.

The applications examined in this thesis provided new data and insights about the matching /mismatching hypothesis where the dependent variable is time-dependent. This analysis was severely limited by learning effect of the route simulators and the golf stroke saver applications proved impractical to actually realistically evaluate.

In conclusion, the studies in this Chapter successfully demonstrated that individual differences such as cognitive styles could be accommodated by means of a test embedded in the interface. However the results obtained were surprising. Matched conditions favoured the Imager, while mismatched conditions favoured the Verbaliser. These results were supported by two other studies with smaller sample sizes and therefore can be considered reliable.

Chapter 10:
Conclusions

10.1 Purpose

The primary purpose of this thesis was to explore how the design of the human computer interface could be varied to accommodate cognitive style.

An analysis of the literature (Chapter 2) suggested that non-adaptive systems at the moment appear to be most appropriate way to accommodate cognitive style. Some initial studies undertaken also supported this view (Chapter 4 - Studies 4.2a, 4.2b and 4.3).

To examine the sensitivity of the design of the human computer interface to individual differences it was important to decide on a test of cognitive style to explore HCI design. Riding's computerised test of cognitive style appeared initially to be the most promising as a basis for designing the interface on the assumption that it was reliable. The other candidate possibilities considered from a large selection of tests were: Witkin's GEFT and Kirton's KAI (1976). Riding's CSA's advantages over those two tests were: a. ease of availability and scoring since it was computerised, b. it was also easy to administer (this was important particularly for large sample sizes) and c. it also provided immediate feedback to the testees as to what their cognitive styles were.

There was a large volume of publications in the literature which gave Riding's CSA an aura of validity, and consequently an inferred reliability that was discovered in the evolution of this thesis to be erroneous. It was also found during work on this thesis that CSA's reliability had not been reported upon previously in the literature. In pursuit of this thesis the reliability of Riding's CSA was examined, and found to be wanting (Redmond, Mullally, Parkinson (2002), (Parkinson, Mullally and Redmond (2004c)).

This discovery in Chapter 5, that Riding's test of cognitive style, (Cognitive Styles Analysis) was unreliable and therefore could not be used as a test of individual difference to be used to modify

the Human-Computer interface, precipitated first a search for another test and ultimately a decision to design a better test, which itself could be more easily tested and used in the human computer interface. This was a significant finding as it had not been demonstrated or reported upon previously in the literature. This finding has since been verified by Rezaei and Katz 2003) although their test-retest interval was quite short (Study 1: one week; Study 2: one month)).

Riding's Verbaliser - Imager dimension proved to be more unreliable than his Wholist-Analyst (W-A) dimension (Chapter 5). These findings raise serious concern about the validity of studies reported in the literature based on the CSA. They also point to the need for a new test of the Verbaliser-Imager dimension which should have a test-retest reliability of 70% or greater. Consequently a new test of the V-I dimension, VItrap, was developed in attempt to fill the gap.

The new test VItrap (discussed in Chapters 6 and 7), is more stable than Riding's CSA in evaluating the V-I dimension when compared with test-retest reliabilities over the same test-retest interval (4 weeks). It also represents a significant improvement on Riding's CSA. Its test-retest reliability (over a 4 week test-retest interval and a 6 week test-retest interval) of 0.33, while much better than Riding's V-I (of about -0.2 for 2 slightly shorter time intervals) is however a long way short of Kline's (1993) desired level of 70%. However it does point in the right direction of a test which is not too long but which is capable of improvement. It allows for the possibility of assessing the V-I dimension using test items in both image and word forms. It also enables the individual doing the test, to dispense with using stickers on two keys, and to use the mouse instead. It provides information relating to each question in a data file unlike Riding's CSA.

This study of VItrap also demonstrated that response time to a test item is significantly affected by a) whether the previous question was answered correctly or incorrectly, and b) the category of test item currently being responded to. The preliminary results suggest that dropping feedback as to correctness or otherwise might improve test-retest reliability. The

findings of both a) and b) were shown to be replicable since they were demonstrated in the retest.

This new test of the V-I dimension - VItrap - is also independent of Intelligence as assessed by Ravens' Matrices and the Millhill Inventory. VItrap was also found to be independent of personality as assessed by Eysenck's EPI and also Spielberger's STAI.

The new test of the V-I dimension was found not to correlate with Smith's Learning Preference Questionnaire. It can be argued that the VItrap consequently does not predict observable behaviour where Smith's learning preference inventory is used as a criterion. Smith had previously reported a significant p-value using an ANOVA model where Riding's CSA and his learning Preference Inventory were components in the ANOVA model. Perhaps these results are questionable given the unreliability of the CSA in the first place. It is interesting to note that Peterson (2003) had also previously found little correlation between Smith's Questionnaire and her VICS test.

Similarly the VVIQ did not correlate with VItrap. There may have been a few reasons for this: a) low test-retest correlation values for VVIQ indicating low reliability and b) the fact that the two tests were somewhat different in design and purpose anyway. This demonstrated the difficulty of establishing the convergent validity of VItrap, as other reliable measures of the V-I dimension are not readily available.

The results of the neurophysiological investigation (Chapter 8) did not strongly support the theory of hemispheric specialisation, the approach used by Riding et al. in their attempt at neuropsyiological evaluation of the CSA. However it must be acknowledged that this study was limited by sample size, type and experimental design but was still a useful exploratory study. The findings from Study 8.1 may not necessarily imply that the new test of the V-I dimension has not got an underlying neurophysiological basis. The EEG approach appears not to be sensitive enough to pick up any interactions of interest, if any. However the Riding, Glass, Butler and Pleydell-Pearce (1997) approach has not proved useful. Consequently, the neurophysiological basis of the VItrap has not been tested and is

still an open question. This also raises the question as to whether cognitive style, from a neurophysiological basis, fails to exist at all.

The prototype applications (Studies 9.1 - 9.3), in Chapter 9 which were developed to scrutinize the feasibility of facilitating cognitive style within the design of the human computer interface, proved advantageous. These applications were exploratory prototypes and were meant to elucidate difficulties on the way to proof of concept. The route simulators proved to be the most successful applications and also enabled the testing of the matching-mismatching hypothesis. Based on these prototypes, the results would suggest that it is sufficient to design for Imagers only in these types of navigational systems. This appears to be a strong conclusion which needs further investigation.

These prototype applications are perhaps the first of their kind to incorporate a test of cognitive style as a component part of the interface. While somewhat constrained, they served to illustrate the potential of similar type applications. These types of applications demonstrated the potential and possibility of the delivery of customised information via the Web and Personal Digital Assistant devices.

The basis of these prototypes of individual difference application to HCI was predicated on the Matching Hypothesis. The hypothesis has its advocates and opponents as already discussed in Chapter 4. This hypothesis has not been satisfactorily validated by the Route Simulator studies due to unforeseen issues. It was found that Matching favours the Imager and Mismatching favours the Verbaliser for the sample used (technical universtity students in their late twenties).

Experimentation such as evaluating individual differences using psychometric tests is a difficult task. The possibility for something unpredictable happening with serious ramifications for the design of the experiment is always high.

The QUIS self-reporting questionnaire proved useful in assessing the usability of the prototype applications. Generally the application, and accompanying tutorial, was found to be easy to

use, but the quality of some of the pictures needed improvement. The WAMMI, an alternative means of evaluating the usability of an interface was found not to be too applicable with these applications (Wammi (2006)).

10.2 Limitations of this Thesis

The applications in this thesis were evaluated with a sample of third and fourth year technical university students.

VItrap test-retest reliability, while a big improvement on Riding's V - I, still needs to be improved further.

There were small, but very typical of the EEG area, sample sizes in the EEG studies.

Also of necessity, many of the studies were done in parallel and not consecutively.

10.3 Contributions to Knowledge

10.3.1 Major Contributions

The unreliability of Riding's CSA: This was demonstrated for the first time in the literature. Similarly validity problems were explored at the same time. When the unreliability of the CSA was shown, it became necessary to design and implement a new test of the V-I dimension. Consequently, the reliability and validity of the VItrap were subsequently investigated.

Design, Construction and Validation of the VItrap: A new test, VItrap, was designed and proved to be more reliable than Riding's V-I dimension over similar test retest intervals.

The test-retest reliability value of the VItrap needs to be significantly increased before the conclusions pertaining to its validity can definitely be held to hold and be valid.

Cross-fertilisation of the two fields of Cognitive Style and Systems Design: The human computer interface can be modified using an integrated test of cognitive style to improve the interface so that individual differences can be accommodated.

The Matching Hypothesis: This thesis demonstrated that, when the Imager is accommodated in matched conditions the Imager's performance is better than that of the Verbaliser in matched conditions.

An unexpected finding is that the Verbaliser's performance is better in mismatched conditions when compared with the Imager. In mixed conditions there is little difference between Verbalisers and Imagers.

In general, designing for the Imager appears to suit both Imagers and Verbalisers in navigational applications

10.3.2 Minor Contributions

Hemispheric Specialisation was Refuted: It was clearly demonstrated that hemispheric specialisation as reported by Riding et al. was not replicable in a similar study of the V-I dimension using VItrap. This is in agreement with recent work reported in the literature on hemispheic specialisation

Design & Implementation of Prototypes: The two fully developed prototypes incorporated cognitive styles for user accommodation. The test of cognitive style had been neurologically investigated using EEG.

10.4 Future Directions

The shortcomings of the new test of the V-I dimension - VItrap - need to be addressed and remedied. Perhaps more novel ways of assessing the Verbaliser-Imager dimension, relying more heavily on neurophysiological approaches, should be explored.
A better test of the V-I dimension can be used more effectively to modify the human-computer interface.

References

Adams, V.M., & McLeod, D.B. (1979) The interaction of field independence with discovery learning in mathematics *Journal of Experimental Education*, 48, 32-35.

Adey, P., Fairbrother, R, and Wilian, D., with Johnson, B., and Jones C, (1999) *Learning styles and Strategies*, OFSTED, Review of research A report commissioned by the office for standards in education , King's College London, School of Education.

Allinson, C. and Hayes, J. (1996) The Cognitive Style Index, *Journal of Management Studies*, 33, 119 - 135.

Allport, G.W. (1937) *Personality: A Psychological Interpretation*. New York Holt and Co.

Altman, D (1991) *Practical Statistics for Medical Research*, Chapman & Hall.

Anderson, P. & Anderson, S.A. (1968) *Physiological Basis of the Alpha Rhythm*, (New York, Appleton-Century-Crofts).

Antonietti, A. and Giorgetti, M. (1998) The Verbalizer-Visualizer Questionnaire: a review, *Perceptual and Motor Skills* Vol. 86 (1) pp 227-240.

Armstrong S.J. (1999) *Cognitive Style and Dyadic Interactions A Study of Supervisors and Subordinates in Working Relationships*, Ph.D. Thesis, University of Leeds December.

Armstrong, S (2000) Individual differences and cognitive style and their potental effects on organisational behaviour: a summary of recent empirical studies In: R. J. Riding & S. G. Rayner, (Eds.), *Cognitive style: international perspectives on individual differences, Cognitive styles (Vol. 1)*, (pp. 315–347). Stamford, Connecticut: Ablex Publishing Corporation

Armstrong, S. (2005) Personal communication, Jan 2005.

Ash, F. (1986) Identifying learning styles and matching strategies for teaching and learning, *Eric document ED270142*.

Atkinson, G. (1988) Reliability of the Learning Style Inventory – 1985. *Psychological Reports*, 62, 755-758

Authorware V6.5 (2002) Macromedia 2002

Autocad (2006)
http://www.axonware.ie/catalog/product_info.php?pName=autocad-lt-2006

Ayersman, D.J. & Minden, A. (1995) Individual differences, computers, and instruction. *Computers in Human Behavior*, 11 (3/4), 371-390.

Banich, M.T. (2004) *Cognitive Neuroscience and Neuropsychology* 2nd Ed. Houghton Mifflin, USA.

Bariff, M.L. and Lusk, E.G. (1977) Cognitive and Personality Tests For The Design of Management Information Systems, *Management Science*, Vol 23, No 8 pp 820 – 829.

Benbasat, I and Taylor, R. N. (1978) The Impact of Cognitive Styles on Information System Design , *MIS Quart.*, Vol 2 pp 43-54

Benjafield, G.G. (1997) Vividness of Visual Imagery: measurement, nature, function and dynamics by S.J. McKelvie *Canadian Journal of Experimental Psychology* Book Review

Benyon, D, R, and Murray, D.M., (1993). Applying user modelling to human computer interaction design. *Artificial Intelligence Review* 6, pp43-69.

Benyon, D, McCrerar A, Wilkinson, S (2001) Individual differences and Inclusive design pp27-46 in *User Interfaces for all , Concepts Methods and Tools*, Edited by Constantine Stephanidis , publishers Lawrence Erlbaum Associates, Inc NJ. ISBN – 0-8058-2967-9

Berger, H. (1932) Über das Electroneenkephalgram des Menschen, *Vierte Nitteilany*. [Translated by G. P. Gloor (1969)

Electroencephalography and Clinical Neurophysiology, Supplement 28, pp. 133–150].

Berger, H. (2006)
http://www.uic.edu/depts/mcne/founders/page0008.html

Betts, G.H. (1909) The distribution and functions of mental imagery. New York: Teachers College, Columbia University referred to in Chara & Verplank

Bieri, J. (1955) Cognitive complexity-simplicity and predictive behavior, *Journal of Abnormal and Social Psychology*, 51, 263 - 268.

Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M. and Prieto, T. (1997) Human brain language areas identified by functional magnetic resonance imaging, *Journal of Neuroscience*, 17, 1, 353-362.

Bisiach and Berti (1990) Waking images and neural activity, in: R. G. Kunzendorf & A. A. Sheikh (Eds.) *The Psychophysiology of Mental Imagery: theory, research and application* (Amityville, NY, Baywood Publishing Company).

Block, J., Block, J.H. and Harrington, D. (1974) Some misgivings about the Matching Familiar Figures Test as a measure of reflection - impulsivity, *Development Psychology*, 10, 611-632.

Block, J., Gjerde, P. and Block H. (1986) More misgivings about the Matching Familiar Figures Test as a measure of reflection - impulsivity: Absence of construct validity in pre-adolescence, *Development Psychology*, 22, (6), November 1986, pp. 820-831, American Psychological Association.

Bloom, B. (1976) *Human Characteristics and Learning* New York McGraw-Hill

Bogen, J.E. (1969) The other side of the brain II. An appositional mind, *Bulletin Los Angeles Neurological Society*, 34, pp. 135–162.

Boles, W.W. and Pillay, H. (1999) Cognitive Styles, Subject content and the design of computer based instruction. Published in the

Proceedings of the fifth International Symposium on Signal processing and its applications, ISSPA 1999,Brisbane Australia , 22-25 August 1999.

Bostrom, R. P. Olfman, l., Sein, M. K. (1988) The importance of individual differences in end user training: the case for learning style. ACM -0-89791-262-4/88/0400-0133.

Bottini, G., Corcoran, R., Sterizi, R., Paulesu, E., Schenone , P., Scarpa, P., Frackowiak, R. and Frith, D. (1994) The Role of the right hemisphere in the interpretation of figurative aspects of languag : A positron emission tomography activation study, *Brain*, 117 1241-1253.

Bradshaw, J. and Rogers, L. (1993) The evolution of lateral asymmetries, language, tool use, and intellect, San Diego, CA: Academic Press, 1993. 463pp.

Brannigan, G.G., Ash, T. and Margolis, H. (1980)) Impulsivity, reflectivity and children's intellectual performance. *Journal of Personality Assessment*, 44, 41-43.

Briggs-Myers, I. (1990). A description of the theory and applications of the Myers-Briggs Type Indicator, California: Consulting Psychologists Press.

Bruner, J.S. (1962). On Knowing: Essays for the left hand. P87 Cambridge, Mass; *Harvard University Press.*

Brusilovsky, P. (1996) Methods and Techniques of Adaptive Hypermedia, *Journal of user modeling and user adapted interaction*. 6, 2-3 pp 87-129

Bunt, A., Conati, C., McGrenere, J., (2004) What role can Adaptive Support Play in an Adaptable System? IUI 2004 January 13- 16, Madeira, Funchal, Portugal.

Bush, V. (1945). As we may think. *Atlantic Monthly*, 76, 10 –108.

Campbell, D.T. and Fiske, D.W. (1959)) Convergent and discriminant validation by the multitrait - multimethod matrix. *Psychological Bulletin*, 56, pp. 81-105.

Carroll, J.B. (1993) *Human cognitive abilities: a survey of factor-analytic studies*, Cambridge University Press.

Carroll, J.N. (2002) *Human - Computer Interaction in the New Millenium*, ACM Press, New York, Addison - Wesley.
Cassidy, Simon (2004) Learning Styles: An overview of theories, models and measures. *Educational Psychology* Vol 24 (4), pp 419-444.

Cattell, R. B. (1971) *Abilities: Their Structure, Growth and Action*, Houghton Mifflin, New York.

Cattell, R.B. (1982) *The Inheritance of Personality and Ability*, Academic Press, London.

Chara, Paul J., Jr. and Verplanck, W.S. (1986) The imagery questionnaire: an investigation of its validity, *Perceptual and Motor Skills, 63*, 915-920.

Clapp, R.G. (1993) Stability of cognitive style in adults and some implications: A longitudinal study of Kirton's Adaption - Innovator Inventory, *Psychological Reports* 73, (1234-1245).

Chen, S.Y, (2002) A cognitive model for non-linear learning in hypermedia programmes, *British Journal of Educational Technology* 33(4), 449-460.

Chen, S.Y. & Macredie, R. D, (2002) Cognitive styles and hypermedia navigation: development of a learning model *Journal of the American Society for Information Science and Technology* 53(1), 3-15.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988) Development of an instrument measuring user satisfaction of the human-computer interface. *Proceedings of CHI 88*, 213-218. New York, NY: ACM. http://www.lap.umd.edu/LAPFolder/papers/cdn.html.

Chou, H.,W. (2001) Influences of cognitive style and training method on training effectiveness. *Computers and Education* Vol 37, Issue 1 pp 11-25.

Churchman, C.W. & Schanblatt (1965) The researcher and the manager: A dialectic of implementation", *Management Science* 11(4) pp69- 87.

Clapp, R.G (1993) The stability of cognitive dtyles in adults :a longitudinal study of the KAI, *Psychological Reports* 73,1235-45.

Clark, R.E., (1985) Confounding in educational computing research. *Journal of Educational Computing Research* 1, pp 137- 147.

Coffield F, Moseley D, Hall E, Ecclestone K, (2004) *Learning Styles and Pedagogy in post -16 learning. A systematic and critical review. (Report B) (The Coffield Report)* Published by the Learning and Skills Research Centre. UK.

Cohen, G. (1982) Theoretical interpretations of lateral asymmetries, in: J. Graham Beaumont (Ed.) *Divided Visual Field Studies of Cerebral Organisation*, pp. 87–111. (London, Academic Press).

Cohen, J., (1977) *Statistical power analysis for the behavioral sciences* , Academic Press, New York (revised edition).

Cohen, J., 1988. *Statistical power analysis for the behavioral sciences.* (2nd ed.), Lawrence Erlbaum & Associates Hillside, New Jersey.

Cohen, B.H. and Saslona, M. (1990)) The Advantage of Being a Habitual Visualizer. *Journal of Mental Imagery* (Fall/Winter) 101-112.

Collins, R.L. (1985) On the inheritance of direction and degree of asymmetry, In *Cerebral Kateralisation in Non-Human Species*, ed. S.D. Glick, 41-72, New York: Academic Press.

Coop, A.H, Sigel C (1977) Cognitive style : Implications for learning and instruction, *Psychology in Schools*, 2, pp. 152-161.

Cooper, C. (2002) *Individual Differences* 2nd. Ed. Arnold, Hodder Headline Group, co-published in the US by Oxford University Press NY ISBN 0340 808160

Court, J.H. and Raven, J. (1975?) *Normative, Reliability and Validity Studies - Raven Manual Section 7: Manual for Raven's Progressive Matrices and Vocabulary Scales: Studies of reliability and validity of RPM*

(no date given - publication date estimated from references) published by Oxford Psychologists Press.

Curry, L. (1983) An organisation of learning styles theory and constructs, *Eric document 235185*.

Curry, L. (1990) A critique of the research on learning styles, *Educational Leadership*, 48 (2), 50 - 56.

Das, B. (1988) Simultaneous successive and planning implications for school learning in Schmeck, R.R. (ed.) *Learning Strategies and Learning Styles* New York, N.Y. U.S.A. Plenum Press

Data Desk V6 (2006) (Computer software), (1980–2006) Ithaca NY 14852: Data Description Inc., P.O. Box 4555 and http://www.kovcomp.co.uk/DataDesk/ and http://www.geoafrica.co.za/reddog/datadesk/datadesk.htm

De Bra, P., Brusilovky, P. and Houben, G.J. (1999). Adaptive hypermedia: from systems to framework. In *ACM Computing Surveys*, December 1999.

De Renzi, E., Faglioni, P. and Sorgato, P. (1982) Modality specific and supra modal mechanisms of apraxia, *Brain*, 105, 301-312.

Desmedt, E. and Valke, M. (2004) Mapping the Cognitive Styles "Jungle": an overview of the literature based on citation analysis, *Educational Psychology* Vol 24 (4) pp 445-464.

Douglas, G., Riding, R.J. (1993) The effect of pupil cognitive style and position of prose passage title on recall. *Educational Psychology* **13**, 385 - 393.

Driver and Mock (1975) in Huber, G. P. (1983) Cognitive Style as a basis for MIS and DSS designs : much ado about nothing. *Management Science* vol 29, No 5.

Dunn, R., Dunn, K. and Price, G.E. (1989) *Learning Styles Inventory*, Lawrence, KS. Price Systems.

Dvorak, R & Sommerville, S.T. (1996) Hypertext/Hypermedia - a review. *QMW Technical Report*.

Egan, D.E., Gomez, L, M (1985) Assaying, isolating and accommodating individual in learning a complex skill. In R. Dillon, *Individual Differences in Cognition* (pp 173- 217). New York Academic Press.

Ehrlichman, H. and Barrett, J. (1983) Right Hemispheric Specialization for Mental Imagery: A Review of the Evidence, *Brain & Cognition* 2, 55-76

Elkstrom, R.B., French, J.W., Harman, H.H. and Dermen, D. (1976). *Manual for kit of factor referenced cognitive tests*, New Jersey: Educational testing service.

Entwhistle, N.J. (1981) *Styles of learning and teaching*. Chichester: Wiley

Eysenck, H.J. and Eysenck, S.G.B. (1975) *The Eysenck Personality Questionnaire*, Hodder & Stoughton, Seven Oaks, U.K.

Eysenck, H.J. and Eysenck, S.G.B. (1991) *Eysenck Personality Scales*, London, Hodder & Stoughton, U.K.

Farah, M.J. (1984) The neurological basis of mental Imagery: A componential analysis, *Cognition*, 18 pp 245- 272.

Farah, M.J. (1986) The laterality of mental image generation: a test with normal subjects, *Neuropsysiogia*, 24, pp. 541-55.

Felder, R.M. and Silverman, L. K, (1988) "Learning and teaching styles in Engineering Education", *Engineering Education*, 78, 7 pp. 674-681 (1988).

Felder, R.M. and Soloman, B.A. (2005) Index of Learning Styles Questionnaire http:www.ncsu.edu/felder-public/ilspage.html

Felder, R.H. and Spurlin, J. (2005) Applications, reliability and validity of the Index of Learning Styles, *International Journal of Engineering education*, Vol. 21, pp. 103, 112.

Findlater, L. and McGrenere, J. (2004). A comparison of Static, Adaptive and Adaptable Menus. *CHI 2004* April 24-29 Vienna, Austria.

Fischer, G. (2001) "User Modeling in Human-Computer Interaction", Contribution to the 10th Anniversary Issue of the Journal "*User Modeling and User-Adapted Interaction (UMUAI)*" Vol. 11, No. 1/2, pp 65-86, 2001
[http://www.cs.colorado.edu/~gerhard/papers/umuai2000.pdf]

Ford, N. (1985) Learning styles and strategies of postgraduate students, *British Journal of Educational Technology*, 16, 65-79.

Ford, N. (1995) Levels and types of mediation in instructional systems: an individual diferences approach, *International Journal of Human Computer Studies*, 43, 241-259.

Ford, N. & Chen, S (2001) Matching/ mismatching revisited: an empirical study of learning and teaching styles, *British Journal of Educational Technology* 32, 1 5-22.

Frank, B. & Davis, J. (1982) Effective field independence match or mismatch on a communication task, *Journal of Educational Psychology* 74, 23-31

Fraisse, P. (1968) Motor and verbal reaction times to words and drawings *Psychon. Sci.* Vol. 12 (6) pp. 235-236.

Furlong, N.E., Lovelace, E.E. and Lovelace, K.L., (2000) *Research Methods and Statistics – an Integrated Approach*, Harcourt, Brace Fort Worth.

Fürst, C.J. (1976) EEG alpha asymmetry and visuospatial performance. *Nature* 260, 254 - 255.

Galin, D. and Ornstein, R (1972) Lateral Specialization of Cognitive Mode: An EEG Study, *Psychophysiology*, Vol.9 No.4 pp 412- 418.

Gardner, R.W., Holzman, P.S., Klein, G.S., Linton, H.B. & Spence, D.P. (1959) Cognitive control: a study of individual consistencies in cognitive behavior. *Psychological Issues*, New York: International Universities Press.

Gazzaniga, M. S., Ivry, R.G and Mangum, G.R (2002) *Cognitive Neuroscience: The Biology of the Mind*, Second Edition. Publisher W.W. Norton and Company N.Y U.S.A.

Gevins, A., Smith, Michael. E, McEvoy Linda, Leong, H. and Le, Jian (1999) Electroencephalographic imaging of higher brain function. *Phil. Trans. Royal Society*, 354, pp. 1125-1134.

Glass, A. (1964) Mental arithmetic and the blocking of the occipital alpha rhythm, *Electroencephalography and Clinical Neurophysiology* 16, 595-603.

Glass, A. and Riding, R.J. (1999) EEG differences and cognitive style, *Biological Psychology*, Vol. 51, Issue (1), pp. 23-41.

Goldstein, K.H and Blackman. S. (1978) *Cognitive Style, five approaches and relevant research*, John Wiley and Sons, New York.

Gordon, R (1949) An investigation into some of the factors that favour the formation of stereotyped images, *British Journal of Psychology* 39, 156-167.

Gregorc, A.R. (1982) *Style Delineator*, Maynard, Ma. Gabriel Systems

Grigorenko, E.L. and Sternberg, R.J. (1995) Thinking styles in Saklofske, D.H. and Zeidner, M (Eds.) *International Handbook of Personality and Intelligence*, 205 - 230, New York, Plenum Press.

Grundnik, J.L. and Krantzler, J.H. (2001) Meta-analysis of the relationship betwen intelligence and inspection time, *Intelligence*, 29. pp. 523-535.

Guilford, J.P. (1967) *The Nature of Human Intelligence*, McGraw-Hill New York.

Guilford (1980) Cognitive styles: What are they?, *Educational and Psychological Measurement* pp 715-735.

Guthrie, G.D. (1967) *Changes in cognitive functioning under stress: a study of plasticity in cognitive controls*, Doctoral dissertation, Clark University, Worcester, MA., Dissertation Abstracts International 28, 212-513.

Halpin, G. & Peterson, H. (1986) Accommodating instruction to learners' field independence/ dependence; a study of effects on achievements and attitudes, *Perceptual and Motor Skills*, 62, pp. 967-974.

Hansen, R. (1983) *The effects of varied text conditions on field independent and field dependent readers' formation of inferences in expository text*, Dissertation Abstracts International, 44, 2722.

Harvey, O.J. (1963) Authoritarianism and conceptual functioning in varied conditions, *Journal of Personality*, 31, 462 - 470 cited in Goldstein & Blackman

Hashway, R.M. and Duke, I.L, (1992) *Cognitive Styles A primer to the literature*, Mellen Research University Press, San Francisco California.

Hatta, T. and Koike, M. (1991) Left-hand preference in frightened mother monkeys in taking up their babies, *Neuropsychologia*, 29(2): 207-209.

Hayes, J. and Allinson, C. W. (1994) Cognitive style and its relevance for management practice, *British Journal of Management* 5, 53-71.

Hayes, J. and Allinson, C. (1996) The implications of learning styles for training and development: a discussion of the matching hypothesis, *British Journal of Management* 1996, Vol 7, Pages 63-73.

Hayes, J. and Allinson, C.W. (1998) Cognitive style and the theory and practice of individual and collective learning in organizations *Human Relations* 51, pp 847-871.

Hayes, J., Allinson, C.W., Christopher, W., Hudson, R. and Keasey, K. (2003) Further reflections on the nature of intuition / analysis and the construct validity of the Cognitive Style Index, *Journal of Occupational and Organisational Psychology* June, Vol 76 (i) 2, p269 - 279.

Head, J.O. (1996) Gender identity and cognitive style in Murphy, P.F. & Gipps, C.V. (Eds.) *Equity in the classroom: Towards effective pedagogy for boys and girls*, London Falmer

Hellige, J. B. (1990) Hemispheric Asymmetry, *Annu. Rev. Psychol.* (41) pp 55-80.

Hellige, Joseph. B. (2001) Perspectives in *Cognitive Neuroscience: Hemispheric Asymmetry, What's Right and What's Left*, Harvard University Press. ISBN 0-674-3870-9.

Hermann, N. (1988) *The creative brain* (fifth edition). Lake Lure ; Publishers NC; Brain books

Hiscock, M. and Kinsbourne, M. (1995) Phylogemy and antogeny of cerebral lateralisation, In R.J. Davidson and K. Hugdahl (Eds.) *Brain Asymmetry*, Cambridge, MA: MIT Press.

Hodgkinson, G. and Sadler-Smith, E. (2003) Complex or unitary? A critique and empirical re-assessment of the Allinson - Hayes Cognitive Style Index, *Journal of Occupational Psychology* 76, pp. 243 - 268.

Holzman, P.S. and Klein, G.S. (1954) Cognitive system principles of levelling and sharpening individual differences in visual time error assimilation effects, *Journal of Psychology*, 37, 105-122.

Honey, P. and Mumford, A., 1992) *The Manual of Learning Styles*, Maidenhead: Peter Honey Publications.

Hook, K (1997) Evaluating the Utility and Usability of an Adaptive Hypermedia System, *Proceedings of the 2nd International Conference on Intelligent User Interfaces, IUI 97* Orlando Florida USA ACM 0-8979-839-8/96/01

Hook, Kristina (1998a): Designing and Evaluating Intelligent User Interfaces. In: Marks, Joe (ed.): *International Conference on Intelligent User Interfaces 1998*. January 6-9, 1998, San Francisco, California, USA. p. 5-6.

Hook, K. (1998b) Evaluating the Utility and Usability of an Adaptive Hypermedia System, in *Journal of Knowledge Based Systems*, Vol 10,(5).

Hook, Kristina (1999): Designing and Evaluating Intelligent User Interfaces. In: Maybury, Mark T. (ed.): *International Conference on Intelligent User Interfaces 1999*. January 5-8, 1999, Redondo Beach, California, USA. p.5-6.

Hook, Kristina (2000): *Steps to Take Before Intelligent User Interfaces Become Real*. In *Interacting with Computers*, 12 (4) p. 409-426

Huber, G. P. (1983) Cognitive Style as a basis for MIS and DSS designs : much ado about nothing. *Management Science* vol 29, No 5

Hudson, L. (1966) *Contrary Imaginations*, London Methuen

Hudson, L. (1968) *Frames of the mind*, Harmondsworth, Penguin Books.

Hunt, R.G., Krzystofiak, F.J., Meindl, J.R and Yousry, R.M (1989) Cognitive style and Decision making. *Organisational Behaviour and Human Decision Processes* 44, 463-453.

Irwin, R.J. (1984) Inspection and its relation to intelligence, *Intelligence*, vol 8, issue (1), 1984, pp. 47-65.

Isaksen, S.G., Lauer, K.J. and Wilson, G.V. (2003) An examination of the relationship between personality type and cognitive style, *Creativity Research Journal*, Vol. 15, No. 4, pp. 343-354.

Jacky, J. (1997) *The Way of Z: Practical Programming with Formal Methods* Cambridge University Press, 1997, ISBN 0521559766.

Jameson, A. (2000) Adaptive interfaces and agents in Jacko, J. and Sears, A. (eds.), *Handbook of HCI*. Mahway NJ USA: Erlbaum.

Jennings, F., Benyon, D. & Murray, D (1991) Adapting systems to differences between individuals, *Acta Psychologica* 78 (1-3) Dec 1991 243-256.

Jennings. M. and Woolridge, N. (1995) Intelligent agents: theory and practice, *Knowledge Engineering Review*, 10 (2), June 13.

Jonassen, D.M. (1989) Hypertext and hypermedia. Englewood Cliffs, N.J.: *Educational Technology Publications.*
Jonassen, D. H. and Grabowski B.L. (1993) *Handbook of Individual Differences, learning and Instruction*, Lawrence Erlbaum Associates Publishers Hillsdale, New Jersey, Hove and London.

Jones, A.E. (1997). Reflection-Impulsivity and Wholistic-Analytic: two fledglings? ... or is R-I a cuckoo? *Educational Psychology*, 17, 65-77.

John, D. and Boucouvalas, A. (1999a) Comparing user performance with interfaces designed to suit their cognitive style. *IEE European*

Workshop on Distributed Imaging, November 18-19, 1999, Savoy Place, London, pp. 5/1-5/6. Place, London.

John, D. and Boucouvalas, A. (1999b) Comparing user cognitive style and performance with text and image user interfaces, *IEE European Workshop on Distributed Imaging*, November 18-19, 1999, pp. 4/1-4/6. Savoy Place, London.

Joust (1999) The Outline Navigation System in Javascript www.ivanpeters.com

Kagan, J., Rosman, B, Day, D., Albert, J. and Philips, W. (1964) Information processing and the child; the significance of analytic and reflective attitudes, *Psychological Monographs* 76, 578

Kagan, J. and Buriel, R. (1977) Field dependence-independence and Mexican American culture and education. In J.L. Martinez, Jr., Ed. *Chicano Psychology* New York Academic Press.

Kaufmann, G. (1989) *The Assimulator Explorer Inventory in Martinsen, O., Cognitive Style and Insight*, Ph.D thesis, Faculty of Psychology, University of Bergen, Norway

Keeble R.J and Macredie R.D. (2000) Assistant agents for the World Wide Web intelligent interface design challenges. *Interacting with computers* – 12(4) pp357- 381.

Kilmann, K. H and Mitroff I.I (1976) Qualitative versus quantitative analysis for management science: Different forms for different psychological types. *Interfaces*, 6, 17-27.

Kim, K. S. (2001) Implications of user characteristics in information seeking on the World Wide Web. *International Journal of Human Computer Interaction*, 13, (3) 323-340.

Kirby, P. (1979) *Cognitive styles, learning style and transfer skill acquisition Information Services no. 195*, Columbus Ohio State University National Center for Research in Vocational Education

Kirton, M.J. (1976) Adaptors and Innovators: a description and measure, *Journal of Applied Psychology* 61, 622-9.

Kirton, M (1980) Adaptors and innovators in organisations, *Human Relations*, Vol. 33. No. 4, 213-224.

Kirton, M.J. (1987) *Adaption Innovation Inventory manual* 2nd Ed. Hatfield Herts. Occupational Research Centre.

Kirton, M.J. and McCarthy, R.M. (1988). Cognitive climate and organizations, *Journal of Occupational Psychology*, 61, 175-184.

Kirton, M.J. (Ed) (1989) *Adaptors and Innovators: styles of creativity and problem solving* (London Routledge).

Kirton, M.J., Bailey, A. and Glendinning, W. (1991) Adaptors and Innovators Preference for Educational Procedures, *The Journal of Psychology* 25 (4) pp 445-455.

Kirton, M.J (2003) *Adaption - Innovation In the Context of Diversity and Change* (London, Routledge).

Kline, P. (1993) *The handbook of psychological testing*, Routledge, London.

Kline, P., (2000). *The handbook of psychological testing*, 2nd. Ed., Routledge, London.

Kohli, W., (1995) *Critical conversations in philosophy of education.*, Miller Publishing Tourtledge, NY.

Kolb, D.A. (1976) *Learning Style Inventory: Technical Manual.* Englewood Cliffs, New Jersey: Prentice Hall.

Korthauer, R.D. & Koubek , R.J. (1994) An empirical evaluation of knowledge, cognitive style, and structure upon the performance of hypertext task. *International Journal of Human-Computer Interaction*, 6 (4), 373-390.

Kosslyn, S. (1987) Seeing and Imagining in the Cerebral Hemispheres: A Computational Approach. *Psychological Review*, Vol 94, No 2 , pp148- 175.

Kosslyn, Stephen, M. (1988) Aspects of cognitive neuroscience of mental Imagery, *Science* , 240, pp 1621-1626.

Kosslyn, S.M., Behrmann, M and Jeannerod, M. (1995a) The cognitive neuroscience of mental imagery. *Neuropsychologia*, 33 (1995) pp. 1335-1344.

Kosslyn, S.M., Maljkovic, V., Hamilton, S.E., Horowitz, G. and Thompson, W.L. (1995b) Two types of image generation: Evidence for left and right hemisphere processes, *Neuropsychologia*, 33 (1995) pp. 1485-1510.

Kozma R.B., (1991) Learning with media, *Review of Educational Research* 61 pp 179-221.

Krantzler, J.H. and Jensen, A.R. (1989) Inspection time and intelligence: a meta-analysis, *Intelligence*, 13, pp. 329-347.

Lajoie, S.P. and Shore, B.M (1987) Impulsivity, reflectivity and High IQ. *Gifted Education International* 4(3) 139-141.

Langhinrichsen, J. and Tucker, D.M. (1990) Neuropsychological concepts of mood, imagery, and performance, in: R. G. Kunzendorf & A. A. Sheikh (Eds) *The Psychophysiology of Mental Imagery: theory, research and application*. (Amityville, NY, Baywood Publishing Company).

Lanier, J. (1996) My problem with agents. http://www.wired.com/wired/archive/4.11/myprob.html and http://www.advanced.org/jaron/agentideeforte.html

Langley, P. (1998) User Modeling in adaptive interfaces. In *Proceedings of the Seventh International Conference on User Modelling*, pp 357- 370. Banff, Alberta.

Leonard, D. and Strauss, S. (1997) Putting your company's whole brain to work, *Harvard Business Review*, 75, 111 - 121.

Livesay, G.A., Dee, K.C., Nauman and Hites, L.S. (2002) Jr, Engineering student learning styles, a statistical analysis using Felder's Index of Learning Styles, presented at the 2002 Annual Conference of the American Society for Engineering Education, Montreal, Quebec.

Lohman, D.F. (1999) Minding our p's and q's: On finding relationships between leaning and intelligence. In P.L. Ackerman, P.C. Kyllonen & R.D. Roberts (Eds.), *Learning and individual differences: Process trait and content determinants.* Washington D.C.: American Psychological Association.

Liu, M. & Reed, M.R. (1994) The relationship between the learning strategies and learning styles in a hypermedia environment. *Computers in Human Behavior,* 1(4), 419-434.

Lynch, R., (1991) Teaching in the 21ST century. *Journal of Educational Psychology* 1, pp 144- 186.

McDonnell, R. and Kemp, K. (1995) *International GIS Dictionary,* Geo-Information International, Cambridge UK

McKay, E (1999) Exploring the effects of graphical metaphors on the performance of learning computer programming concepts in adult learners: a pilot study, *Educational Psychology* Vol 19 Issue 4.

McKelvie S.J. (1990) The Vividness of the Visual Imagery Questionnaire: commentary on the Marks – Chara debate. *Perceptual and Mental Skills,* 70, 551-560.

McKelvie S.J. (1994) The Vividness of the Visual Imagery Questionnaire as a predictor of facial recognition memory performance *British Journal of Psychology* 85 (Pt 1) 93-104 Feb.

McKelvie S.J. (1995) The VVIQ as a psychometric test of individual differences in visual imagery vividness: a critical quantitative review and plea for direction. *Journal of Mental Imagery,* 19: 1-106.

McKenna, F.P (1984) Measures of field dependence: cognitive style or cognitive ability? *Journal of Sociology and Social Psychology* 47,593-603.

McKenna, F. (1990) Learning Implications of Field Dependence – independence Cognitive style versus Cognitive Ability. *Applied Cognitive Psychology*, Vol 4, pp 425-437.

McLeod, D.B., Carpenter, T.P., McCormack, R.L., & Romuladas, S. (1978). Cognitive style and mathematics learning: The interaction of field independence and instructional treatment in numeration systems. *Journal for Research in Mathematics Education*, 9, 163-174.

Marks, D.F. (1973) Visual imagery differences in the recall of pictures. *British Journal of Psychology*, 64, 17-24.

Marks, D.F. (1989) Bibliography of research utilizing the Vividness of Visual Imagery Questionnaire. *Perceptual and Motor Skills* 69, 707-718.

Martin, Neil, G (1998) *Human Neuropsychology*, Prentice Hall. UK. ISBN Number 0-13-802331-X.

Mayer, R.E. and Massa, L.J. (2003) Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology*, 95 (4), pp. 833-846.

Massa, J.L. and Mayer, E.R. (2005) Three obstacles to validating the Verbal-Imager subtest of the Cognitive Styles Analysis, *Personality & Individual Differences* 39 (2005) 845-848.

Messick, S. and Associates (1976) *Individuality in Learning*, San Francisco; Jossey Bas Publishers.

Microsoft [Computer Software] Visual Interdev 1998-2003 Microsoft Corporation, Redmond, Oregon USA.

Mintzberg, H. (1976) Planning on the left side and managing on the right *Harvard Business Review*, July - August, 49 - 58.

Moran, A. (1985) Unresolved issues in research on Field dependence - independence *Social Behaviour and Personality* 3 (2) pp. 119-125.

Moran, A. (1993) Conceptual and methodological issues in the measurement of mental imagery skills in athletes *Journal of Sport Behavior* Sept 1993 v16 n3 p156 (15)

Morgan A.H., MacDonald P.J. and MacDonald, H. (1971) Differences in bilateral alpha activity as a function of experimental task, with a note on lateral eye movement and hypnotizability *Neuropsychologica* 383, 812-5.

Negroponte, N. (1995) *Being Digital* Hodder & Stoughton London GB 1995.

Nettelbeck, T. and Lally, M. (1976) Inspection time and measured intelligence, *British Journal of Psychology*, 67, pp. 17-22.

Nettelbeck, T. (2001) Correlation between inspection time and psychometric abilities: a personal interpretation, *Intelligence* 29, pp. 459-474.

Nunnally, J.O. (1978) *Psychometric Theory*, McGraw-Hill, New York.

Oltman, P.K. (1968) A portable rod-and-frame apparatus, *Perceptual and Motor Skills*, 26, 503 - 506.

Opperman, R (1994) Adaptively Supported Adaptability, *International Journal Of Human Computer Studies*, 40: 455-472.

Osborne, K. and Gale, A. (1976)) Bilateral **EEG** differentiation of stimuli, *Biological Psychology, 14, pp. 185–196.*

Otter, M. & Johnson, H. (2001): Lost in hyperspace: metrics and mental models. *Interacting with Computers* 13 (1), 1-40.

Packer, J. and Bain, J.D. (1978) Cognitive style and teacher – student compatibility. *Journal of Educational Pyschology*, 70 , 864-871.

Paivio, A. (1971) Styles and strategies of learning, *British Journal of Educational Psychology*, 46, 128 - 148.

Paivio, A. and Begg, I, (1974) Pictures and words in visual search. *Memory and Cognition*, Vol 2, No. 3, pp. 515-521.

Paivio, A and Harshman, R. (1983) Factor Analysis of a Questionnaire on Imagery and Verbal Habits and Skills. *Canadian Journal of Psychology* Vol 37 pp 461-483.

Papert, S (1996) *The Connected Family - Bridging the Generation Gap.* Longstreet Press Inc. USA p47.

Parasuraman, R., Mouloua, M., Hilburn, B. (1998) Adaptive Aiding and Adaptive Task Allocation Enhance Human – Machine Interaction. In M.W Scerbo and M. Mouloa (Ed), *Proceedings of the Third conference on Automation Technology and Human Performance* (119-123). Norfolk, VA.

Parkinson, A.P., & Redmond, J.A (2001) The effect of cognitive styles and educational computer environments on learning performance" *Proceedings of the 6th Annual Conference of the European Learning Styles Information Network (ELSIN)* ed. M Graff et al 25-27th June 2001 Glamorgan Wales.

Parkinson, A.P., & Redmond, J.A (2002a) "TCD CS-2002-18 Results from a study on the affects of Cognitive Styles on Learning performance in 3 different computer environments", April 2002 Computer Science, Dept. of Trinity College Dublin Ireland.

Parkinson, A. & Redmond, James A. (2002b) Do cognitive styles affect learning performance in different computer media? Inroads - SIGCSE Bulletin Vol 34 No. 3 (pp 39-43) Sept 2002 ACM Press NewYork Association of Computing Machinery NewYork also in Proceedings of the 7th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education June 24-28, Aarhus, Denmark 2002 and in the ACM Digital Library.

Parkinson, A.B., Mullally, A.A.P., and Redmond, J.A. (2002c) "The affect of working memory and cognitive styles on learner performance" Learning Styles: Reliability and Validity *Proceedings of The 7th Annual ELSIN Conference* (pp 265-273) June 26-28 2002, Ghent, Belgium

Parkinson, Adrian and Redmond, James A. (2003) "An Adaptive Web Interface to Accommodate User Cognitive Style" *Adjunct Proceedings of the Tenth Annual conference on Human - Computer Interaction* Hersonnisos Crete June 22nd - June 27th 2003 ed Constantine Stephanidis pp 127-128 Parkinson, Adrian and Redmond, James A.

Parkinson, A., Redmond, J. & Walsh, C. (2004a) Accommodating Field-dependence: a Cross-over Study Inroads - SIGCSE Bulletin Vol

36, No 3 September (2004) ACM Press New York Adrian Parkinson, James A. Redmond & Cathal Walsh also in *Proceedings of ITiCSE 2004 ACM SIGCSE 9th Annual Conference on Innovation and Technology in Computer Science Education* Leeds, England, June 28-30, 2004 and also published in the online ACM Digital Library (http://portal.acm.org).

Parkinson, A, Mullally, A.P. and Redmond, J.A. (2004b) The internal reliability of the VVIQ Elsin Seminar, University of Durham Sept 2004

Parkinson, A, Mullally, A.A.P. & Redmond, James A. (2004c) Test-Retest Reliability of Riding's Cognitive Styles Analysis Test *Personality and Individual Differences* Vol 37/6 pp 1273-1278 Oct 2004 and also in Science Direct (Elsevier) database http://www.sciencedirect.com/.

Parkinson, A & Redmond, James A. (2005) The Accommodation of the Field-dependent Learner in Web Design, *The Psychology of Education Review* Vol 29 No 1 March 2005 pp 43-53.

Parkinson, A., Mullally, A. & Redmond, J. (2005) A test-retest of the VItrap test of the Verbaliser-Imager dimension Adrian Parkinson, Alan Mullally & James A Redmond *Proceedings of the 10th Annual Conference of the European Learning Styles Information Network (ELSIN)* ed. E Sadler-Smith et al 13-15th June 2005 Guildford Surrey England

Parkinson, A., Mullally, A, Moore, C & Redmond, J (2005) A test of the external convergent validity of VItrap using the VVIQ, *Proceedings of the 10th Annual Conference of the European Learning Styles Information Network (ELSIN)* ed. E Sadler-Smith et al. 13-15th June 2005 Guildford Surrey England.

Parkinson, A., Redmond, J.A. & Mullally, A. (2006) Is cognitive style as assessed by VItrap independent of Intelligence and Personality?, *Proceedings of the 11th Annual Conference of the European Learning Styles Information Network (ELSIN), 11th Annual Conference of the European Learning Styles Information Network (ELSIN) 2006, University of Oslo, Norway, June 12 - 15, 2006.*

Pask, G. and Scott, B.C. (1972) Learning strategies and individual competence, *International Journal of man-Machine Studies*, 4, 217-253.

Pask, G. (1973) CASTE: a system for exhibiting learning strategies and regulating uncertainties, *International journal for Man-Machine Studies*, 5, 17-52.

Pask, G. (1976) Styles and strategies of learning, *British Journal of Educational Psychology*, 46, 128-148.

Peterson, S.E., Fox, P.T., Posner, M.I. , Mintum, M. and Raichle, M.E. (1988) Positron emission tomographic studies of cortical anatomy of single word processing, *Nature* 331, 585-589.

Peterson, E.R. (2003) *Cognitive Style: The reliability and validity of verbal-imagery and wholistic-analytic cognitive style*. Ph.D thesis, University of Edinburgh.

Peterson, E.R., Deary, I.J. and Austin, E.A., (2003a) The Reliability of Riding's Cognitive Style Analysis Test, *Personality and Individual Differences* 34 5, pp. 881–891.

Peterson, E.R., Deary, I.J. and Austin, E.A. (2003b) On the Assessment of Cognitive Style: Four Red Herrings, *Personality and Individual Differences*, Vol. 34, 5, April 2003 pp. 899-904.

Peterson (2004) Elsin Winter Newsletter "Are we on a quasi evangelical crusade to transform all levels of education?" Some thoughts on the Coffield et al. learning style report pp 5-6.

Peterson, E.R., Deary, I.J. and Austin, E.J. (2005a) A new measure of Verbal–Imagery Cognitive Style: VICS *Personality and Individual Differences*, Volume 38, Issue 6, April 2005, Pages 1269-1281.

Peterson, E.R., Deary, I.J. and Austin, E.J. (2005b) Are intelligence and personality related to the verbal-imagery and wholistic-analytic cognitive styles? *Personality and Individual Differences*, Volume 39, Issue 1, July 2005, Pages 201-213.

Pillay, H. (1988) An Investigation of the effect of Individual Cognitive preferences on Learning through Computer- based Instruction. *Educational Psychology*, Vol 18, No 2 1998, pp 171-181.

Piombo, C., Batatia, H. and Ayache, A. (2003). A Framework for Adapting Instruction to Cognitive learning styles, *Proceedings of the Third IEEE International conference on Advanced learning Technologies*

Potter, M.C. and Fulconar, B.A. (1975) Time to understand pictures and words, *Nature*, 253, pp 437-438.

Pritchard, D.A. (1975) Levelling - sharpening revisited, *Perceptual and Motor Skills*, 40, 111-117.

QUIS (2006) http://www.cs.umd.edu/hcil/quis/ and also at http://www.lap.umd.edu/QUIS/index.html

R (2006) The R project for statistical computing http://www.r-project.org/

Ramirez, M. (1973) Cognitive styles and cultural democracy in education. *Social Science Quarterly*. 53 pp 895-904.

Reber, A.S. Reber, E. (2001) *The Penguin Dictionary of Psychology* Third Edition Published by the Penguin Group London, UK.

Redmond, J.A., Mullally, A.A.P., and Parkinson, A.B. (2002) "Test-retest reliability of Riding's CSA" Learning Styles: Reliability and Validity *Proceedings of The 7th Annual ELSIN Conference* (pp 303-313) June 26-28 2002, Ghent, Belgium.

Redmond, James A., Walsh, C and Parkinson, A. (2003) Equilibriating Instructional Media for Cognitive Styles *Inroads - SIGCSE Bulletin* Vol 35, No 3 September (2003) pp 55-59 ACM Press NewYork Association of Computing Machinery, NewYork also in *Proceedings of the 8th Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE 2003)* ed. David Finkel June 30 - July 2 Thessaloniki, Greece (2003) and in the ACM Digital Library.

Redmond, J.A., Parkinson, A (2004a) Designing a Web Interface Based on Cognitive Style Scores James Redmond and Adrian Parkinson, 112th Annual Convention of the American Psychological Association Division 21 - Applied Experimental and Engineering Psychology: 3039 Paper Session: *Effects of Individual, Group and Cultural Differences on Human Performance Honolulu, Hawaii July 28 – August 1, 2004*

Redmond, J.A., Parkinson, A & Alan Mullally, (2004b) A new test of the Verbaliser – Imager dimension - VItrap, Elsin Seminar, University of Durham September 2004.

Redmond, J.A., Parkinson, A. & Mullally, A. (2006) VItrap Test Validity: Use of Curry's Onion Model, *Proceedings of the 11th Annual Conference of the European Learning Styles Information Network (ELSIN), 11th Annual Conference of the European Learning Styles Information Network (ELSIN) 2006, University of Oslo, Norway, June 12 - 15, 2006[.*

Rezaei, A. and Katz, L (2004) Evaluation of the reliability of the Cognitive Styles Analysis, *Personality and Individual differences*, Vol 36, issue 6, April pp. 1,317 - 1327.

Rezler, A.G. and Rezmovic, V. (1981) The Learning Preference Inventory, *Journal of Allied Health*, Feb. 10(1) pp. 28-34.

Richardson, A (1977) Verbaliser -Visualizer A cognitive style dimension *Journal of Mental Imagery* pp 109-125

Riding, R.J. & Taylor, E.M. (1976) Imagery performance and prose comprehension in 7-year old children, *Educational Studies* 2, 21-27.

Riding, R.J. & Wicks, B.J. (1978) The effects of extraversion and presentation order on learning from picture commentary sequences by children , *Educational Review*. Vol. 30, n 3, pp 255-257.

Riding, Richard J. and Watts, Mary, (1979) The effect of cognitive style on the preferred format of instructional material, *Educational Psychology* March – June, Vol 17, pp 179-183

Riding, R.J., Ashmore J. (1980) Verbaliser-imager learning style and children's recall of information presented in pictorial versus written form, *Educational Psychology*, vol 6 , no 2 June 1980 – pp 141-145.

Riding, R. and Dyer, V. (1980) The relationship between extraversion and verbal imagery in twelve year old children, *Personality & Individual Differences*, 1, 273-279.

Riding, R. and Calvey, I. (1981) The assessment of verbal imagery learning styles and their effect on the recall of concrete and abstract

prose passages by eleven year old children, *British Journal of Psychology* 72, 59-64.

Riding, R.J. and Anstey, L. (1982) Verbal Imagery learning learning style and reading attainment in eight year old children, *Journal of Research in Reading* 5, 57-66.

Riding, R.J., Buckle, C.F., Thompson, S. & Hagger, E. (1989) The Computer Determination of Learning Styles as an Aid to Individualised Computer-Based Training. *Educational and Training Technology International*, 26, 293-298

Riding, R.J. (1991a) *Cognitive Styles Analysis*, Birmingham. Learning and Training Technology.

Riding, R. J., (1991b) *Cognitive Styles Analysis - Research Administration*. Birmingham: learning and training technology.
Riding, R. and Cheema, I. (1991) Cognitive styles, an overview and integration, *Educational Psychology*, Vol 11, Nos 3 and 4 pp. 193-215.

Riding R.J., Sadler –Smith E (1992), Type of Instructional Material Cognitive style and learning Performance, *Educational Studies* Vol 18, No 3, 1992 pp 323-40.
Riding, R.J. and Borg, M.G. (1993) Teacher stress and cognitive style, *British Journal of Educational Psychology*, Jun, 63, Pt (2) 271-286.

Riding, R.J. and Douglas, G. (1993) The effect of cognitive style and mode of presentation on learning performance, *British Journal of Education Psychology*, 63, 297-307.

Riding, R.J., Glass, A, and Douglas, G. (1993) Individual differences in thinking: cognitive and neurophysiological perspectives, *Educational Psychology*, Vol. 13, pp 267-279.

Riding, R., & Pearson, F. (1994). The relationship between cognitive style and intelligence. Educational Psychology, 14, 413–425.

Riding R.J. and Rayner, S. (1995) The Information Superhighway and Individualised Learning, *Educational Psychology* Vol 15, No 4 1995.

Riding R.J. and Wheeler, H. (1995) Occupational stress and cognitive style in nurses, *British Journal of Nursing*, 4, 160-168.

Riding R.J., Read, G. (1996) Cognitive style and Pupil Learning Preferences, *Educational Psychology*, Vol 16, No 1 1996.

Riding R.J. and Agrell, T., (1997) The effect of cognitive style and cognitive skills on school subject performance, *Educational Studies*. 23, pp 311-323.

Riding, R.J., Glass, A., Butler, S.R. and Pleydell-Pearce, C.W. (1997) Cognitive style and individual differences in EEG alpha during information processing, *Educational Psychology*, Vol. 17, pp 219-34.

Riding R.J. and Watts M., (1997) The effect of Cognitive Style on the preferred format of instructional material, Educational *Psychology*. 01443410, Mar-June 97, Vol 17, pp 179-183.

Riding R.J. and Wigley, S. (1997) The relationship between cognitive style and personality in further education students, *Personality and Individual differences*, Vol. 23, 379-89.

Riding, R.J. and AI Sanabani, S. (1998) The effect of cognitive style, age, gender and structure on the recall of prose passages, *International Journal of Research Education*, Vol. 29, issue 3.

Riding, R. and Rayner, S. (1998) *Cognitive styles and learning strategies, Understanding Style Differences in Learning and Behaviour*. David Fulton Publishers London.

Riding, R. J., (1999) *Cognitive Styles Analysis - Research Administration*. Birmingham: learning and training technology
Riding, R.J. & Grimley, M. (1999) Cognitive style and learning from multimedia materials in 11 year old children. *British Journal of Educational Technology* 30, 43-56.

Riding, R. J., (2000a). *Cognitive Styles Analysis - Research Administration*. Birmingham: learning and training technology.

Riding, R. J., (2000b) Cognitive style - a review In: R. J. Riding & S. G. Rayner, (Eds.), *Cognitive style: international perspectives on individual differences*, Cognitive styles (Vol. 1), (pp. 315–347). Stamford, Connecticut: Ablex Publishing Corporation.

Riding, R.J. (2000c). Information Processing Index - Research Administration and Application, Learning & Training Technology Bermingham UK 2000.

Riding, R. and Rayner, S. (2000) Cognitive style: international perspectives on individual differences, Cognitive styles (Vol. 1), (pp. 315–347). Stamford, Connecticut: Ablex Publishing Corporation.

Riding, R.J. (2001) *CSA Cognitive styles analysis - Research Administration and Application*, Learning & Training Technology Bermingham UK 2001.

Riding, R.J., 2003. On the assessment of cognitive style: a commentary on Peterson, Deary and Austin, *Personality and Individual Differences* 34 5, pp. 893–897.

Rogers, L.J. (1980) Lateralisation in the avian brain, *Bird Behaviour*, 2, 1-12.

Rossion, B. and Pourtois, G. (2001) Revisiting Snodgrass and Vanderwart's object database: Colour and texture improve object recognition. In Sarasota: *Vision Sciences.*

Rosen, R (1985) Anticipatory Systems – philosophical, Mathematical and Methodological Foundations. Pergamon Press, 1985 reviewed in Rothrock, l., Koubeck R., Fuchs F., Haas M. and Salvendy, G. (2002) Review and Reappraisal of adaptive interfaces toward biologically – inspired paradigms. *Theoretical Issues in Ergonomic Science* pp 1- 82. Rothrock, l., Koubeck R., Fuchs F., Haas M. and Salvendy, G. (2002) Review and Reappraisal of adaptive interfaces toward biologically – inspired paradigms. *Theoretical Issues in Ergonomic Science* pp 1- 82.

Rowan, James, A, and Tolunsky, E (2003) *Primer of EEG* . Publishers Buttterworth Heinemann Philadelphia, Pennsylvania, USA. ISBN - 0-7506-7476-8.

Rust, J. and Golombok, S. (1999) *Modern Psychometrics: the science of psychological assessment*, 2nd. Ed., Routledge, London and New York.

Sadler-Smith, E. and Badger, B. (1998) Cognitive style, learning and innovation, *Technology Analysis and Strategic Management*, 10, pp. 247-265.

Sadler-Smith, E. and Riding. R.J. (1999) Cognitive Syle and Instructional Preferences, *Journal of Instructional science*, 27, pp. 355-371.

Santostefano, S. (1964) Developmental study of the cognitive control "levelling - sharpening", *Merrill Palmer Quarterly*, 10, 343-360.

Saracho, O.N. and Dayton, C.M (1980) Relationships of teachers' cognitive styles to pupils' academic gains *Journal of Educational Psychology*, 72 .544-549.

Saracho, O.N. (1993) The effects of teachers' cognitive styles on their Sstudents' academic achievement, *International Journal of Early Childhood*, 25 (2) 37-40.

Saracho, O.N. (2000) A framework for effective classroom teaching: matching student and teachers' and students' cognitive styles In: R. J. Riding & S. G. Rayner, (Eds.), *Cognitive style: international perspectives on individual differences, Cognitive styles (Vol. 1),* (pp. 315–347). Stamford, Connecticut: Ablex Publishing Corporation

Satterly, D.J., & Telfer, I. G. (1979) Cognitive style and advance organizers in learning and retention. *British Journal of Educational Psychology*, 49, 169-178.

Schmeck, R., Ribich, F. and Ramanaish, N. (1977) Development of self-report inventory for associating individual differences in learning processes. *Applied Psychological Measurement*, 1(3).

Sergent, J., (1990) Furtive incursions into bicameral minds: integrative and co-ordinating role of subcortical structures, *Brain*, 113, 537-568.

Sheenan P, W. (1967) A shortened form of Bett's Questionnaire, *Journal of Clinical Psychology* 34/23 pp 386-389.

Sheridan, T.B. and Verplanck, W.L (1978) Human and computer control of undersea teleoperators, Cambridge , MA: MIT Man – Machine Laboratory reviewed in Rothrock, l., Koubeck R., Fuchs F., Haas M. and Salvendy, G. (2002) Review and Reappraisal of adaptive interfaces toward biologically –inspired paradigms. *Theoretical Issues in Ergonomic Science pp. 1- 82.*

Shneiderman, B, (1995) Looking at the bright side of user interface agents, *Interactions*, 2(1) 13-15.

Shneiderman, B (1997) Direct manipulation for comprehensible, predictable and controllable user interfaces, *Proceedings of the 2nd international Conference on Intelligent User Interfaces* Orlando, Florida, United States pp 33-39 ISBN:0-89791-839-8 ACM Press New York, NY,

Shneiderman, B., Maes, P. (1997): Direct manipulation vs. interface agents, *Interactions* 4(6): 42-61 (1997).

Shneiderman, B. (1998) *Designing the User Interface, Strategies for Effective Human-Computer Interaction*, Third edition, Addison Wesley Reading, Mass. ISBN 0-201-69497-2.

Shneiderman, B. (1999). Human values and the future of technology: a declaration of responsibility *ACM SIGCAS Special Issue: SIGCAS at 30* 29 (3) 5 - 9. Sept 1999

Shneiderman, B. (2004a) Personal communication.

Shneiderman, B. (2004b) *Designing the User Interface, Strategies for Effective Human-Computer Interaction*, Fourth edition, Addison Wesley Reading, Mass. ISBN 0-201-69497-2.

Shneiderman, B. (2005) Personal communication.

Smith, M.C. and Magee, E.L. (1980) Tracing the time curve of picture - word processing, *Journal of Experimental Psychology* Vol. 109, No. 4 pp. 373-392.

Somberg, B.L. (1987) A comparison of ruled based and positionally constant arrangements of computer menu items. *Proceedings of CHI & GI 87 Conference on Human factors in Computing Systems*. New York ACM.

Sobol, S. and Stones, C. (2005) Evaluation Usage of Dynamic Web Site Structures through Visualisation in *Adaptable and Adaptive Hypermedia Systems* , Sherry Y. Chen, George D. Magoulas (2005). published by IRM Press Hersely, London, Melbourne and Singapore.

Spence, W.J., Tsai, R.J (1997) On human cognition and the design of information systems, *Information & Management* 32 pp 65-73.

Speilberger, C.D, (1977) State and Trait Anxiety Inventory Form Y-1. Palo Alto California: *Consulting Psychologists' Press.*

Sperry, R.W. (1964) The great cerebral commissure, *Scientific American*, Jan, 42 - 52.

SPSS V12 (Computer Software) (2002) Systat Software Inc. 2002 http://www.spss.com/scienceproducts.htm

Springer, S.P. and Deutsch, G. (2003) *Left Brain, Right Brain: Perspectives from Cognitive Neuroscience*, 5th Ed. W.H. Freeman & Co., Worth Publishers

Sterman, B., (2002) Personal communication and notes, Philadelphia, USA.

Sterman B, and Malkovitz D, (2002) The accuracy of the automatic artifacting in SKIL when compared with traditional methods, unpublished report, Philadelphia.

Sternberg, R.J. (1997) *Thinking Styles*, Cambridge University Press, NY.

Sternberg, R. & Grigorenko, E. (1997) Are cognitive styles still in style? *American Psychologist Vol* 52, No. 7, p700 – 712.

Sternberg R.G and Zhang L.F. (2001) *Perspectives on Thinking, Learning, and Cognitive styles*, Lawrence Erlbaum Associates, Publishers, Mahwah, NJ and London.

Sternberg, R.J. (2002) *Thinking Styles* (second edition) Cambridge University Press UK.

Sternberg, R. (2005) Personal communication.

Tan, B.W. & Lo, T.W. (1991) The impact of interface customization on the effect of cognitive style on information system success. *Behaviour & Information Technology* 1991 10 (4), 297-310.

Tennant, M. (1988) *Psychology and adult learning* London Routledge.

Tetlock, P.E. (1983) Cognitive style and political ideology, *Journal of Personality and Social Psychology*, 45, pp. 118-126

Tiedemann, J. (1989) Measures of cognitive style: a critical review. *Educational Psychology* 24, 261-75.

Torenvliet, GL, Jamieson, GA, & Vicente, KJ (2000). Making the most of ecological interface design: The role of individual differences. *Fourth Symposium on Human Interaction with Complex Systems,1998*.http://doi.ieeecomputersociety.org/10.1109/HUICS.1998.65999

Triantafillou, E., Pomportsis, A and Demetriadis, S. (2003) The design and Formative evaluation of an adaptive educational system based on cognitive styles. *Computers & Education* Vol 41, Issue 1

Tullett, A.D. and Kirton, M.J. (1995) Further evidence for the independence of the Adaptive - Innovative (A-I) cognitive style from national culture, *Personality & Individual differencse*, 19 (3) 393-396.
Tullett, A.D. (1997) Cognitive style: not culture's consequence *European Psychologist* Vol 2 No. 3 Sept pp 258-267.

Van den Broeck, H., Vanderheyden, K., and Cools, E. (2003) The Field of Cognitive Styles from a theoretical review to the construction of the cognitive style inventory. *Vlerick Leuven Gent Working Paper Series* 2003/26, D/3003/6482/27. Published by Vlerick Leuven Gent, Management School. http://www.vlerick.be/research/workingpapers/vlgms-wp-2003-26.pdf

Vick, O. and Jackson, D. (1967) Cognitive styles in the schematizing process: a critical evaluation, *Educational and Psychological Measurement* 27, 267-286.

Virvou, M. (1999) Automatic reasoning and help about human errors in an operating system. *Interacting with Computers*, 11, 545-573.
Wammi (2006) http://www.wammi.com/

Weibelzahi, S. (2005) Problems and Pitfalls in the evaluation of adaptive systems in *Adaptable and Adaptive Hypermedia Systems* , Sherry Y. Chen, George D. Magoulas (2005) published by IRM Press Hersely, London, Melbourne and Singapore.

Weller, Herman G., Repman, J., William, L. and Rooze G, (1995). Improving the effectiveness of learning through Hypermedia and based instruction. The Importance of Learner Characteristics. *Computers in Human Behaviour*, Vol 11, No. 3-4, pp 451-465.

Wickens, C.D. (1992) *Engineering Psychology and Human Performance* (2nd ed.). New York: Harper Collins.

Wilborn, M.R. (1981). *An investigation of the relationships among proportional reasoning, field-dependence/independence, sex, and grades in science of eight grade ISCS students* (Doctoral Dissertation, Georgia State University), Dissertation Abstracts International, 42, 92.

Williamson, S.J., Kaufman, L, Lu, Z-L, Wang, J-Z and Karron, D., (1977) Study of occipital alpha rhythm and alpha suppression, *Internationa Journal of Psychophysiology*, 26, 63-76.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.

Witkin, H.A. & Asch S. E., (1948). Studies in space orientation: IV. Further experiments perception of the upright with displaced visual fields *Journal of Experimental Psychology* Vol 38, 762-782.

Witkin, H.A., (1965) Some implications of research on cognitive style for problems of education, *Archivio di Psicologia, Neurologia e Psichatrio*, 26: 27-54.

Witkin, H.A., Lewis, H.B and Weil, E (1968) Affective actions and patient-therapist interactions among more differentiated patients early in therapy. *Journal of Nervous and Mental Disease*, 146 pp 193-208.

Witkin, H.A., Oltman, P., Raskin, E., Karp, S. (1971). *A Manual for the Embedded Figures Test.* Palo Alto, CA: Consulting Psychologists Press.

Witkin, H. Cognitive style in academic performance and in teacher student relations in Messick, S. (1976). *Individuality in Learning: Implications of Cognitive Styles and Creativity For Human Development.* San Francisco, CA: Jossey Bass.

Witkin, H .A. and Goodenough, D.R. (1977) Field dependence and interpersonal behaviour. *Psychological Bulletin*, 84 pp661-689

Witkin, H. A., Moore. C.A., Goodenough, D. R and Cox, P.W (1977) Field dependent and field independent cognitive styles and their educational implications. *Review of Educational Research*. 47,1-64.

Workman, M. (2004) Performance and perceived effectiveness in computer based and computer –aided education: do cognitive styles make a difference *Computers in Human Behaviour* Vol 20 , Issue 4, pp 517 – 534.

Zatorre, R.J. (1992) Evans, A.C., Meyer, E. and Gjedde, A. (1992) Lateralisation of phonetic and pitch discrimination in speech processing, *Science*, 256, 846-849.

Zhang, L. and Sternberg, R.J. (2005) A threefold model of intellectual styles, *Educational psychology Review*, 17 (1) pp. 1-53.

Zmud, R.W. (1979) Individual differences and MIS success: A review of the empirical literature, *Management Science* Vol 25, No 10, October pp 966-977.

Zywno, M.S., (2003) A contribution of validation of a score meaning for Felder-Solaman's Index of Learning Styles, *Proceedings 2003 of the Annual ASEE Conference.*

Appendix 1-2.1.A Rothrock et al.'s Generic Adaptive Model - Table A

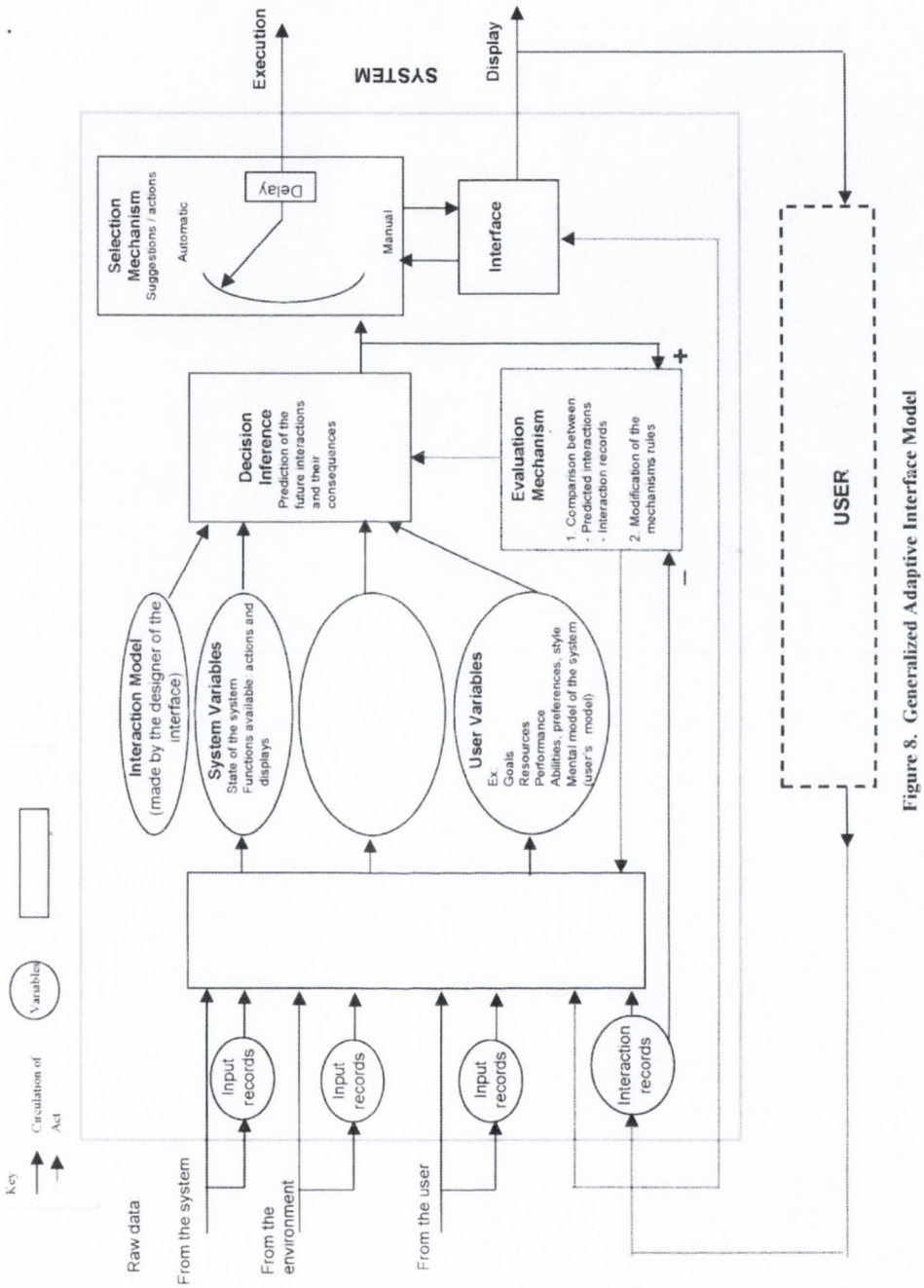


Figure 8. Generalized Adaptive Interface Model

Taken from "Review and reappraisal of adaptive interfaces toward biologically-inspired paradigms" by Ling Rothrock, Richard Koubek, Frederic Fuchs, Michael Haas and Gavriel Salvendy

Appendix 2-2.2.B
Rothrock et al.'s Review of Eighteen Adaptive Systems - Table B

Table 5. Survey of Interfaces for Real-time Adaptability.

Author of the article	Application	Input	Identification Mechanism	Calling Variables	Decision Mechanism	Output	Primary dimension of the adaptation		
							Abilities	Tasks	Situation
Beaumont (1994)	Anatom-Tutor: anatomy tutoring system	Keyboard and mouse	Weighted rules (LISP)	User's knowledge = correct answers to key questions	Rules (LISP)	Stretch texts	X		
Bergstrom et al. (2000)	Data mining via a WAP application, which needs to minimize the length of the information displayed.	Phone keyboard	Genetic programming (Prolog)	User's goal = keywords defining his search	Genetic programming and semantic networks	Selection of information			X
Boyle and Encarnacion (1994)	MetaDoc: online documentation system of Unix.	Keyboard and mouse	Rules (LISP)	User's knowledge = initial stereotype (form) + prior user's requests	Rules	Stretch texts. Less explanations and more details for the experienced users	X		
Brusilovsky (1998)	Interbook delivers adaptive electronic textbooks on the web.	Keyboard and mouse	Rules	User's knowledge of concepts	Rules	Selection of information + link annotation: Guidance, navigation support, help	X		
De Bra and Calvi (1998)	AHA (Adaptive Hypermedia Architecture), a generic adaptive hypermedia system to deliver courses through the web. Also possible: on-line help, on-line information system, institutional hypermedia, personalized views.	Keyboard and mouse	Rules	User's knowledge = answers to tests or pages read (pairs of concepts and Boolean values)	Rules	Selection of texts. And link annotation or link hiding, according to the user's choice	X		

(table continues)

Table 5. (continued)

Francisco-Revilla and Shipman (2000)	The Mars Medical Assistant will give medical information to crewmembers flying to Mars. The medical information is selected according to the user's model, the situation model and the task model (describing procedures, aiding diagnosis or provide information on health concerns)	Keyboard and mouse?	None	User, situation and task model are selected explicitly.	Predicate logic including conflict resolution algorithms	Selection of the content and adaptation of the presentation.		X	
Goecks and Shavlik (2000)	This agent models the user's interest in a web page. The user's interest in previously visited pages is inferred by his activity while reading the page (i.e., number of hyperlinks clicked, amount of scrolling...). A page is broken down into a bag-of-words.	Records of keyboard and mouse actions + records of visited pages	Neural network (fully connected, three-layer)	User's interest = Key words	Not implemented yet	Not implemented yet		X	
Gong and Salvendy (1995)	Gentle "push" for the user to learn the command-base interface. First, the menus prompt the corresponding command. Progressively, only the command mode is available.	Computer keyboard and mouse	Predicate logic	Knowledge levels = frequency of use of functions	Predicate logic	Deactivate the menus / hide the menus	X		
Harper, Guarino and Zacharias (2000)	Air Traffic Controller Interface to support controller situation awareness	Event cues from user interaction and from monitoring the task environment	Fuzzy logic and Belief Network	User's mental model	Rules	Display of flight plans, anticipated conflicts, and proposed solutions for conflict resolution			X
Höök (1998)	PUSH selects information by inferring the user's goal from his interactions.	Computer keyboard and mouse. Record of the user's actions.	Plan recognition with predicate logic (Prolog)	User's goal = one of the prelisted information seeking tasks	Rules	Selection of what to hide and what to show on a page using a stretch text technique.		X	

(table continues)

Table 5. (continues)

Horvitz, E. (1998)	The Microsoft Office Assistant gives contextual help according to the program state and the user's goal. The Lumiere research project (which engendered MOA) also inferred the user's knowledge.	Program state, records of "atomic" sequences of user's actions and queries.	Bayesian statistics	User's goal (thousands of different goals identified)	Rules	Selection of information		X	
Iba and Gervasio (1999)	INCA (interactive crisis assistant) helps users develop effective crisis response plans and schedules in a timely manner.	Mouse and keyboard records: previous solutions chosen by the user	Different classifier algorithms experimented with largely similar results	User's preference: the next step the user is the most likely to select	Classifier algorithm	Highlight of a selected information			X
Jaczinsky and Trousse (1998)	Broadway is a world wide web browsing adviser that reuses past navigations from a group of users.	Records of navigations of a group of users: URL, extracted keywords, user's evaluation	Rules	Group profile= Keywords and appreciation of related web pages. It is compared with the user's goal (= keywords of the request).	Rules	The user can choose between continuous recommendation or recommendation upon request.		X	
Jameson et al. (1999)	Ready is an experimental system that adapts the explanations and advices given to user's in emergency cases, for instance phone calls to firemen or auto repairs.	Natural language	Rules	User's time constraints, work memory load, knowledge level and emotional stress	Bayesian network (Netica)	Selection and quantity of information.	X		
Miah and Alty (2000)	Adaptive desktop window management: the inactive windows vanish progressively after a time-out.	Time during which a window is not used.	Rules	User's goal	Rules	Progressive reduction of the window size, movement to the sides of the screen.		X	

(table continues)

Table 5. (continues)

Pazzani and Billsus (1997)	Intelligent agent to identify informative words from web pages and can determine the interest of new pages.	First, user's feedback on a set of previously visited sites, then record of visited pages.	Bayesian statistics + 5 other algorithms tested	Long term user's goal	Naïve Bayesian classifier	Selection of web pages		X	
Rogers, Fiechter and Langley (1999)	The Adaptive Route Advisor gives recommendation of driving routes. It is designed for in-car use. Criteria taken into account for each section of the roads: length, estimated driving time, road class, etc.	Records of previous user's choices (entered by a keyboard)	Differential perceptron	User's preferences = weighted characteristics of 9 attributes of the routes	Rules	Highlight of the recommended route on a map, and sort of alternative routes			X
Wilson (2000)	Experiment to measure the performance with real-time physiologically controlled adaptive aiding. The task is the NASA Multiple Attribute Task Battery.	6 EEG channels + ECG + EOG + respiration	Neural Network	User's workload	Rule	Task allocation: two subtasks are turned-off in high workload conditions.	X		

Appendix 3-4.1.A - Pilot Study (Study 4.1)

The lower end of the Fd/Fi continuum (i.e. low values) indicated Field-dependence; the higher end indicated Field-independence.

Table 4.1.A Simple Linear Regression Model of Exam Score vs. Fd/Fi in Preliminary Study (Study 4.1) (n = 15).

Dependent variable is: **Exam**
No Selector
R squared = 57.3% R squared (adjusted) = 54.1%
s = 6.269 with 15 - 2 = 13 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	686.505	1	686.505	17.5
Residual	510.828	13	39.2945	

Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	42.8710	4.111	10.4	≤ 0.0001
Fd-Fi	1.51882	0.3634	4.18	0.0011

The higher the score on the Fd/Fi continuum, the more field-independent the respondee is. Field-independence is estimated by the GEFT in these studies.

The results show that the Fd/Fi score was statistically significant at the $p = 0.005$ level (since $p = 0.001 < 0.005$ for Fd/Fi). This is confirmed by an R squared of 57.3% which shows that more than 57% of the variability in the model is accounted for, by Fd/Fi. The sample size is small but since there is only one predictor, Fd/Fi, it exceeds the recommendation by Altman of 10 samples per predictor (Altman (1991)).

Appendix 4-4.2.B Summary Statistics for Geft Scores for Studies 4.1, 4.2a and 4.2b

Table B The Mean, Median and Standard Deviation for GEFT scores

Geft Scores	Pilot Study 4.1	Study 4.2a	Study 4.2b
n	15	34	38
Mean	10.4	9.5	10.2
Median	11	10	11
S.D	4.6	4.9	4.8

Appendix 5-4.3.C Summary Statistics for Examination Scores for Studies 4.1, 4.2a and 4.2b

Table D The Mean, Median and Standard Deviation for Examination scores

Exam Scores	Pilot Study 4.1	Study 4.2a	Study 4.2b
n	15	34	38
Mean	60%	13.9	12.5
Median	60%	14	13
S.D	7.1	3.4	3.4
Type of test	Multiple choice	Objective style	Objective style

Appendix 6-5.1.A

The mean values for W-A and V-I Study 1 and 2 for speed index and percentage correct are presented in Table 2.

Table 2. Speed index and percentage correct for W-A and V-I for Study 1 and 2

	Study 1:(n = 51) 14-day mean interval	Study 2: (n = 27) 23-month interval
W-A speed index Session 1	6.45	4.37
W-A speed index Session 2	8.14	5.07
W-A percentage correct Session 1	95.8	95.6
W-A percentage correct Session 2	97.4	94.5
V-I speed index Session 1	3.71	2.57
V-I speed index Session 2	4.95	3.0
V-I percentage correct Session 1	93.6	94.5
V-I percentage correct Session 2	95.7	91.6

Appendix 7-5.2.B Study 5.2.2 Additional Analysis

Product moment Correlation between W-A and V-I with Tests of Intelligence (APM 2 and Mill Hill)

Pearson Product-Moment Correlation

No Selector

	APM2	MHTot	WA	VI
APM2	1.000			
MHTot	0.396	1.000		
WA	-0.053	0.148	1.000	
VI	0.145	0.086	-0.220	1.000

The IQ tests correlate very poorly with both dimensions of the CSA as would be expected (n = 51).

Appendix 8-6.1.A Spearman Browne Prophecy Formula

$N = rd(1 - r0)/(r0(1 - rd))$

N = number of tests of the correct version's length

rd = required level of reliability

r0 = observed level of reliability

r0 = 0.2 based over a 4-week test-retest interval

rd = 0.7

$n = 0.7 * 1.2/(0.2*0.3) = 0.84/0.06 = 14$

Using this formula, this indicates 14* 48 = 672

(48 test items used by Riding in his V-I dimension)

Therefore 672 test items required to establish a test-retest reliability of 0.7.

This is too many test items to be practical due to testee fatigue.

Appendix 9-6.2.B Mean Inspection Times for Total Verbal Forms vs. Image Forms

```
Paired t-Test of μ(1 - 2 )
No Selector
Individual Alpha Level 0.05
Ho: μ(1 - 2 ) = 0 Ha: μ(1 - 2 ) ≠ 0

I_A+V_ASA - V_A+I_A:
Test Ho: μ(I_A+V_ASA-V_A+I_A) = 0 vs Ha: μ(I_A+V_ASA-V_A+I_A) ≠ 0
Mean of Paired Differences = 2.3240524 t-Statistic = 26.02 w/155 df
Reject Ho at Alpha = 0.05
p ≤ 0.0001
```

This mean of the verbal forms of the questions was compared (mean of verbal form were added together, the mean of the image forms were added together) for A. By inspecting the summary statistic table above it would suggest that there was a difference between these two sets of inpection times. The paired t test confirms this. Similarly this would also appear to be the case for Session B.

```
Paired t-Test of μ(1 - 2 )
No Selector
Individual Alpha Level 0.05
Ho: μ(1 - 2 ) = 0 Ha: μ(1 - 2 ) ≠ 0

V_B+I_B - I_B+V_B:
Test Ho: μ(V_B+I_B-I_B+V_B) = 0 vs Ha: μ(V_B+I_B-I_B+V_B) ≠ 0
Mean of Paired Differences = -1.9069771 t-Statistic = -25.47 w/155 df
Reject Ho at Alpha = 0.05
p ≤ 0.0001
```

Table 6.2b: The results of this paired t test indicate that there is a difference between the means in both sessions as was the case in session A.

These results indicate that the image test items were completed significantly faster than the word tests items in Session A and Session B.

Appendix 10-6.3.C The total correct for each test item type for Session A (test) and Session B (retest)

Section – in results.

Sum of each question correct type per session one and two.

Paired t tests.

Paired t-Test of μ(1 - 2)

No Selector

Individual Alpha Level 0.05

Ho: μ(1 - 2) = 0 Ha: μ(1 - 2) ≠ 0

SumIQB - SumIQA:

Test Ho: μ(SumIQB-SumIQA) = 0 vs Ha: μ(SumIQB-SumIQA) ≠ 0

Mean of Paired Differences = 0.57142857 t-Statistic = 5.918 w/153 df

Reject Ho at Alpha = 0.05

p ≤ 0.0001

Paired t-Test of μ(1 - 2)

No Selector

Individual Alpha Level 0.05

Ho: μ(1 - 2) = 0 Ha: μ(1 - 2) ≠ 0

SumVQA - SumVQB:

Test Ho: μ(SumVQA-SumVQB) = 0 vs Ha: μ(SumVQA-SumVQB) ≠ 0

Mean of Paired Differences = -0.058441558 t-Statistic = -0.6466 w/153 df

Fail to reject Ho at Alpha = 0.05

p = 0.5189

Paired t-Test of μ(1 - 2)

No Selector

Individual Alpha Level 0.05

Ho: μ(1 - 2) = 0 Ha: μ(1 - 2) ≠ 0

SumVPB - SumVPA:

Test Ho: μ(SumVPB-SumVPA) = 0 vs Ha: μ(SumVPB-SumVPA) ≠ 0

Mean of Paired Differences = -0.025974026 t-Statistic = -0.4638 w/153 df

Fail to reject Ho at Alpha = 0.05

p = 0.6434

Paired t-Test of μ(1 - 2)

No Selector

Individual Alpha Level 0.05

Ho: μ(1 - 2) = 0 Ha: μ(1 - 2) ≠ 0

SumIPB - SumIPA :

Test Ho: μ(SumIPB-SumIPA) = 0 vs Ha: μ(SumIPB-SumIPA) ≠ 0

Mean of Paired Differences = 0.0064935065 t-Statistic = 0.132 w/153 df

Fail to reject Ho at Alpha = 0.05

p = 0.8951

Paired t-Test of μ(1 - 2)

No Selector

Individual Alpha Level 0.05

Ho: μ(1 - 2) = 0 Ha: μ(1 - 2) ≠ 0

SumVPB - SumVPA :

Test Ho: μ(SumVPB-SumVPA) = 0 vs Ha: μ(SumVPB-SumVPA) ≠ 0

Mean of Paired Differences = -0.025974026 t-Statistic = -0.4638 w/153 df

Fail to reject Ho at Alpha = 0.05

p = 0.6434

Appendix 11-6.4.D Means of the mean times, median, standard deviation, range and standard error for the for each of the four test item categories for Session A (Test) and Session B (Retest).

	Variable	Mean	Median	Std Dev	Range	Std Error
Session A	Vq	3.20	3.07	0.86	4.29	0.07
Session A	Iq	3.11	3.00	0.78	5.14	0.06
Session A	Vp	2.18	2.12	0.53	2.96	0.04
Session A	Ip	1.82	1.78	0.40	2.53	0.03
Session B	Vq	2.64	2.57	0.70	3.62	0.06
Session B	Iq	2.44	2.41	0.56	2.73	0.04
Session B	Vp	1.70	1.63	0.44	2.79	0.03
Session B	Ip	1.48	1.43	0.35	2.59	0.03

Table 6.4: Presents mean of the mean times, median, standard deviation, range and standard error for the for each of the four test item categories for Session A (Test) and Session B (Retest).

Section 6.3.1 Distribution of reaction ratios in Session A

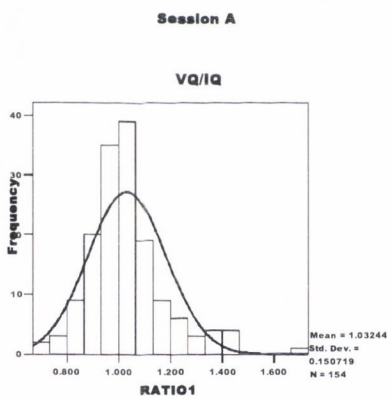


Figure 6.3.C.1 : This figure is a histogram to which a normal curve has been fitted for the ratio VQ/IQ. The distribution is approximately normal indicating a good spread of data.

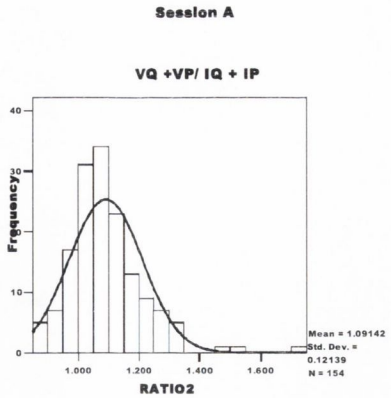


Figure 6.3.C.2 : This figure is a histogram to which a normal curve has been fitted for the ratio VQ+VP/IQ+ IP. The distribution is approximately normal also indicating a reasonable distribution of the data.

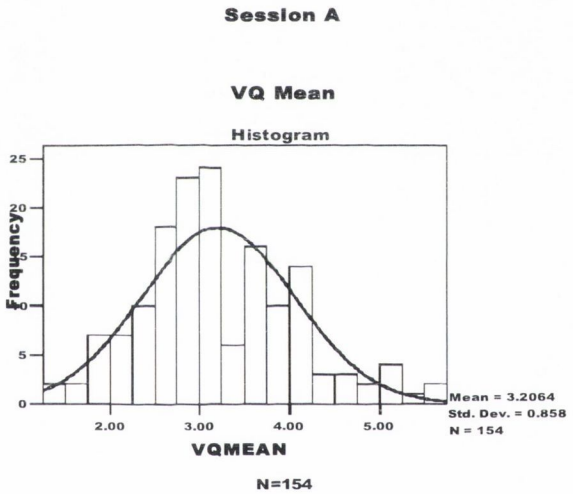


Fig. 6.3.C.3

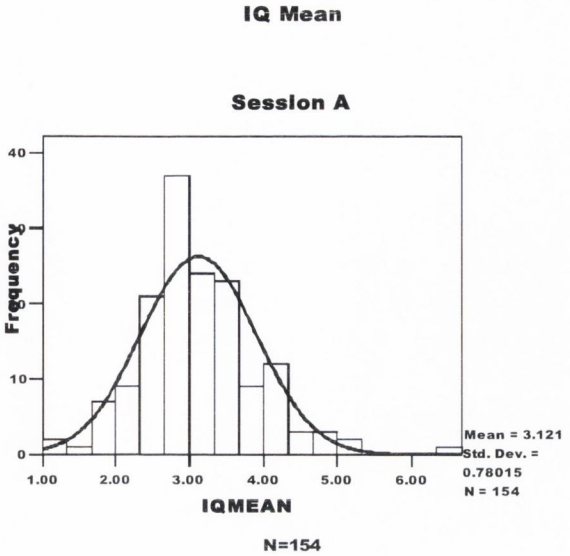


Fig. 6.3.C.4

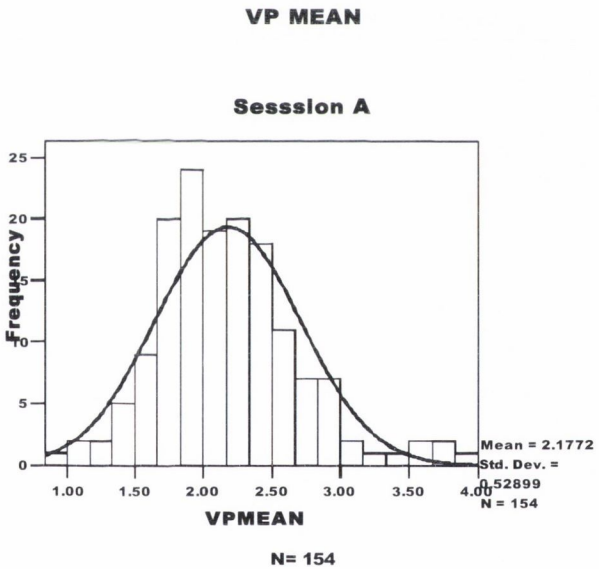


Fig. 6.3.C.5

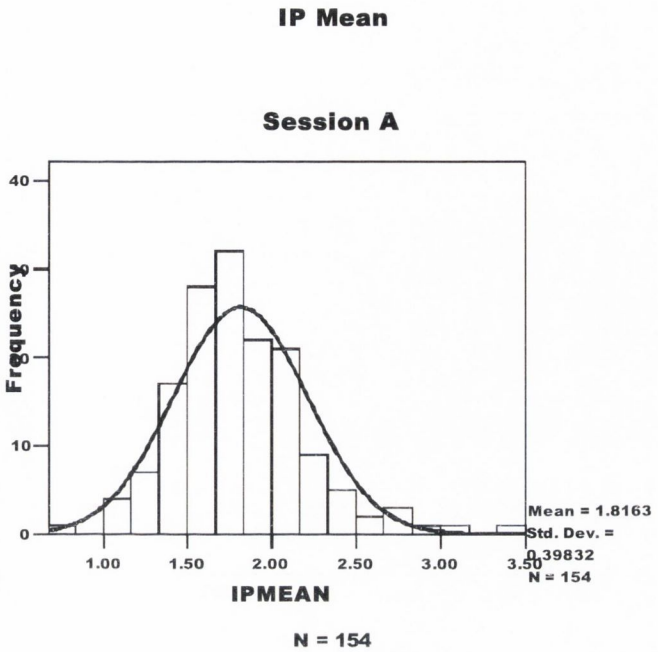

IP Mean
Session A
Frequency
40
30
20
10
0
1.00
1.50
2.00
2.50
3.00
3.50
IPMEAN
Mean = 1.8163
Std. Dev. = 0.39832
N = 154
N = 154

Fig.6.3.C.6

Appendix 12-6.5.E Survey of Potential test Item clarity

This Table presents results on the ambiguity of potential test items. The items were tested on a group of 40 IS students. The highlighted questions signal those that were removed from the potential test item pool. 120 test items were selected from the potential pool for actual use.

	+ +	+	0	-	- -
FOOTBALL and DAISY are the same Type?	0	0	0	5	35
BLACK and GREEN are the same Type?	23	11	1	1	4
CAT and LION are the same Type?	25	11	2	1	1
WHISKEY and ICECREAM are the same Type?	2	5	4	10	19
TENNIS RACKET and GUN are the same Type?	6	6	2	9	17
SPARROW and ROBIN are the same Type?	28	10	0	2	0
HOUSE and BICYCLE are the same Type?	0	2	5	10	23
CRAYONS and MARKERS are the same Type?	24	15	0	1	0
PEAR and PLUM are the same Type?	32	6	0	1	1
POTATO and TOMATO are the same Type?	6	16	4	9	5
BUTTERCUP and EGG CUP are the same Type?	0	0	0	11	29
CAR and SUBMARINE are the same Type?	4	15	5	10	6
MAN and DOG are the same TYPE?	2	9	6	11	12
JUPITER and EARTH are the same Type?	20	16	0	3	1
MISSILE and BLACKBIRD are the same Type?	0	5	4	10	21
SPADE and SWORD are the same Type?	2	12	10	8	8
COAL and WOOD are the same Type?	12	23	0	5	0
TRUCK and CAR are the same Type?	20	19	0	1	0
GUITAR and HARP are the same Type?	28	10	1	1	0
BANANA and ONION are the same Type?	4	12	4	12	8
HAMMER and SCREWDRIVER are the same Type?	15	23	1	1	0
MILK and TEA are the same Type?	9	18	4	7	2
HAZELNUTS and ALMONDS are the same Type?	27	12	0	1	0
SALMON and EAGLE are the same Type?	3	5	7	14	11
LEMON and GRAPEFRUIT are the same Type?	25	13	1	1	0
GRAPE and WINE are the same Type?	3	14	2	14	7
CARROT and PEAR are the same Type?	2	14	5	12	7
SALMON and TROUT are the same Type?	30	8	1	1	0
WATER and CHEESE are the same Type?	1	3	8	10	18
DOG and FOX are the same Type?	14	16	4	3	3
STRAWBERRY and CABBAGE are the same Type?	1	10	3	15	11
DAFFODIL and CAULIFLOWER are the same Type?	1	8	3	15	13
KITE and PARACHUTE are the same Shape?	1	13	4	8	16
NEWSPAPER and BOOK are the same Shape?	13	20	1	5	3

BASKETBALL and GLOBE are the same Shape?	23	10	2	4	3
CHAIR and STOOL are the same Shape?	7	18	1	10	6
CARROT and PARSNIP are the same Shape?	23	12	2	3	2
PITCH FORK and TRIDENT are the same Shape?	15	6	13	6	2
ARROW and DART are the same Shape?	16	17	1	8	0
ROCKET and AIRPLANE are the same Shape?	5	13	1	16	7
LION and TIGER are the same Shape?	20	17	1	2	2
SCREW and NAIL are the same Shape?	13	18	1	6	4
MOON and JUPITER are the same Shape?	20	15	1	5	1
GOLF CLUB and WALKING STICK are the same Shape?	9	14	3	12	4
MOUSE and RAT are the same Shape?	19	18	2	2	1
MARBLE and BOWLING BALL are the same Shape?	21	15	0	5	1
CIGAR and LOLLIPOP are the same Shape?	2	2	2	18	18
PAINTBRUSH and TOOTHBRUSH are the same Shape?	1	3	0	18	20
KNIFE and FORK are the same Shape?	3	0	1	11	27
RUGBY BALL and CANNONBALL are the same Shape?	1	1	0	13	27
PENCIL and BIRO are the same Shape?	23	16	0	1	2
DRINKING STRAWS and CIGERETTE are the same Shape?	15	16	0	4	7
HOT AIR BALLOON and UMBRELLA are the same Shape?	3	3	4	11	21
CHESSBOARD and DARTBOARD are the same Shape?	2	5	1	10	24
LIGHTBULB and APPLE are the same Shape?	0	6	4	12	20
SHOVAL and SPOON are the same Shape?	7	17	4	10	4
EGG and WATERMELON are the same Shape?	10	18	1	8	5
GOLF BALL and TENNIS BALL are the same Shape?	23	12	1	3	3
PEACH and PINEAPPLE are the same Shape?	2	5	1	15	19
REVOLVER and RIFLE are the same Shape?	2	8	2	15	15
DOUGHNUT and RING are the same Shape?	22	16	2	2	0
HORSE and ZEBRA are the same Shape?	27	13	0	2	0
APPLE and ORANGE are the same Shape?	16	13	1	9	3
AXE and HACKSAW are the same Shape	0	2	0	10	30
RED and WHITE are the same Type?	29	10	0	0	1
OSTRICH and EMU are the same Type?	26	13	0	1	0
MOUSE and RAT are the same Type?	28	12	0	0	0
FROG and TOAD are the same Type?	29	10	1	0	0
EAGLE and VULTURE are the same Type?	28	12	0	0	0
PEN and PENCIL are the same Type?	27	11	2	0	0
ORANGE and PEACH are the same Type?	24	16	0	0	0
CABBAGE and POTATO are the same Type?	20	15	2	3	0
SCISSORS and KNIFE are the same Type?	15	11	5	7	2
SUN and STAR are the same Type?	24	11	0	4	1
WOOD and PAPER are the same Type?	9	12	4	12	3
FERRY and CANOE are the same Type?	11	27	1	1	0
SAXOPHONE and FLUTE are the same Type?	18	19	1	1	1

SAW and DRILL are the same Type?	9	16	3	7	5
TEA and COFFEE are the same Type?	20	19	0	0	1
PORRIDGE and CEREAL are the same Type?	22	17	0	1	0
WASP and BEAR are the same Type?	2	3	2	16	17
HONEY and STONE are the same Type?	0	0	1	12	27
AEROPLANE and TIGER are the same Type?	0	0	0	9	31
SPOON and MACE are the same Type?	0	1	3	10	26
BEER AND ORANGE are the same Type?	3	9	1	10	17
GOLD and CHEESE are the same Type?	0	0	1	12	27
FLOWER and FLOUR are the same Type?	2	2	0	12	24
BICYCLE and BOAT are the same Type?	2	14	2	10	12
HUMAN and INSECT are the same Type?	0	5	4	13	18
TOOTH and CHALK are the same Type?	0	2	3	10	25
CREAM and BOWL are the same Type?	0	0	1	10	29
SUGAR and WINE are the same Type?	0	1	2	14	23
COMPUTER and BOTTLE are the same Type?	0	0	0	9	31
MONKEY and KEY are the same Type?	0	1	0	8	31
TULIP and BREAD are the same Type?	0	0	0	11	29
SWIMMING and FOOD are the same Type?	0	0	0	8	32
PLUG and SLUG are the same Shape?	0	0	4	4	34
HORSESHOE and RAINBOW are the same Shape?	13	19	2	0	8
FOOTBALL and PLANET are the same Shape?	19	18	0	3	2
GUN and BULLET are the same Shape?	0	2	0	13	27
RING and HOOLA-HOOP are the same Shape?	24	14	1	2	1
BED and ARMCHAIR are the same Shape?	0	5	1	10	26
ONION and GARLIC are the same Shape?	11	15	3	10	3
SICKLE and CHAINSAW are the same Shape?	0	2	8	9	23
NEWSPAPER and MAGAZINE are the same Shape?	18	19	0	5	0
FOX and WOLF are the same Shape?	16	20	1	5	0
SCARF and BANDAGE are the same Shape?	14	14	4	6	4
AEROPLANE and HELICOPTER are the same Shape?	0	4	0	14	24
PLATE and FRISBEE are the same Shape?	22	16	1	1	2
SHIRT and BLOUSE are the same Shape?	21	15	0	4	2
BOTTLE and BOOK are the same Shape?	1	1	0	7	33
FROG and TOAD are the same Shape?	22	15	1	1	3
CABBAGE and LETTUCE are the same Shape?	22	15	1	2	2
MELON and HAMMER are the same Shape?	1	1	0	6	34
SUN and MOON are the same Shape?	24	13	1	2	2
SPOON and CUP are the same Shape?	0	1	0	9	32
CD and TAPE are the same Shape?	0	0	0	6	36
TELEVISION and SOFA are the same Shape?	0	0	0	7	35
MONKEY and APE are the same Shape?	19	19	0	1	3
SUBMARINE and CAR are the same Shape?	0	1	1	7	33

SWORD and DAGGER are the same Shape?	19	18	0	2	3
DOOR and KEY are the same Shape?	0	1	0	5	36
CHICKEN and EGG are the same Shape?	0	0	0	3	39
BANANA and ORANGE are the same Shape?	1	2	0	3	36
PEN and PENCIL are the same Shape?	25	13	2	0	2
PIPE and CIGARETTE are the same Shape?	4	1	1	8	28
COCONUT and EGG are the same Shape?	13	16	5	4	4
TREE and AXE are the same Shape?	1	1	0	5	35

Appendix 13-6.6.F - Summary Statistics for the VItrap Test Session B VItrap (Retest) and Session A (Test)

Summaries
No Selector

Mean	Median	StdDev	Range	StdErr
1.08499	1.0627	0.166352	1.21578	0.013405
1.1524	1.14061	0.119843	0.721356	0.00965721
1.10797	1.09018	0.123252	0.873656	0.00993194
2.64787	2.56837	0.698072	3.624	0.0562522
2.44817	2.41705	0.558144	2.72687	0.0449765
1.70251	1.63812	0.434704	2.78997	0.0350294
1.47967	1.42765	0.350375	2.59403	0.028234
2.22745	2.17575	0.540477	3.385	0.0435529
2.06668	2.013	0.451441	2.6185	0.0363782
1.45679	1.4145	0.287959	1.6225	0.0232044
1.30271	1.272	0.244349	1.462	0.0196902

Vq/Iq
Vp/Ip
Vq+Vp/Iq+Ip

Vq mean
Iq mean
Vp mean
Ip mean

VQ median
IQ median
VP median
IP median

- order of Summary variables as given in above table.

Summaries
No Selector

Mean	Median	StdDev	Range	StdErr
1.03244	1.01257	0.150719	0.982071	0.0121452
1.20098	1.18424	0.148857	0.856999	0.0119952
1.09142	1.07034	0.12139	0.867017	0.00978191
3.2064	3.07343	0.858	4.2929	0.0691396
3.12097	3.00112	0.780145	5.13893	0.0628659
2.17724	2.12402	0.528991	2.96107	0.0426273
1.81625	1.7785	0.39832	2.5259	0.0320975
2.79142	2.644	0.718196	3.8805	0.0578739
2.71263	2.60875	0.678157	4.1955	0.0546474
1.85367	1.82275	0.404026	2.97	0.0325573
1.58807	1.567	0.300829	1.7875	0.0242415

Session A - VItrap

Appendix 14-6.7.G Post Hoc Test Results for Variables which affect Test Item response Times - Sessions A and B

Session A (Test Results).

A linear repeated measures model being applied to investigate the effect of incorrect/ correct questions on subsequent inspection times to questions. Another factor examined by this model is the effect of question type on subsequent inspection times. Where necessary post hoc tests are included to investigate a significant p value in the general analysis of variance table. For these Tables the dependent variable is the log of inspection. (N.B. for these tables the labels inspection, response time and reaction were used equally).

Table 6.7.G.1 Session A: Post hoc tests for effect of Question (correct /incorrect) on log of inspection times in Session A.

General Results

18480 total cases

ANOVA

Results for factor LCt

Coefficients

Coefficients of: LogTSA on LCt

Level of LCt	Coefficient	std. err.	t Ratio	prob
0	0.09539	0.004299	22.19	≤ 0.0001
1	-0.09539	0.004299	-22.19	≤ 0.0001

Expected Cell Means

Scheffe Post Hoc Tests

	Difference	std. err.	Prob
1 - 0	-0.190782	0.008599	0

Table 6.7.G.2 Results for question type: on the effect of question type on log of inspection times in Session A.

General Results

Results for factor QTy

Coefficients

Coefficients of: LogTSA on QTy

Level of QTy	Coefficient	std. err.	t Ratio	prob
IP	-0.1294	0.002118	-61.12	≤ 0.0001
IQ	0.09124	0.002116	43.12	≤ 0.0001
VP	-0.05924	0.002119	-27.95	≤ 0.0001
VQ	0.09743	0.002124	45.87	≤ 0.0001

Expected Cell Means

Scheffe Post Hoc Tests

Session B (Retest Results).

A linear repeated measures model being applied to investigate the effect of incorrect/correct questions on subsequent inspection times to questions. Another factor examined by this model is the effect of question type on subsequent inspection times. Where necessary post hoc tests are included to investigate a significant p value in the general analysis of variance table.

Table 6.7.G.3 Results for question type: on the effect of question type on log of inspection times in Session B.

DESIGN

Dependent variables

Factors

No Modifications

RESULTS

General Results

18480 total cases

ANOVA

Analysis of Variance For LTimeSB
No Selector

Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Const	1	1219.9	1219.9	46037	≤ 0.0001
QTy	3	172.86	57.62	2174.5	≤ 0.0001
Prn	153	122.838	0.802864	30.299	≤ 0.0001
LCB	1	9.34717	9.34717	352.75	≤ 0.0001
Error	18322	485.502	0.0264983		
Total	18479	796.944			

Results for factor QTy

Table 6.7.G.4 Post Hoc Tests on Incorrect/ Correct Responses on Log of Time in Session B.

DESIGN

Dependent variables

Factors

No Modifications

RESULTS

General Results

Results for factor LCB

Coefficients

Coefficients of: LTimeSB on LCB

Level of LCB	Coefficient	std. err.	t Ratio	prob
0	0.09032	0.004809	18.78	≤ 0.0001
1	-0.09032	0.004809	-18.78	≤ 0.0001

Expected Cell Means

Scheffe Post Hoc Tests

Table 6.7.G.5 Post hoc Tests for Effect of Question Types on log of inspection times in Session B.

DESIGN

Dependent variables

Factors

No Modifications

RESULTS

General Results

Results for factor QTy

Coefficients

Coefficients of: LTimeSB on QTy

Level of QTy	Coefficient	std. err.	t Ratio	prob
IP	-0.122	0.002075	-58.78	≤ 0.0001
IQ	0.07894	0.002075	38.04	≤ 0.0001
VP	-0.06655	0.002076	-32.05	≤ 0.0001
VQ	0.1096	0.002076	52.79	≤ 0.0001

Expected Cell Means

Scheffe Post Hoc Tests

Repeated measures means that the same subject is being measured multiple times. Thus, in the analysis explicit account must be taken of the subject. This is done in the ANOVA by putting the PrN (Person) into the model - this is a factor which accounts for the average effect for each person. Such an analysis is called a repeated measures design.

From Table 6.7.G.3 we note that the statistically significant variables on the repeated measures analysis are Qty, Lcb and Prn. Thus, whether a subject gets the last question correct has an effect on the next question response time. Similarly the Question Type has an impact on average response time.

In terms of the size of this effect, since the response is the log of time, the coefficient represents the proportionate delay (gain) for a given factor. To get this, one needs to exponentiate the coefficient (and for a confidence interval the coefficient +- 1.96*standard error). For example, for IP exp(-0.1294) = 0.88, i.e. 88% with the 95%CI of (0.875,0.882).

So, if you had the last question wrong, the time takes 20% longer - specifically it takes 121% of the average - thus 1.21 - 95%CI (1.19,1.23).

Appendix 15-7.1.A Vividness of Visual Imagery Questionnaire

(David Marks (1970) Middlesex University)

In this test, think of the numbers you give yourself as ranks – first place, second place etc. In other words, 1 = good and 5 = poor. The lower your total score, the better your visual imagery. You are rating your visual images in terms of how well they correspond to 'mental pictures', how clearly you are 'seeing in the mind's eye'.

First, get this scoring system clear in your head:

1 The image is perfectly clear and as vivid as normal vision

2 The image is clear and reasonably vivid

3 The image is moderately clear and vivid

4 The image is vague and dim

5 There is no image at all; you only 'know' that you are thinking of an object

Remember, 1 is best imagery, 5 is worst.

For the first 4 questions, think of some relative or friend whom you frequently see (but who is not with you at present) and consider carefully the picture that comes before your mind's eye.

Name: _____ Date: _____ Class: _____

scene rating (1 to 5)

1 The exact contour of face, head, shoulders and body. _____

2 Characteristic poses of head, attitudes of body, etc. _____

3 The precise carriage, length of step, etc. in walking. _____

4 The different colours worn in some familiar clothes. _____

Next, visualise the rising sun. Consider carefully the picture that comes before your mind's eye.

5 The sun rising above the horizon into a hazy sky. _____

6 The sky clears and surrounds the sun with blueness. _____

7 Clouds. A storm blows up with flashes of lightning. _____

8 A rainbow appears. _____

Think of the front of a shop which you often go to. Consider the picture that comes before your mind's eye.

9 The overall appearance of the shop from the opposite side of the road. _____

10 A window display including colours, shapes and details of individual items for sale. _____

11 You are near the entrance. The colour, shape and details of the door. _____

12 You enter the shop and go to the counter. The counter assistant serves you. Money changes hands. _____

Finally, think of a country scene which involves trees, mountains and a lake. Consider the picture that comes before your mind's eye.

13 The contours of the landscape. _____

14 The colour and shape of the trees. _____

15 The colour and shape of the lake.

16 A strong wind blows on the trees and on the lake, causing waves. _____

Appendix 16-7.2.B Summary Statistics For Felder's Scale

Summary of **ACTREF**
No Selector

Mean	5.74854
Median	6
StdDev	2.12297
Range	9
StdErr	0.162347

Summary of **SEQGLO**
No Selector

Mean	6.44444
Median	6
StdDev	1.90699
Range	9
StdErr	0.145831

Summary of **VISVRB**
No Selector

Mean	4.4152
Median	4
StdDev	2.17678
Range	10
StdErr	0.166462

Summary of **SENINT**
No Selector

Mean	6.38596
Median	6
StdDev	2.55588
Range	11
StdErr	0.195453

Felder's - questionnaire data

Summary Statistics for the Mill Hill Inventory IQ Test

Summary of **millhill**
No Selector
182 total cases of which 4 are missing

Mean	40.6292
Median	40.5
StdDev	6.98486
Range	37
StdErr	0.523538

STAI Summary Statistical Data

Summary of **stat**
No Selector

Mean	34.8362
Median	34
StdDev	8.88923
Range	41
StdErr	0.666276

Summary of **tra**
No Selector

Mean	40.4859
Median	39
StdDev	9.07726
Range	53
StdErr	0.68037

APM Summary Statistical data

Summary of **APM2**
No Selector
175 total cases of which 1 is missing

Mean	28.7299
Median	29
StdDev	4.68937
Range	27
StdErr	0.3555

Statistical Data For Study n = 154

Summary of **age**
No Selector

Mean	20.6092
Median	20
StdDev	2.1648
Range	14
StdErr	0.163643

Statistical Data for Eysenck's Three Subscales

Summary of **E**
No Selector

Mean	13.2737
Median	14
StdDev	4.23972
Range	25
StdErr	0.316891

Summary of **L**
No Selector

Mean	2.11173
Median	2
StdDev	1.61413
Range	8
StdErr	0.120646

Summary of **N**
No Selector

Mean	10.8883
Median	11
StdDev	4.57753
Range	21
StdErr	0.34214

Eysenck Personality Data

Appendix 17-8.1.A Sample EEG Scans from the SKIL Analysis Software

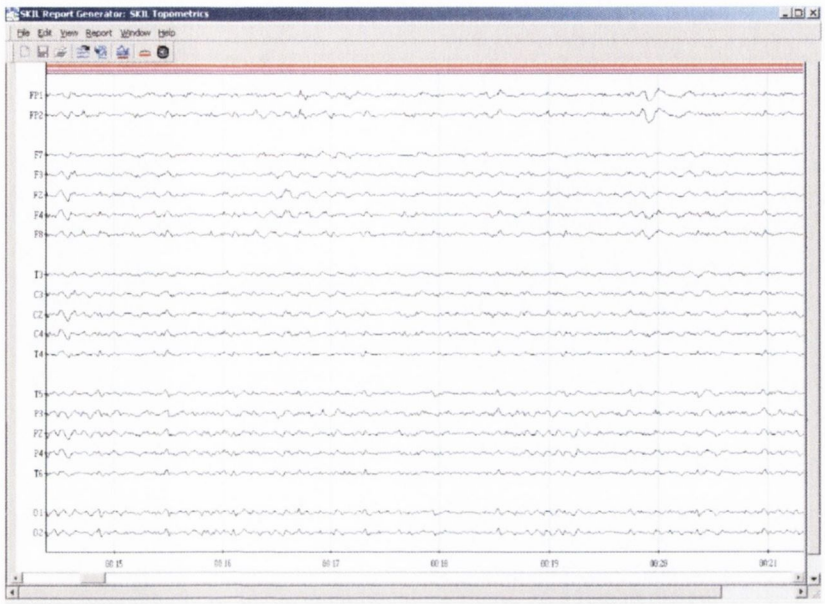


Fig 8.1.A Eyes Closed (EC) EEG Recording
Note the eye movement artifact at FP2 (node) at time 00.30

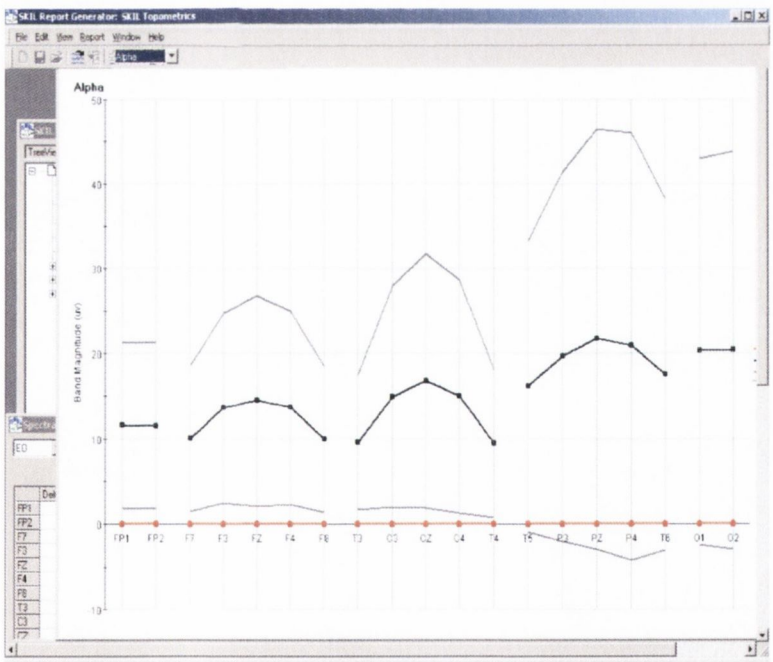


Fig 8.2.B Eyes Closed (EC) EEG Recording Referenced to SKIL database (Mid-group represents mean value of subject's recording. Other curves represent +/- 3 s.d.'s

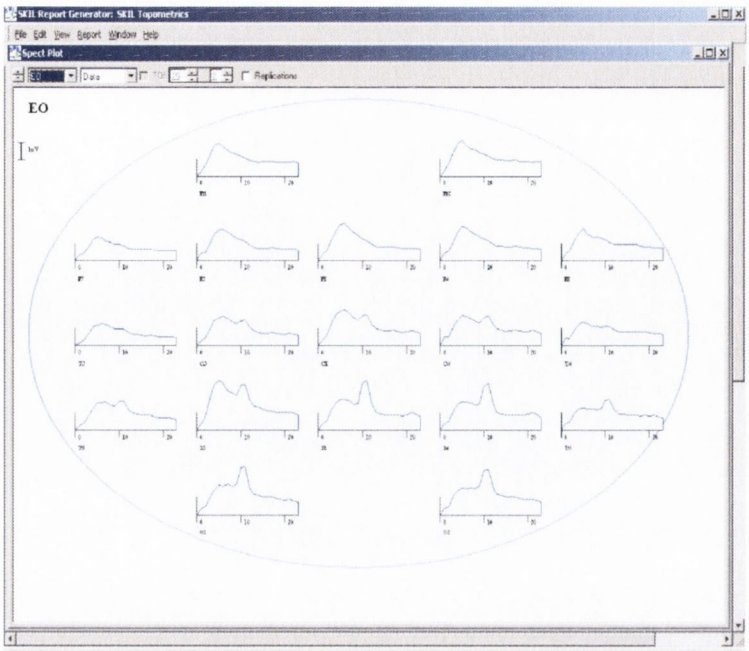


Fig 8.3.C Topographic Map - Eyes Closed (EC) EEG Recording

Note the asymmetry!

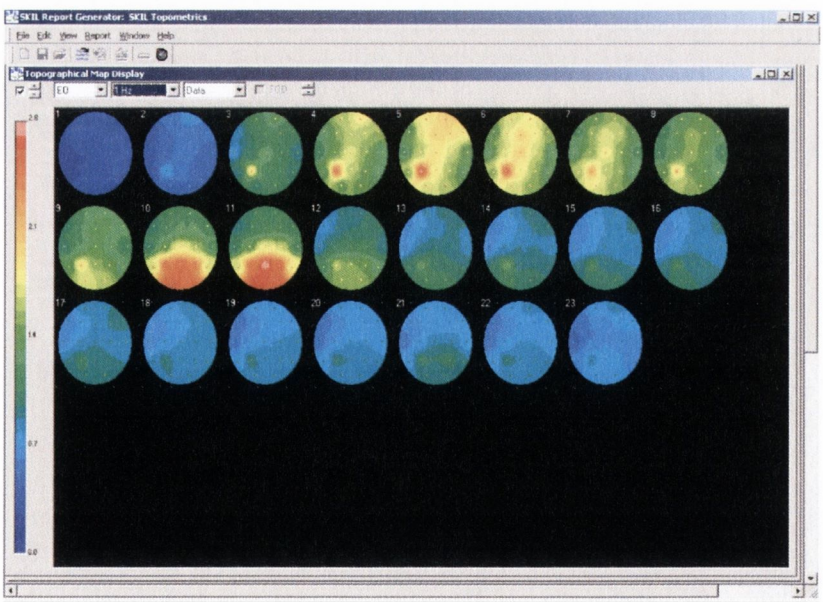


Fig 8.5.E Spectral Plot - Eyes Closed (EC) EEG Recording

Appendix 18-9.1.A Usability histograms pertaining to Study 9.1

The following analysis of the Quis relates to Study 9.1. (n = 168).

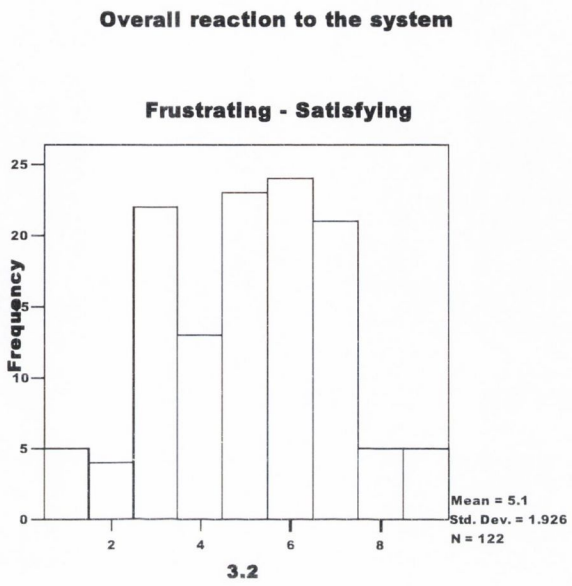


Fig. 9.1.11.1 seems to be evenly spread and symmetrical?

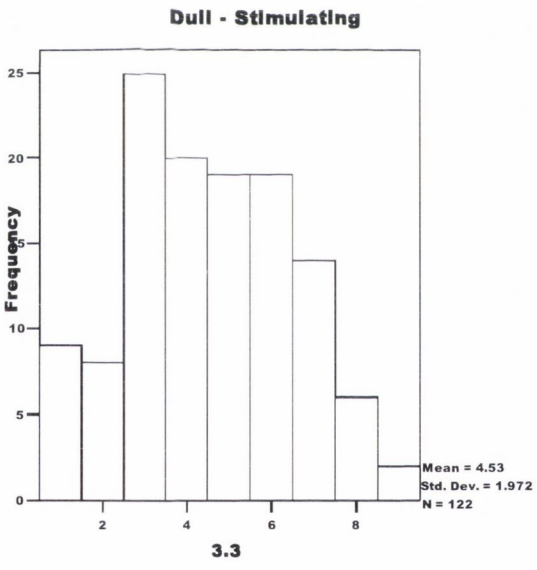


Fig. 9.1.11.2 indicates that the users found the system somewhat monotonous. This might be due to the number of repetitions.

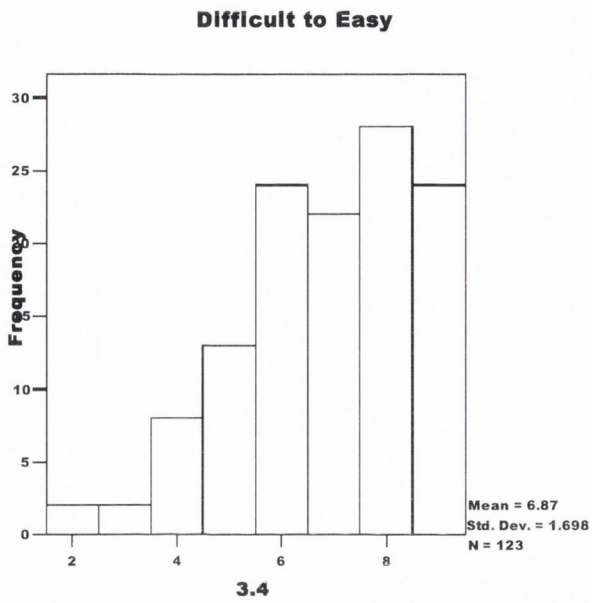


Fig. 9.1.11.3 This indicates that the application was easy to use

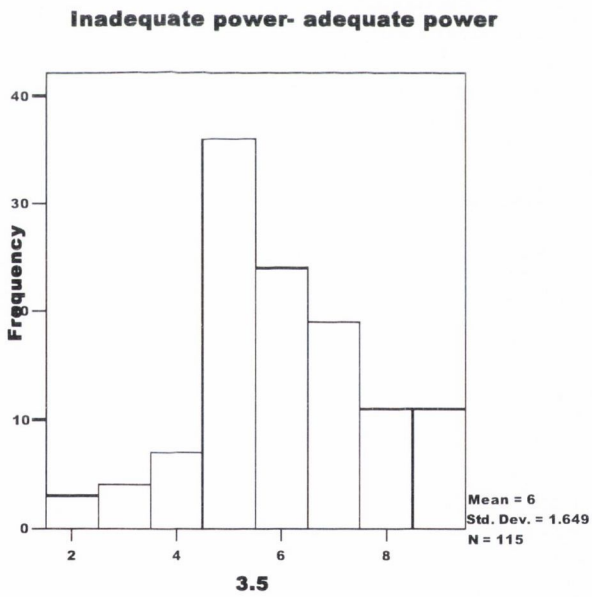


Fig. 9.1.11.4 difficult to interpret adequately - somewhat ambiguous.

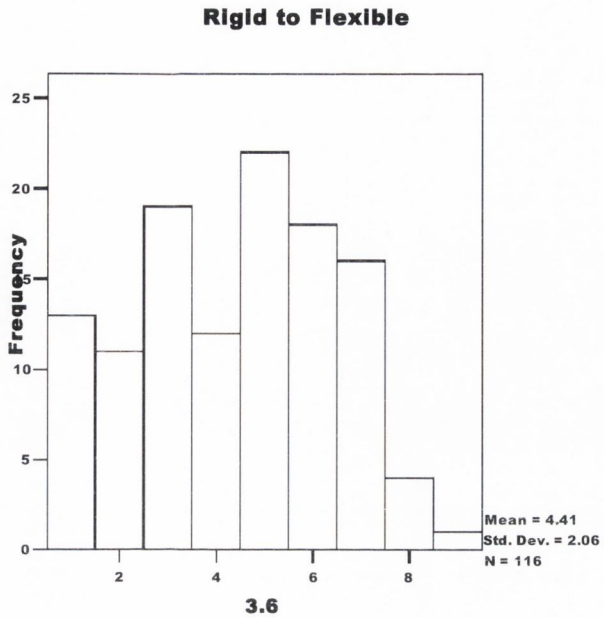


Fig. 9.1.11.5 The experimental nature of the application interface was somewhat restricted.

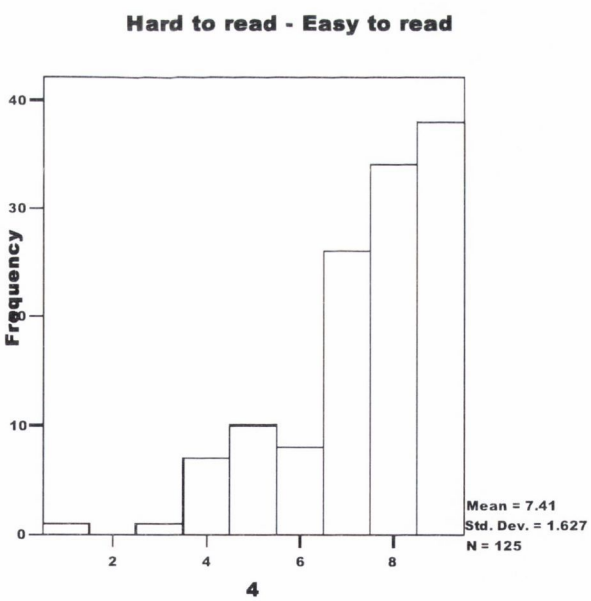


Fig. 9.1.11.6 The application interface was easy to read.

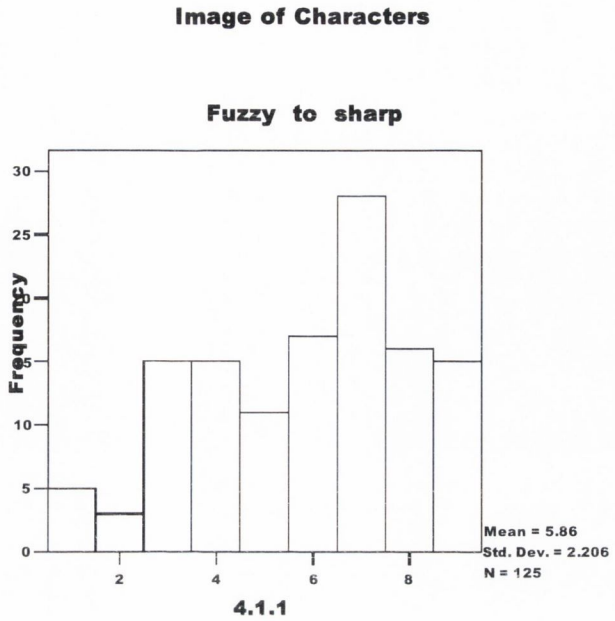


Fig. 9.1.11.7 The quality of images varied somewhat.

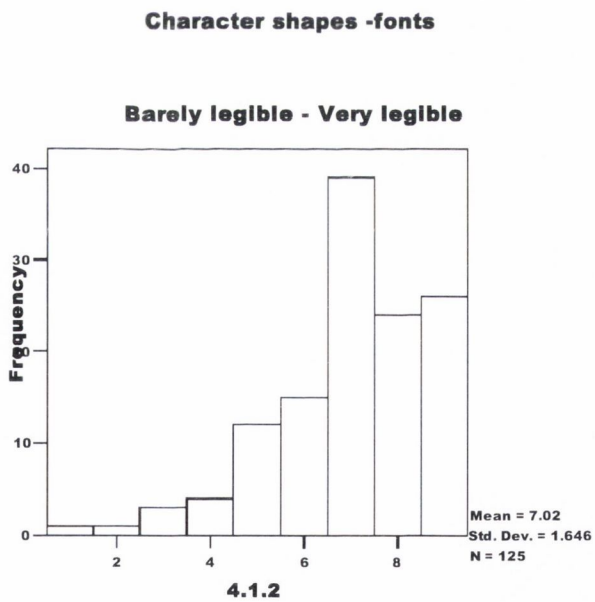


Fig. 9.1.11.8 The quality of fonts was easy to read.

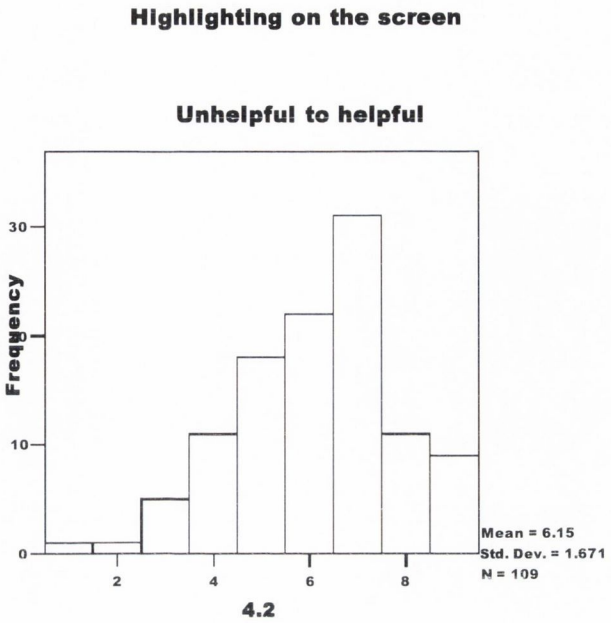


Fig. 9.1.11.9 The highlighting on the screen was helpful

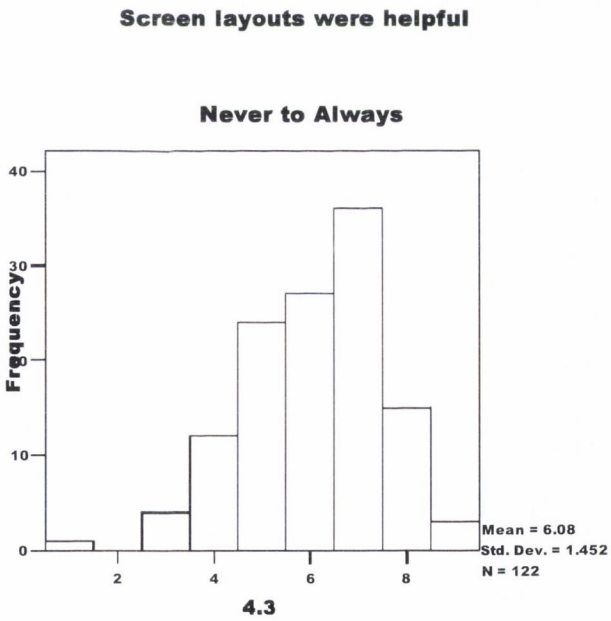


Fig. 9.1.11.10 The screen layouts were also helpful.

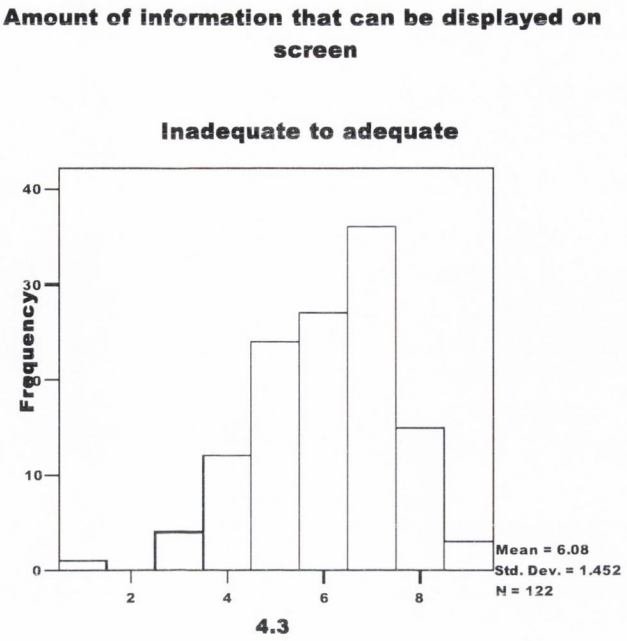


Fig. 9.1.11.11 The amount of information displayed was adequate.

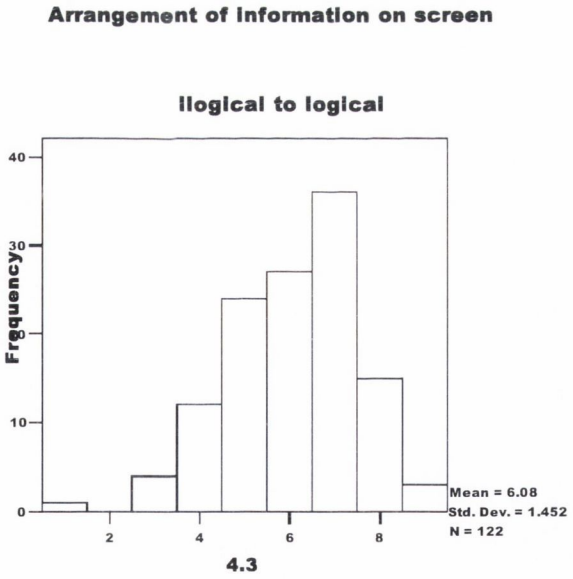


Fig. 9.1.11.12 The arrangement of information on the screen was logically rational.

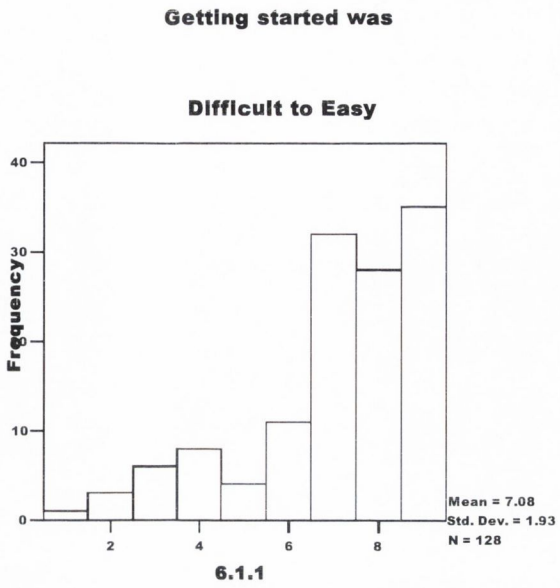


Fig. 9.1.11.13 Getting started was very easy.

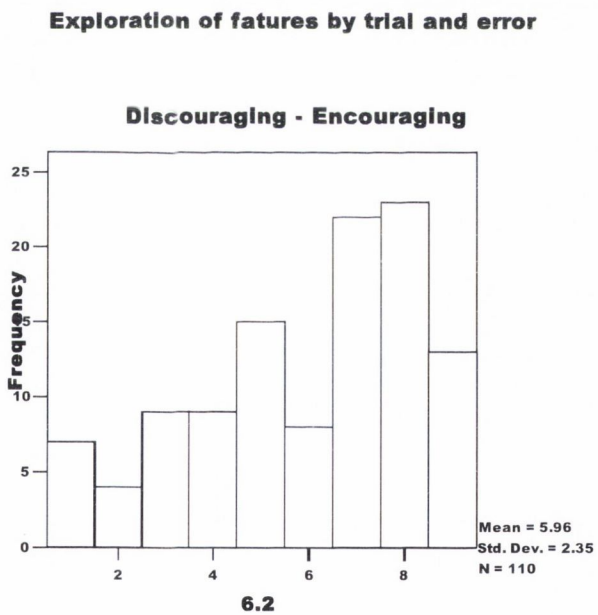


Fig. 9.1.11.14 The exploration of features of the system by trial and error was encouraging

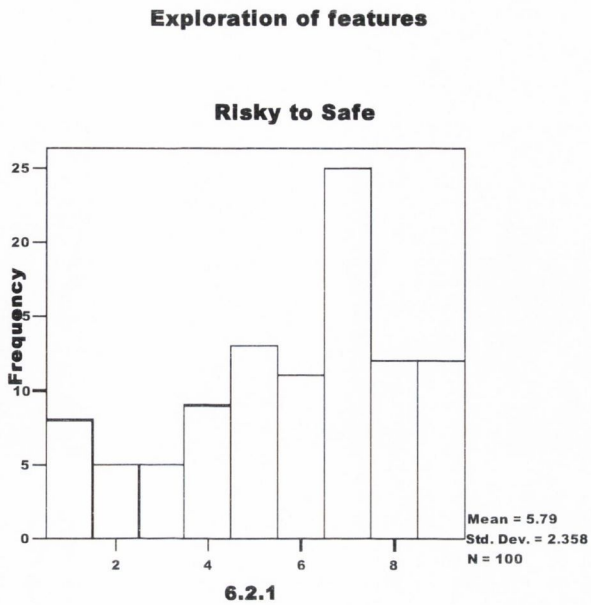


Fig. 9.1.11.15 The exploration of features was generally safe.

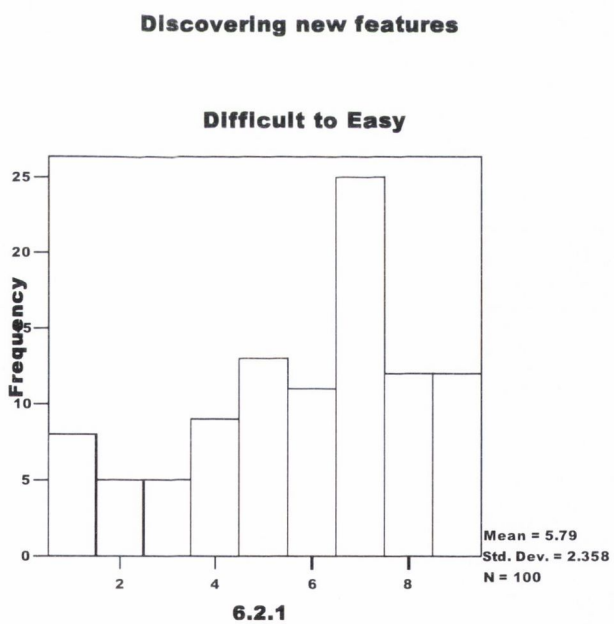


Fig. 9.1.11.16 Discovering new features was relatively easy.

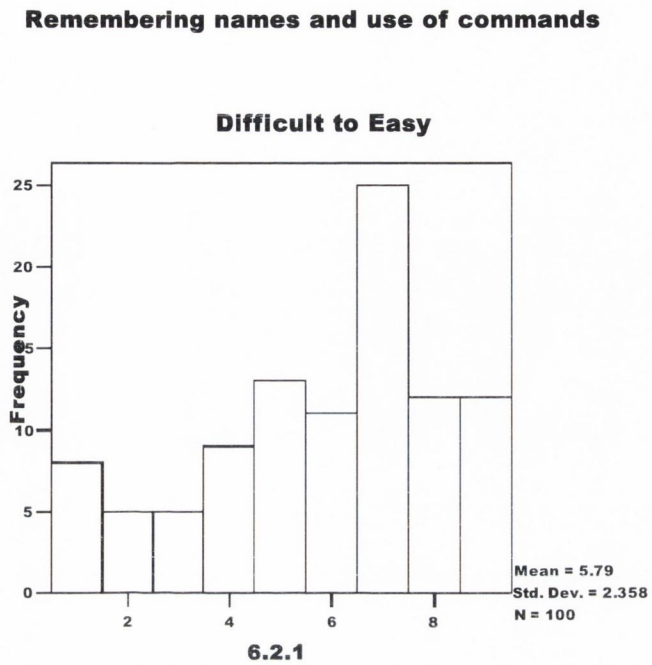


Fig. 9.1.11.17 Remembering names and use of commands was relatively easy.

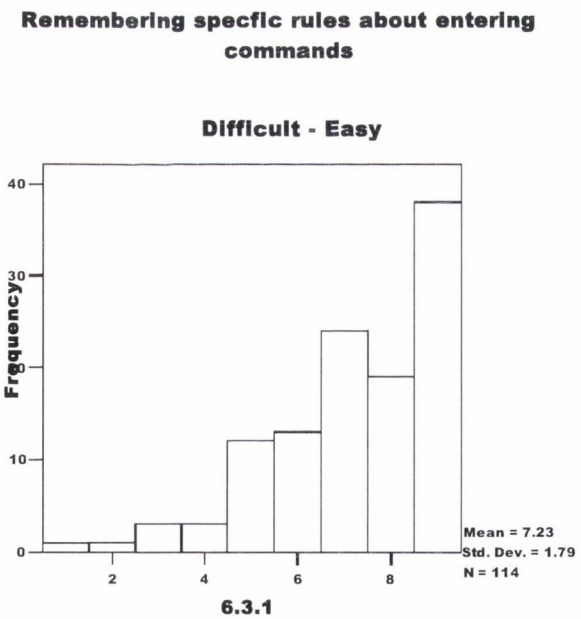


Fig. 9.1.11.18 Remembering specific rules for entering commands was very easy.

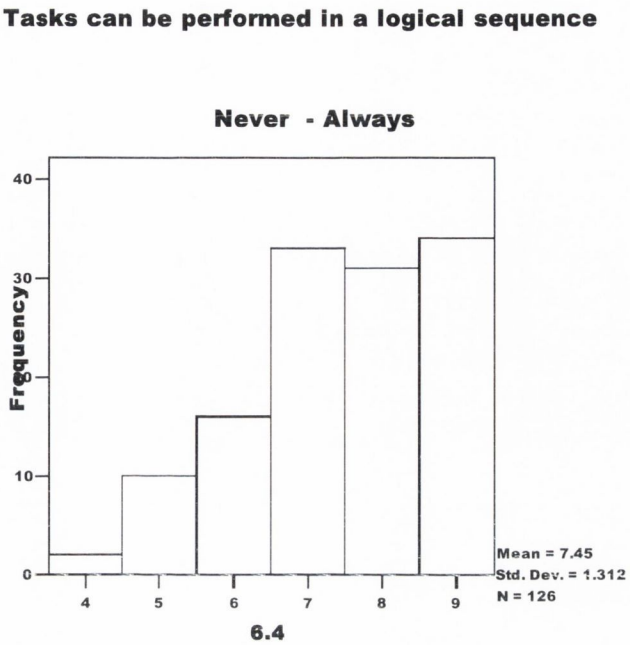


Fig. 9.1.11.19 Tasks could be performed in an adequate logical sequence.

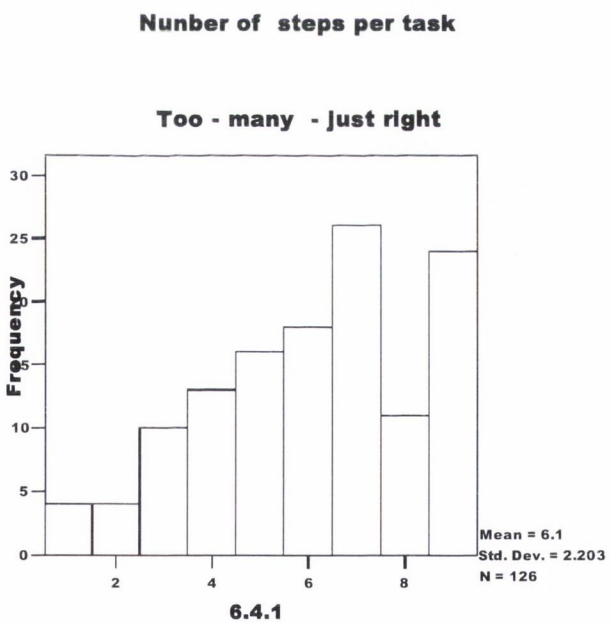


Fig. 9.1.11.20 The number of steps per task was just about right.

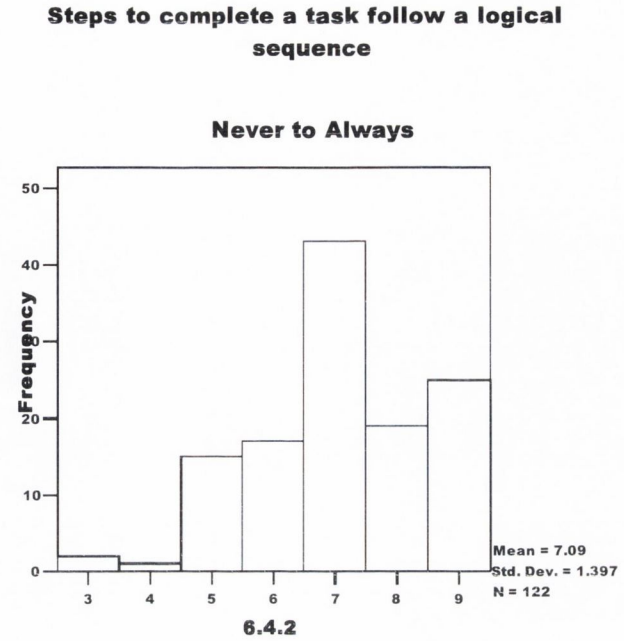


Fig. 9.1.11.21 Steps to complete a task were usually presented in a rational sequence

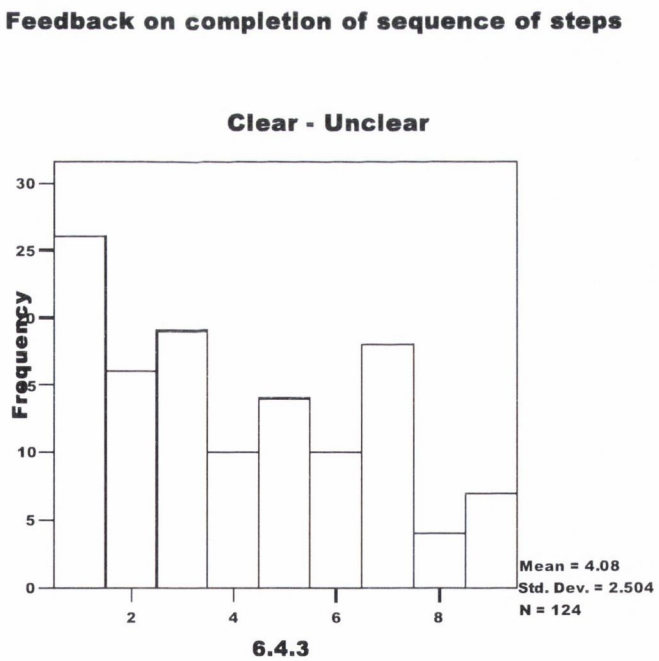


Fig. 9.1.11.22 The feedback on the completion of a sequence of steps could have been better.

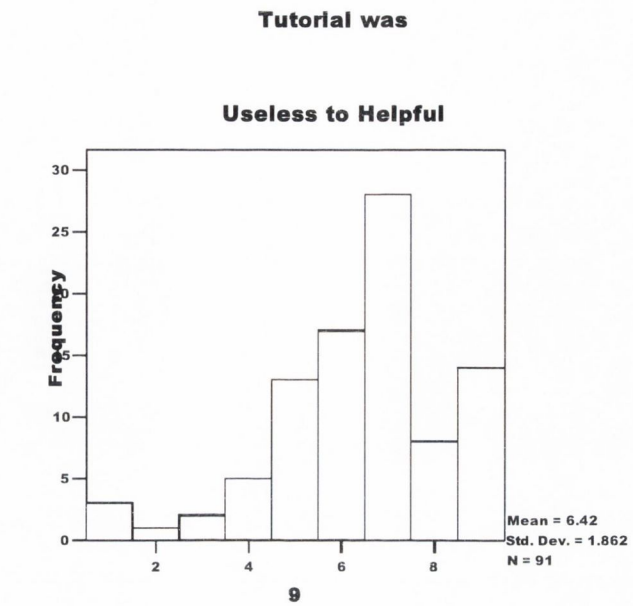


Fig. 9.1.11.23 The Tutorial was helpful,

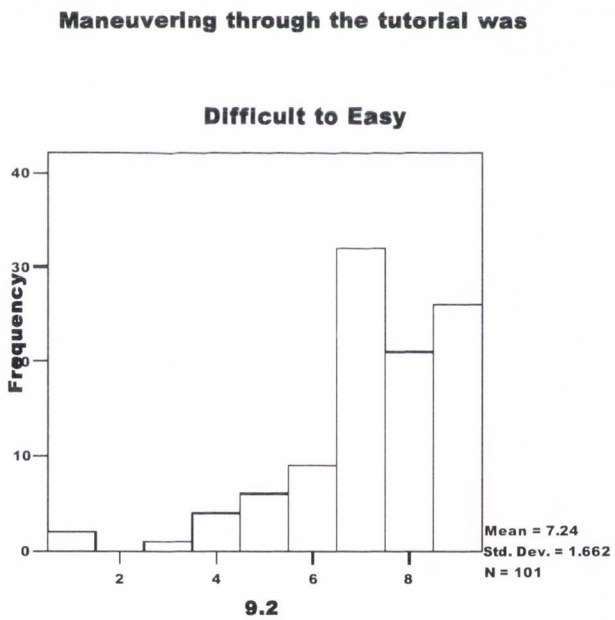


Fig. 9.1.11.24 Manouevring through the Tutorial was easy.

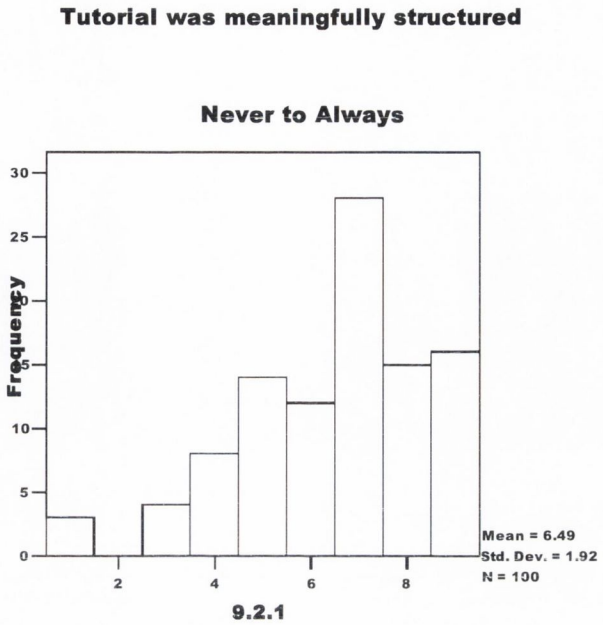


Fig. 9.1.11.25 The Tutorial was meaningfully structured.

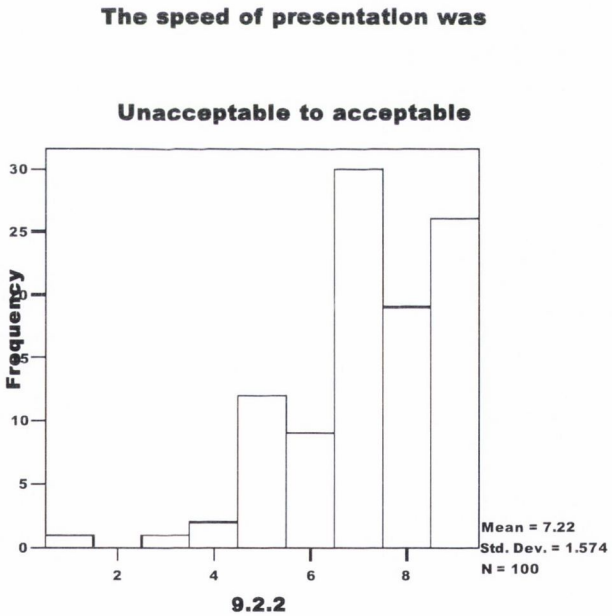


Fig. 9.1.11.26 The speed of presentation was acceptable.

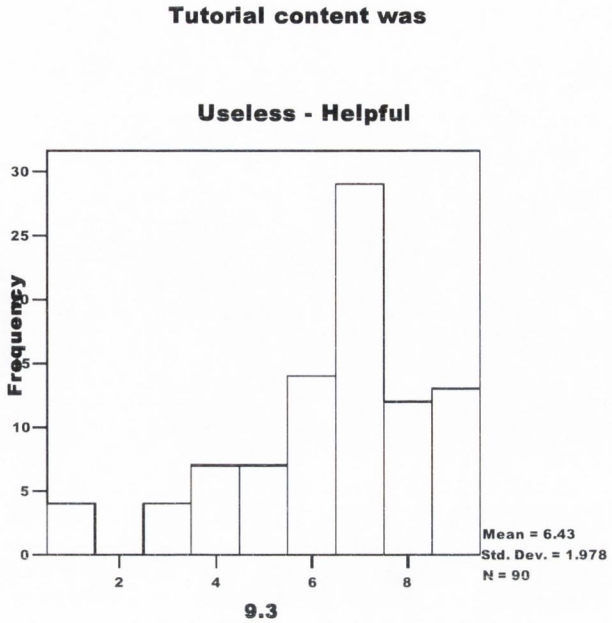


Fig. 9.1.11.27 The Tutorial Content was helpful.

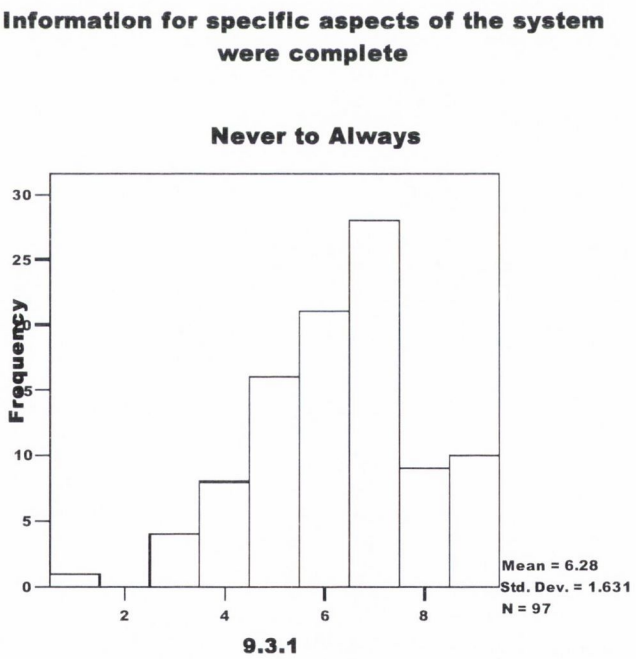


Fig. 9.1.11.28 Information for specific aspects of the system were generally complete.

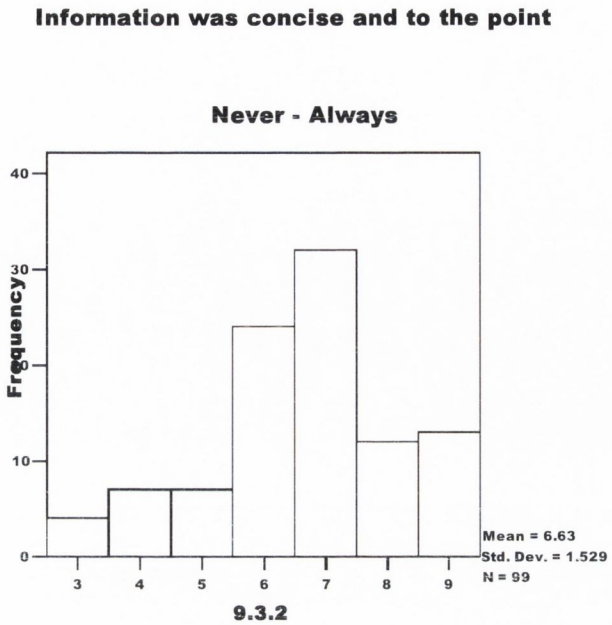


Fig. 9.1.11.29 Information was usually concise and to the point.

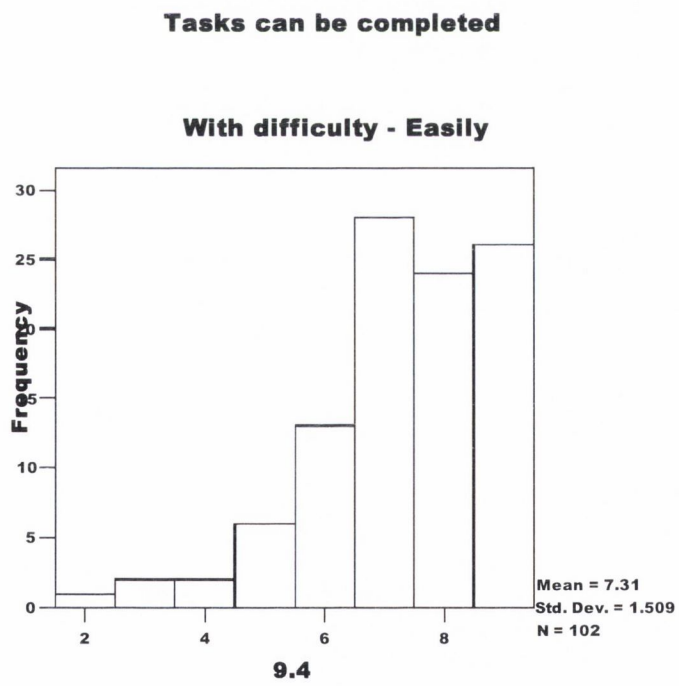


Fig. 9.1.11.30 Tasks were completed easily.

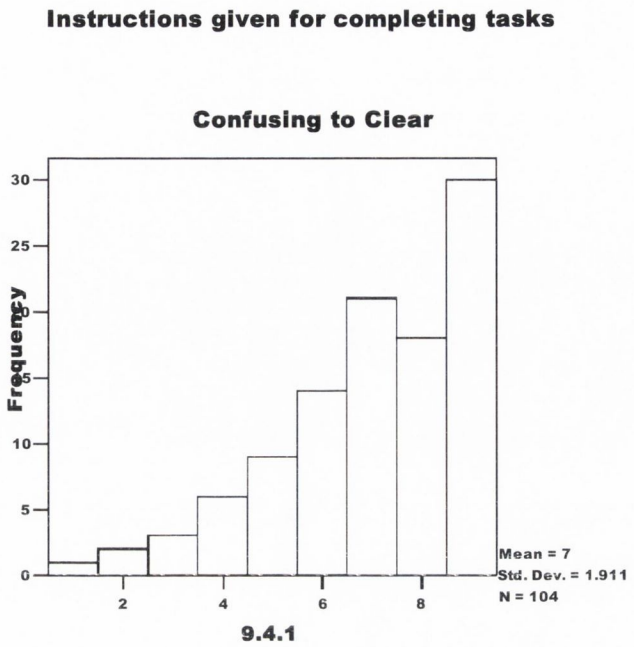


Fig. 9.1.11.31 The instructions for completing tasks were clear.

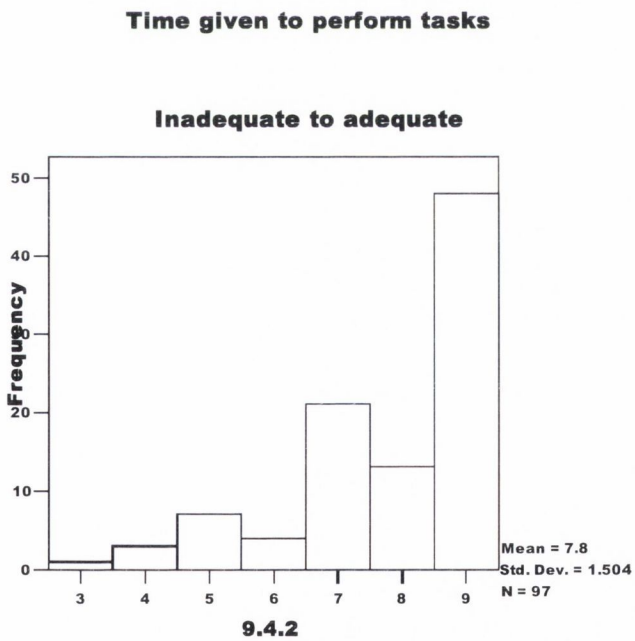


Fig. 9.1.11.32 Time given to perform tasks was adequate.

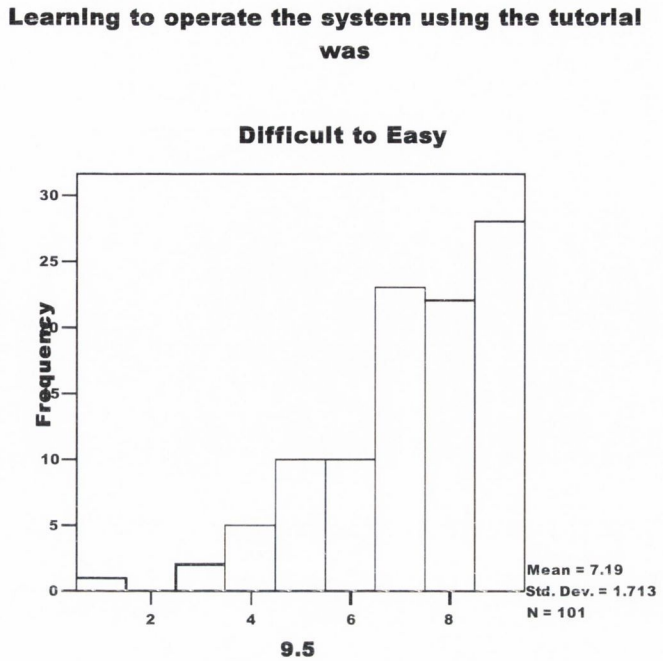


Fig. 9.1.11.33 Learning to operate the system via the Tutorial was relatively easy.

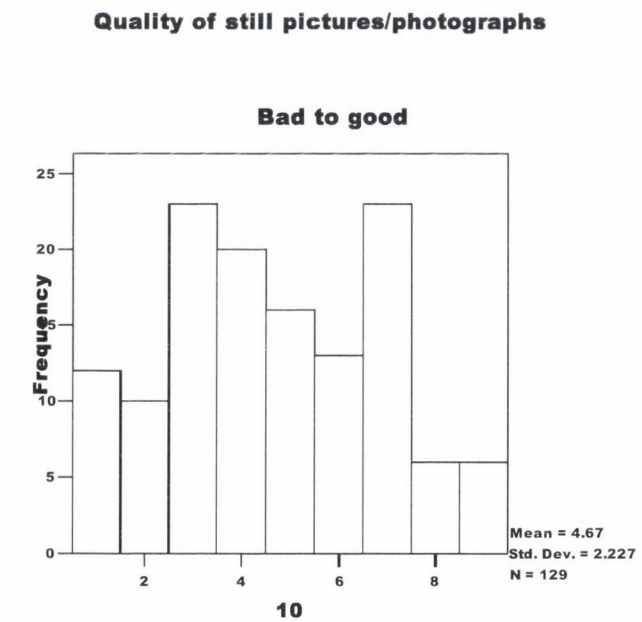


Fig. 9.1.11.34 The quality of still photographs could be better.

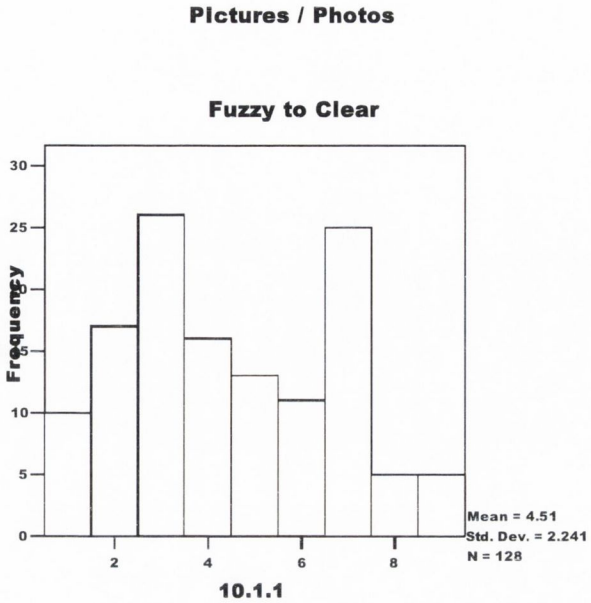


Fig. 9.1.11.35 The photos could have been clearer.

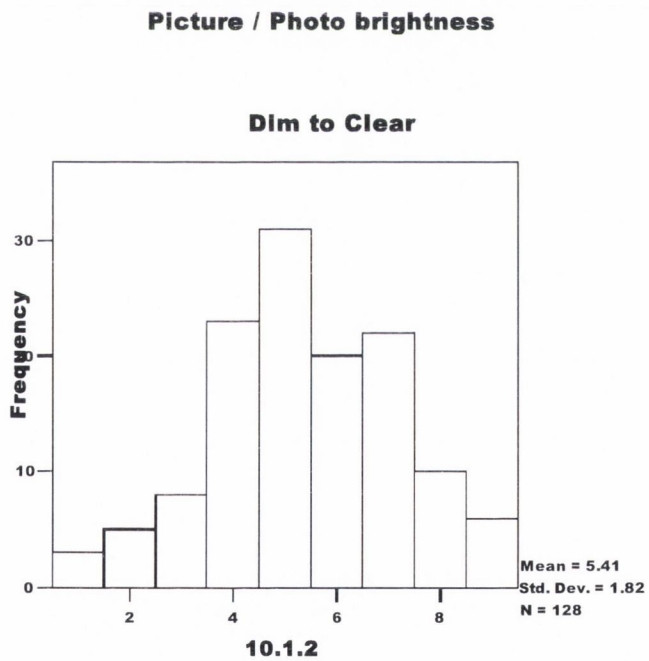


Fig. 9.1.11.36 Photo brightness could have been clearer.

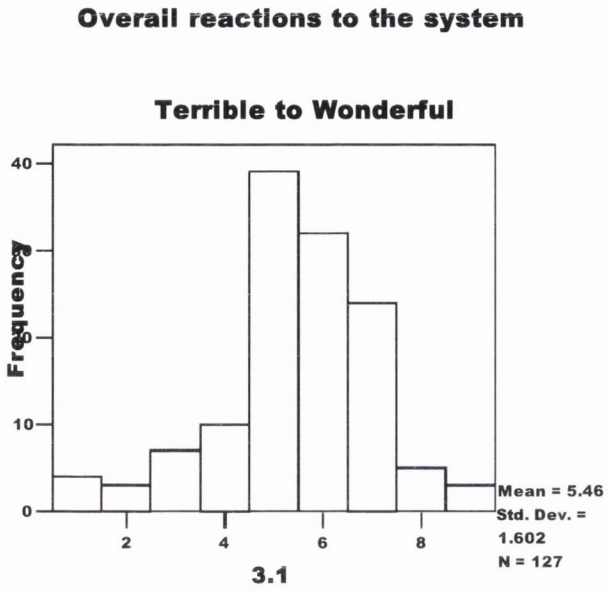


Fig. 9.1.11.37 Overall reaction to the system was average,

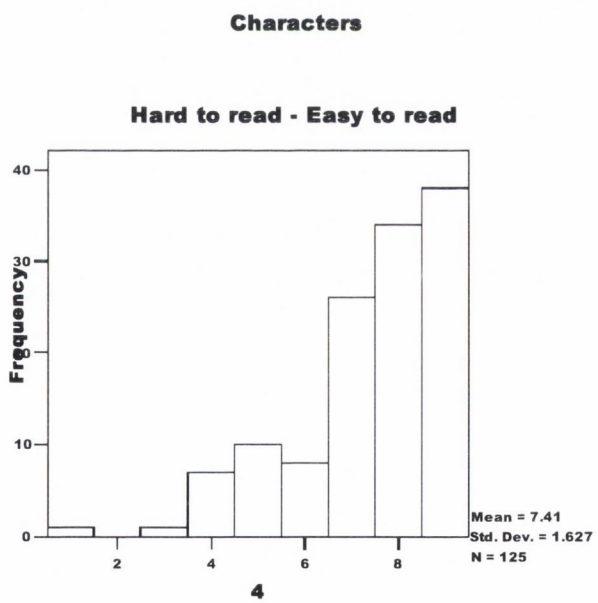


Fig. 9.1.11.38 The characters were very easy to read.

Appendix 19-9.2.B Golf Applications 1 (Ultan)

9.2.1 Study 9.4 Third Application - Golfing Stroke Saver on Cd-Rom

9.2.2 Introduction - Study 9.4:

The third application developed was a golfing simulator essentially an advanced stroke saver. It was developed using Authorware (2006).

9.2.3 Design of Golf Application 1 (Ultan)

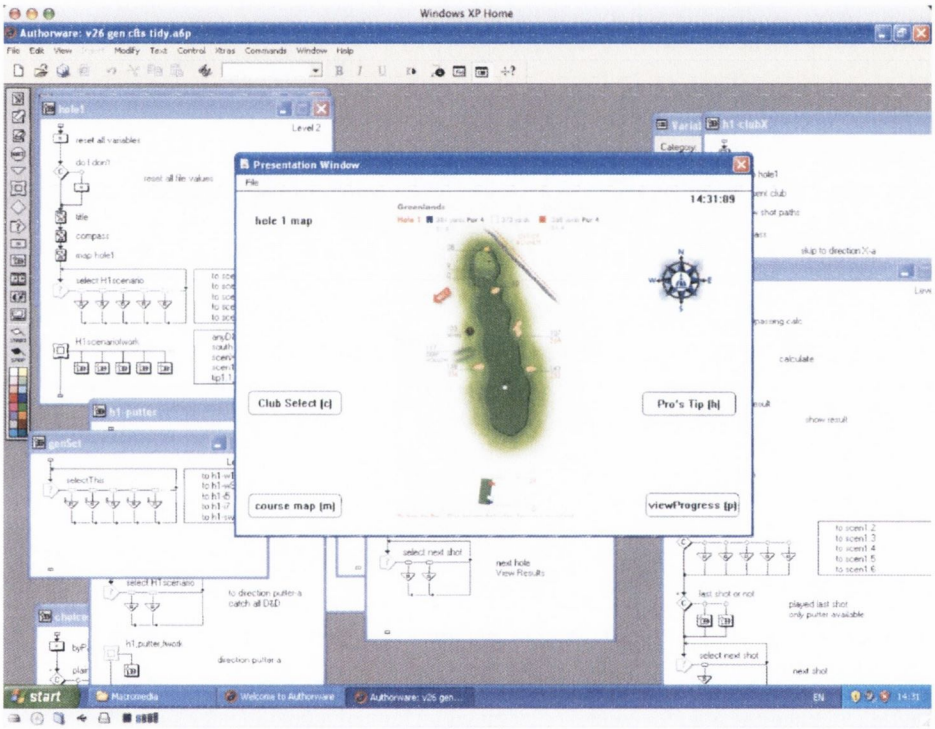


Fig. 9.2.1 Example screenshot of the Golf Simulator in Authorware

9.2.4 Results



Fig. 9.2 Imported image of rosse's Point golf links

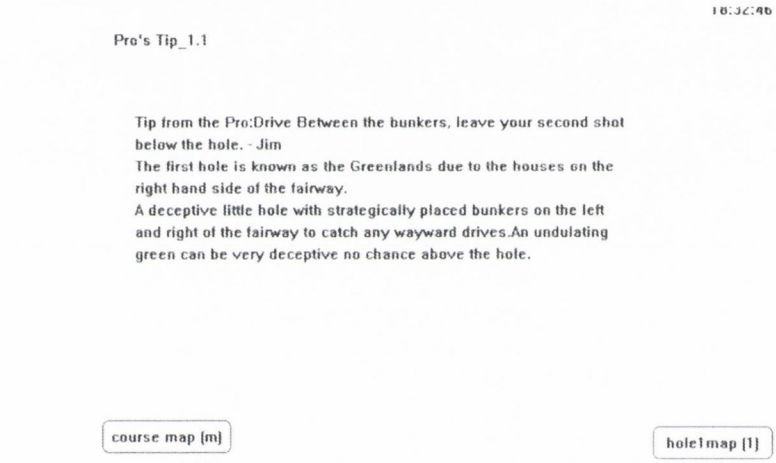


Fig. 9.3 Sample Pro's tip



Fig. 9.4 Map of Hole 1 in graphic mode

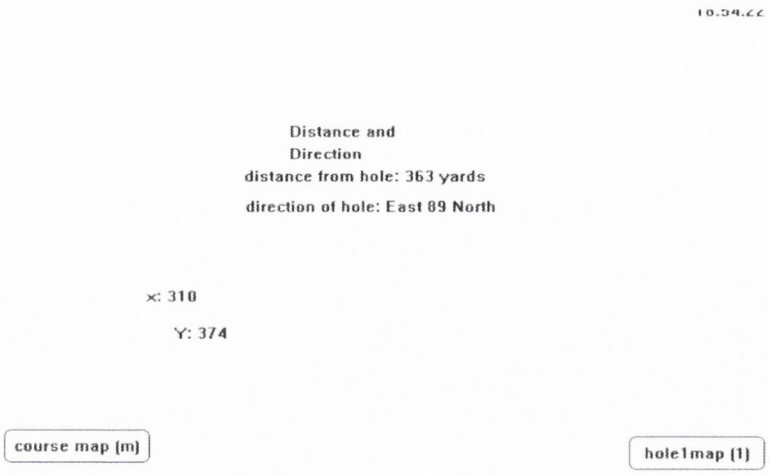


Fig. 9.5 Distance Guidance for Hole 1



Fig. 9.6 Picture of Hole 1 at Rosse's Point



Fig. 9.7 Picture of Hole 1 at Rosse's Point

Appendix 20-9.3.C Golf Application 2 (Piotr)

9.3.1 Study 9.5 Application 4 – Golf Stroke Saver on PDA/ Web

9.3.2 Introduction - Study 9.5

This was similar application to the previous one except it was for a Personal Digital Assistant (PDA) attached to a golf cart and developed using different technologies (Tomcat server, XML) to the previous one.

9.3.3 Analysis

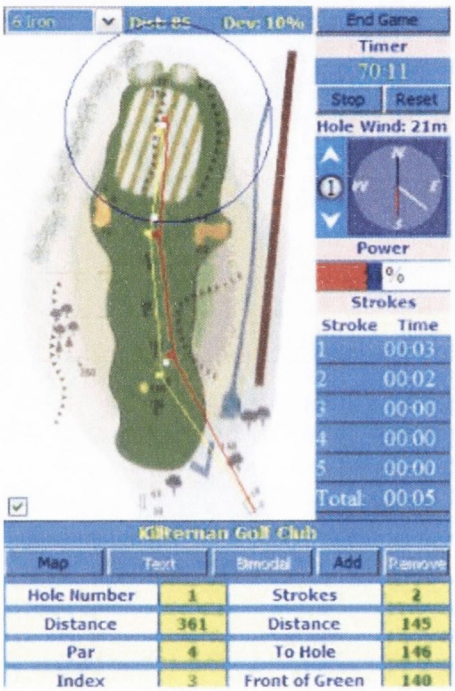


Figure 9.3.1: Example of the Map Mode interface for the Imager

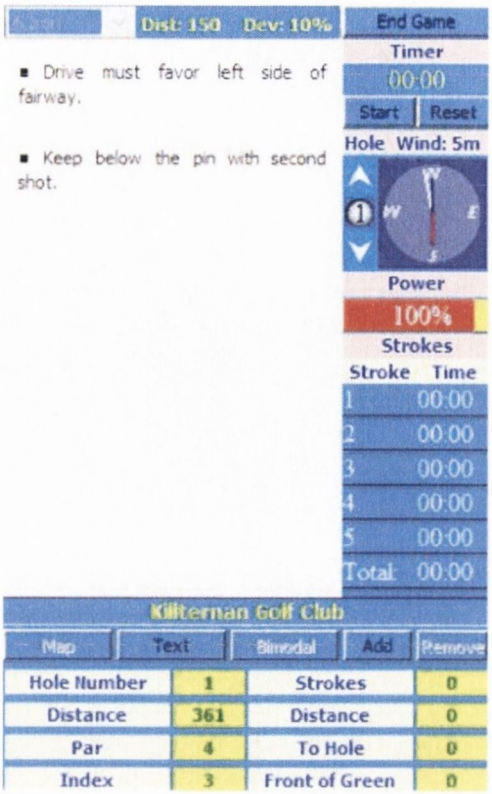


Fig 9.3.2. Sample textual interface for the Bimodal.

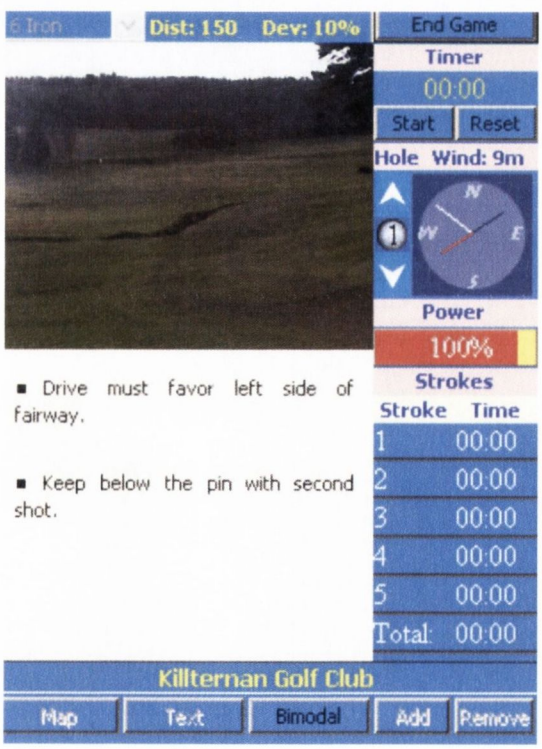


Fig 9.3.3 Sample Graphical interface for the Bimodal

Conclusions Relating to Golf Applications - Studies 9.4 & 9.5

Two golf stroke saver applications were designed: one was web based the other Multimedia Cd-Rom based. Each application presented three different views of each hole on two different golf courses.

Authorware was used to develop the multimedia Cd-Rom. The application was designed for Imagers, Bimodals and Verbalisers. The Golf Stroke Saver gave the golfer an overview of the each hole and the distance between different points of interest and the location of the pin. The application proved impractical to test in a meaningful way.

The purpose of the Golf applications was to investigate whether Golf Stroke Savers could be effectively designed to allow for the user's cognitive style. The golf based applications proved too difficult as platforms for testing the hypothesis. An examination of golfing performance as affected by the use of a stroke saver is somewhat insignificant as there are many more decisive factors involved. The testing of this application would also have required the use of the stroke saver in the field.

The use of the first application for the palmtop designed and evaluated in Study 9.4 also proved difficult to evaluate, as it too simulated a stroke saver but was not designed for use as an advanced virtual golf simulator. However this type of application demonstrated the potential for these types of commercial software applications

Appendix 21-9.4.D Web screenshots of the Route Navigator

This Appendix shows sample screenshots of the Route Navigator Interface.

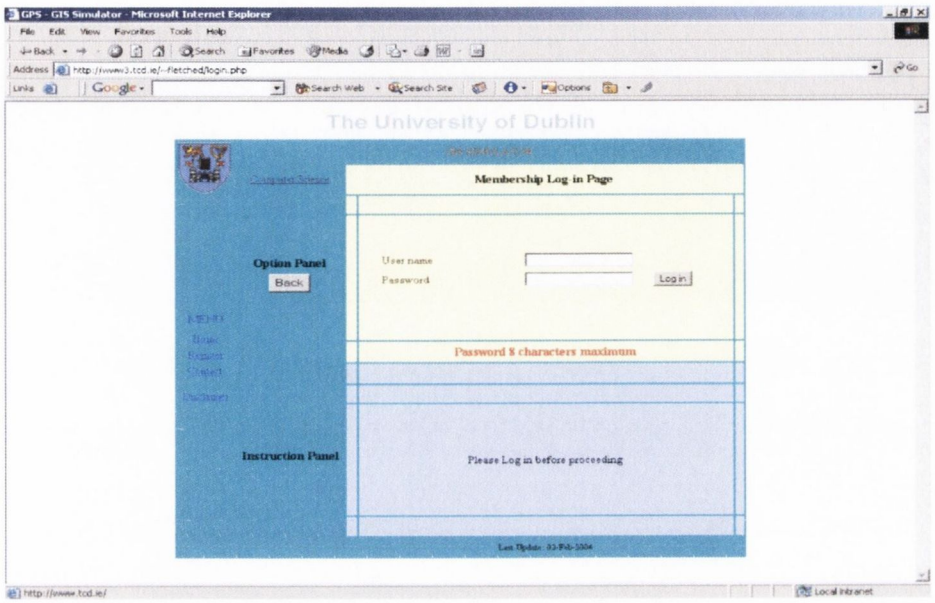


Fig. 9.4.1 Logon Screen for web-based Route Navigator

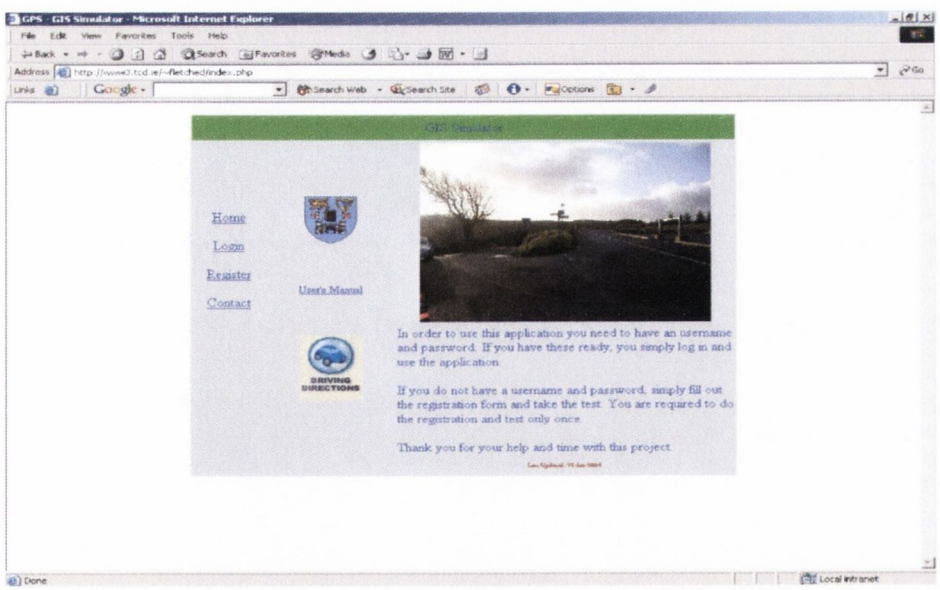


Fig. 9.4.2 Sample Screen for web-based Route Navigator

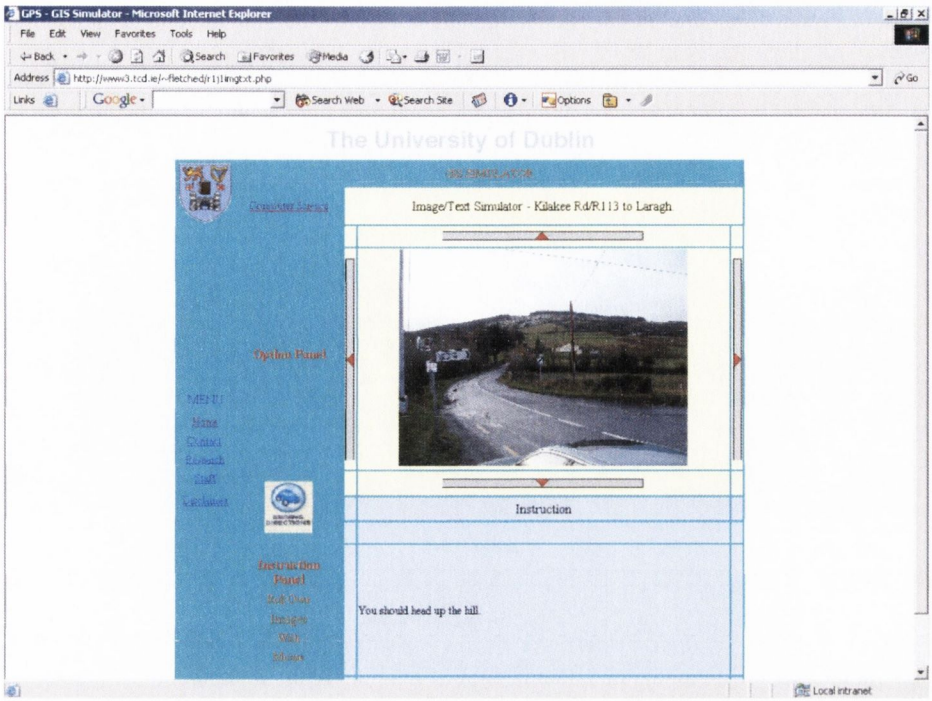


Fig. 9.4.3 Sample Screen for web-based Route Navigator

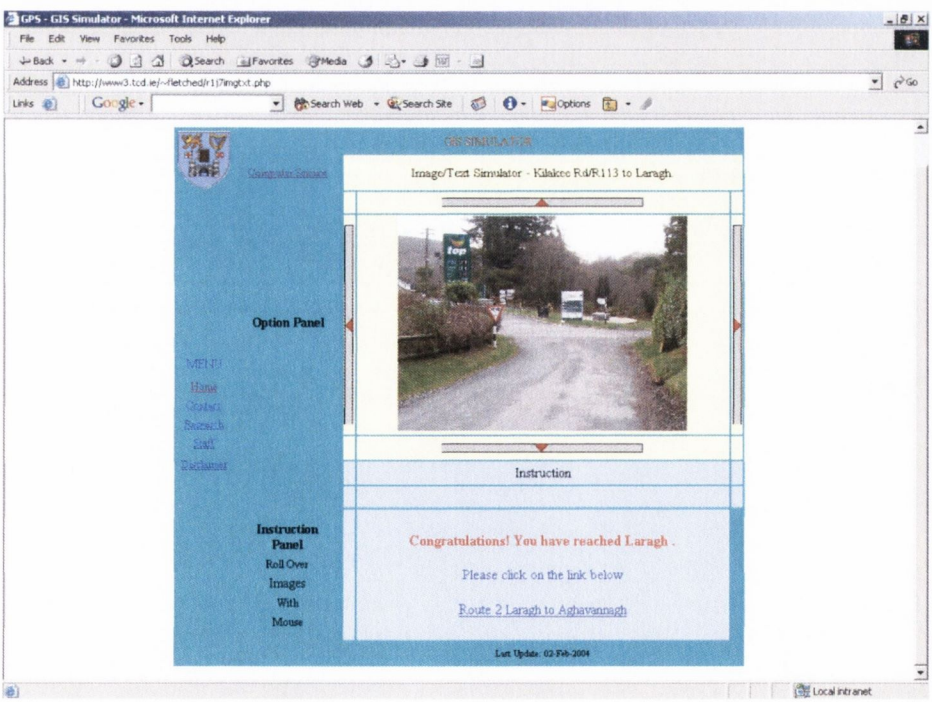


Fig. 9.4.4 Sample Screen for web-based Route Navigator

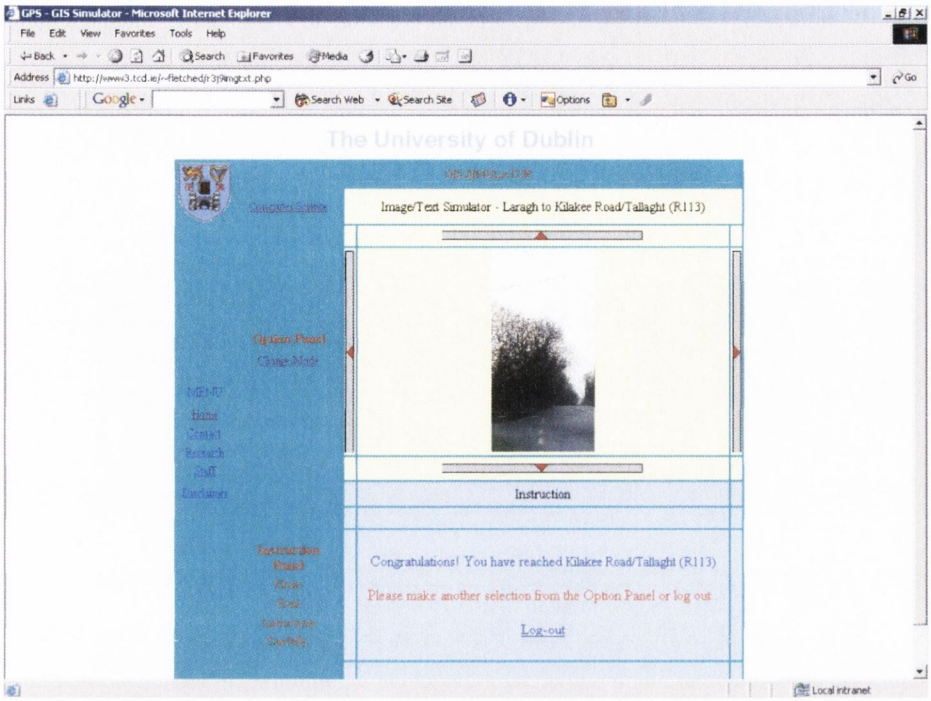


Fig. 9.4.5 Sample Screen for web-based Route Navigator

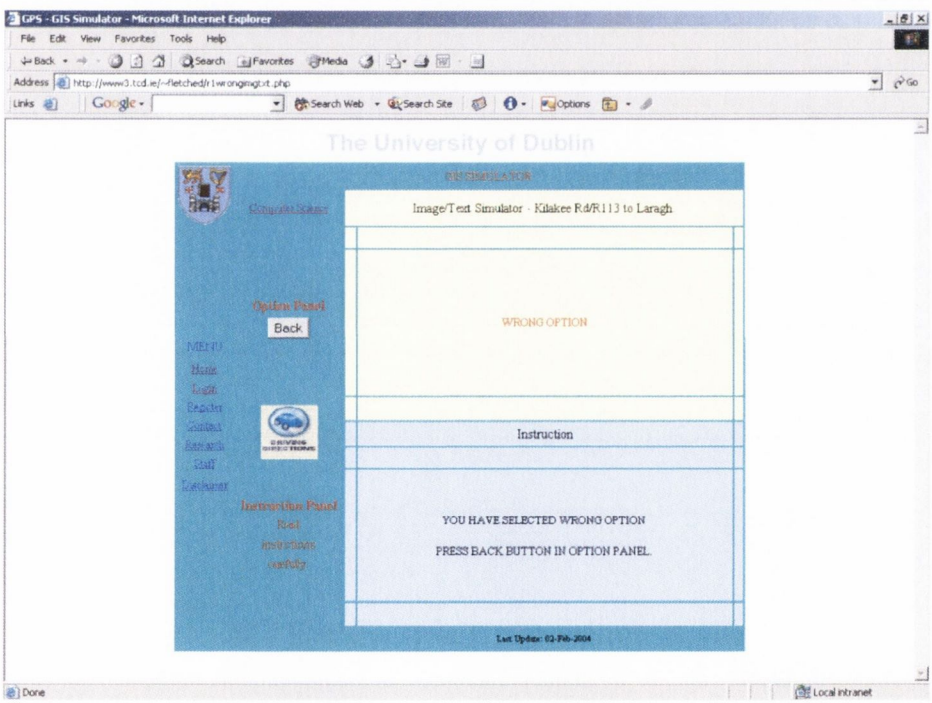


Fig. 9.4.6 Sample Screen for web-based Route Navigator

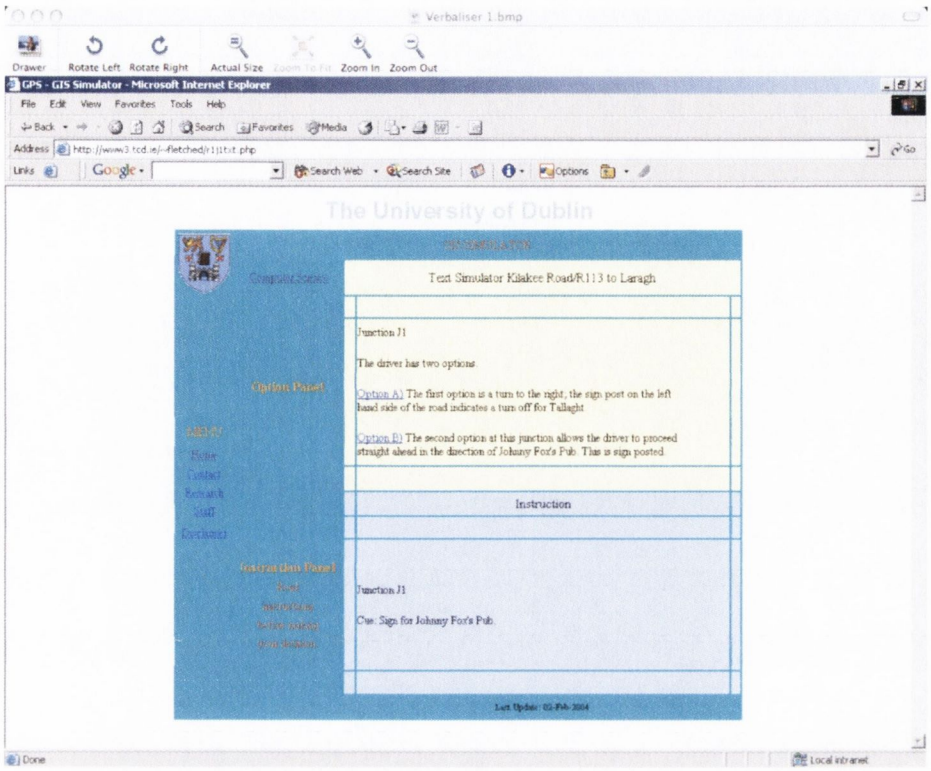


Fig. 9.4.7 Sample Screen for web-based Route Navigator

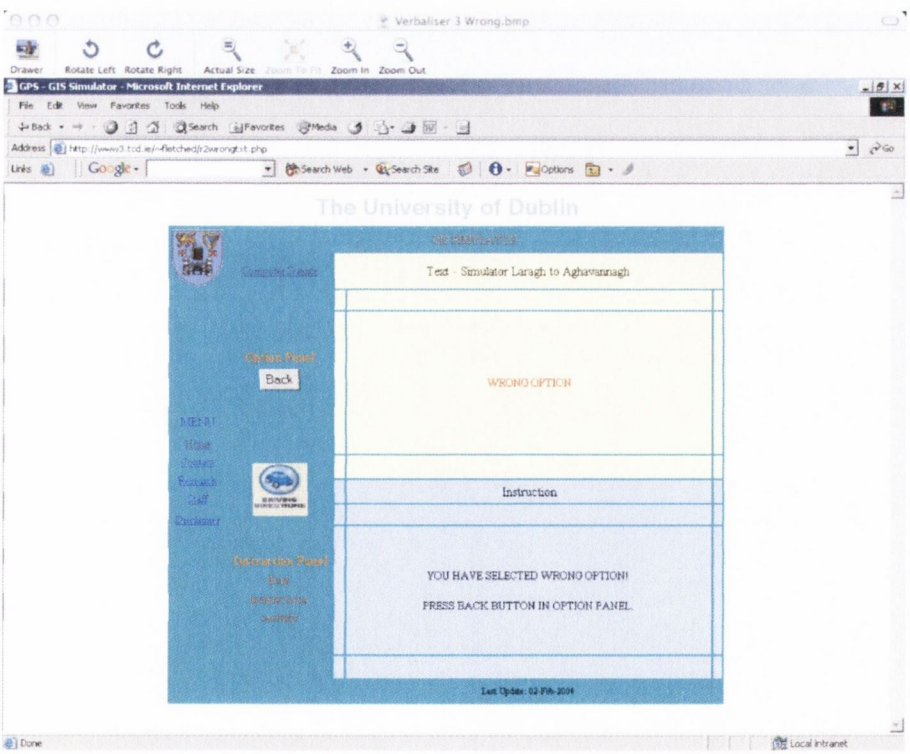


Fig. 9.4.8 Sample Screen for web-based Route Navigator

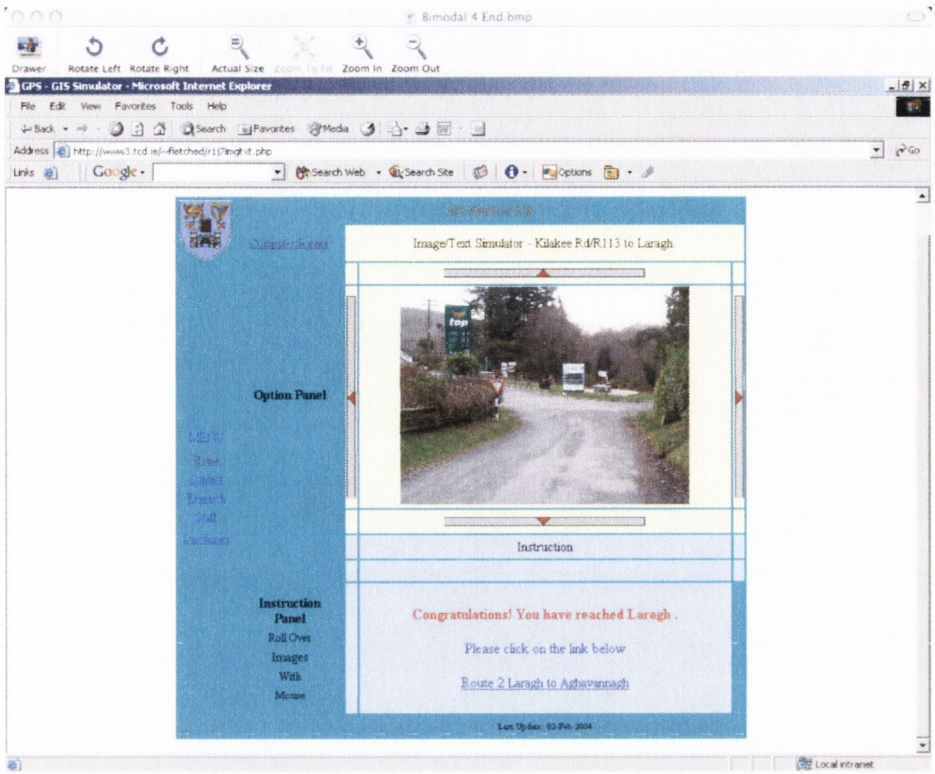


Fig. 9.4.9 Sample Screen for web-based Route Navigator

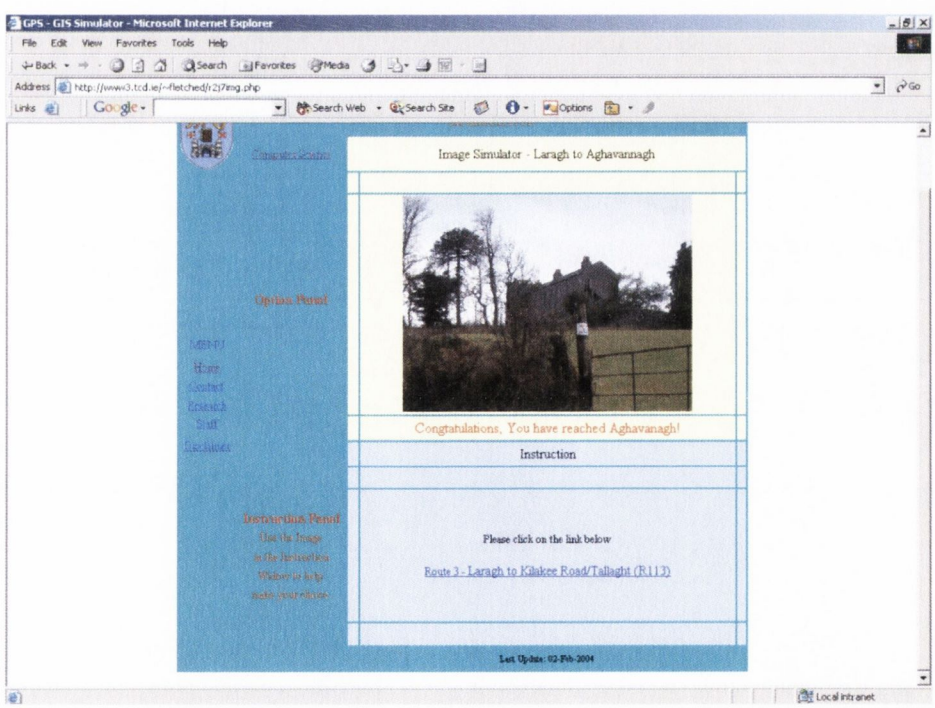


Fig. 9.4.9 Sample Screen for web-based Route Navigator

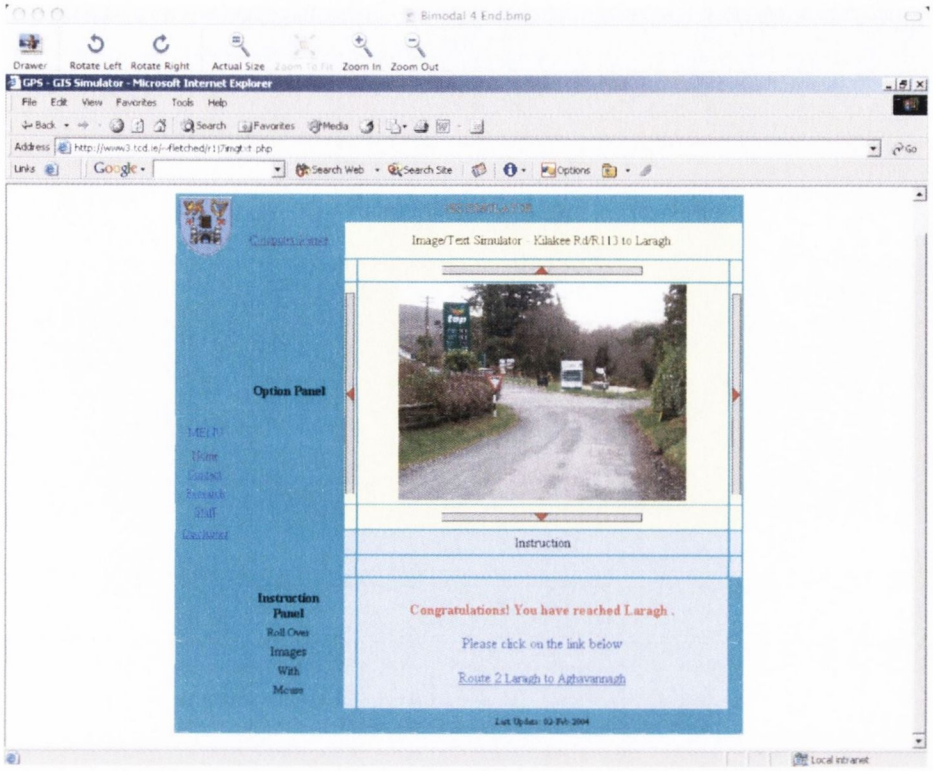


Fig. 9.4.10 Sample Screen for web-based Route Navigator

Appendix 22-9.5.E Quis

PART 3: Overall User Reactions

Please circle the numbers, which most appropriately reflect your impressions about using this computer system.

Not Applicable = NA.

3.1	Overall reactions to the system:	terrible	wonderful	
		1 2 3 4 5 6 7 8 9		NA
3.2		frustrating	satisfying	
		1 2 3 4 5 6 7 8 9		NA
3.3		dull	stimulating	
		1 2 3 4 5 6 7 8 9		NA
3.4		difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
3.5		inadequate power	adequate power	
		1 2 3 4 5 6 7 8 9		NA
3.6		rigid	flexible	
		1 2 3 4 5 6 7 8 9		NA

PART 4: Screen

4.1	Characters on the computer screen	hard to read	easy to read	
		1 2 3 4 5 6 7 8 9		NA
4.1.1	Image of characters	fuzzy	sharp	
		1 2 3 4 5 6 7 8 9		NA
4.1.2	Character shapes (fonts)	barely legible	very legible	
		1 2 3 4 5 6 7 8 9		NA
4.2	Highlighting on the screen	unhelpful	helpful	
		1 2 3 4 5 6 7 8 9		NA
4.2.1	Use of reverse video	unhelpful	helpful	
		1 2 3 4 5 6 7 8 9		NA
4.2.2	Use of blinking	unhelpful	helpful	
		1 2 3 4 5 6 7 8 9		NA
4.2.3	Use of bolding	unhelpful	helpful	
		1 2 3 4 5 6 7 8 9		NA

4.3	Screen layouts were helpful	never	always	
		1 2 3 4 5 6 7 8 9		NA
4.3.1	Amount of information that can be displayed on screen	inadequate	adequate	
		1 2 3 4 5 6 7 8 9		NA
4.3.2	Arrangement of information on screen	illogical	logical	
		1 2 3 4 5 6 7 8 9		NA
4.4	Sequence of screens	confusing	clear	
		1 2 3 4 5 6 7 8 9		NA
4.4.1	Next screen in a sequence	unpredictable	predictable	
		1 2 3 4 5 6 7 8 9		NA
4.4.2	Going back to the previous screen	impossible	easy	
		1 2 3 4 5 6 7 8 9		NA
4.4.3	Progression of work related tasks	confusing	clearly marked	
		1 2 3 4 5 6 7 8 9		NA

PART 6: Learning

6.1	Learning to operate the system	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
6.1.1	Getting started	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
6.1.2	Learning advanced features	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
6.1.3	Time to learn to use the system	slow	fast	
		1 2 3 4 5 6 7 8 9		NA
6.2	Exploration of features by trial and error	discouraging	encouraging	
		1 2 3 4 5 6 7 8 9		NA
6.2.1	Exploration of features	risky	safe	
		1 2 3 4 5 6 7 8 9		NA
6.2.2	Discovering new features	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
6.3	Remembering names and use of commands	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA
6.3.1	Remembering specific rules about entering commands	difficult	easy	
		1 2 3 4 5 6 7 8 9		NA

6.4 Tasks can be performed in a straight-forward manner — never always
1 2 3 4 5 6 7 8 9 NA

6.4.1 Number of steps per task — too many just right
1 2 3 4 5 6 7 8 9 NA

6.4.2 Steps to complete a task follow a logical sequence — never always
1 2 3 4 5 6 7 8 9 NA

6.4.3 Feedback on the completion of of steps — clear unclear
1 2 3 4 5 6 7 8 9 NA

PART 7: System Capabilities

7.1 System speed — too slow fast enough
1 2 3 4 5 6 7 8 9 NA

7.1.1 Response time for most operations — too slow fast enough
1 2 3 4 5 6 7 8 9 NA

7.1.2 Rate information is displayed — too slow fast enough
1 2 3 4 5 6 7 8 9 NA

7.2 The system is reliable — never always
1 2 3 4 5 6 7 8 9 NA

7.2.1 Operations are — undependable dependable
1 2 3 4 5 6 7 8 9 NA

7.2.2 System failures occur — frequently seldom
1 2 3 4 5 6 7 8 9 NA

7.2.3 System warns you about potential problems — never always
1 2 3 4 5 6 7 8 9 NA

7.3 System tends to be — quiet noisy
1 2 3 4 5 6 7 8 9 NA

7.3.1 Mechanical devices such as fans, disks, and printers — noisy quiet
1 2 3 4 5 6 7 8 9 NA

7.3.2 Computer generated sounds are — annoying pleasant
1 2 3 4 5 6 7 8 9 NA

7.4 Correcting your mistakes — difficult easy
1 2 3 4 5 6 7 8 9 NA

7.4.1 Correcting typos — complex simple

1 2 3 4 5 6 7 8 9 NA

7.4.2 Ability to undo operations — inadequate / adequate
1 2 3 4 5 6 7 8 9 NA

7.5 Ease of operation depends on your level of experience — never / always
1 2 3 4 5 6 7 8 9 NA

7.5.1 You can accomplish tasks knowing only a few commands — with difficulty / easily
1 2 3 4 5 6 7 8 9 NA

7.5.2 You can use features/shortcuts — with difficulty / easily
1 2 3 4 5 6 7 8 9 NA

PART 9: On-line Tutorials

9.1 Tutorial was — useless / helpful
1 2 3 4 5 6 7 8 9 NA

9.1.1 Accessing on-line tutorial — difficult / easy
1 2 3 4 5 6 7 8 9 NA

9.2 Maneuvering through the tutorial was — difficult / easy
1 2 3 4 5 6 7 8 9 NA

9.2.1 Tutorial is meaningfully structured — never / always
1 2 3 4 5 6 7 8 9 NA

9.2.2 The speed of presentation was — unacceptable / acceptable
1 2 3 4 5 6 7 8 9 NA

9.3 Tutorial content was — useless / helpful
1 2 3 4 5 6 7 8 9 NA

9.3.1 Information for specific aspects of the system were complete and informative — never / always
1 2 3 4 5 6 7 8 9 NA

9.3.2 Information was concise and to the point — never / always
1 2 3 4 5 6 7 8 9 NA

9.4 Tasks can be completed — with difficulty / easily
1 2 3 4 5 6 7 8 9 NA

9.4.1 Instructions given for completing tasks — confusing / clear
1 2 3 4 5 6 7 8 9 NA

9.4.2 Time given to perform tasks — inadequate / adequate
1 2 3 4 5 6 7 8 9 NA

9.5	Learning to operate the system using the tutorial was	difficult easy 1 2 3 4 5 6 7 8 9	NA
9.5.1	Completing system tasks after using only the tutorial	difficult easy 1 2 3 4 5 6 7 8 9	NA

PART 10: Multimedia

10.1	Quality of still pictures/photographs	bad good 1 2 3 4 5 6 7 8 9	NA
10.1.1	Pictures/Photos	fuzzy clear 1 2 3 4 5 6 7 8 9	NA
10.1.2	Picture/Photo brightness	dim bright 1 2 3 4 5 6 7 8 9	NA
10.2	Quality of movies	bad good 1 2 3 4 5 6 7 8 9	NA
10.2.1	Focus of movie images	fuzzy clear 1 2 3 4 5 6 7 8 9	NA
10.2.2	Brightness of movie images	dim bright 1 2 3 4 5 6 7 8 9	NA
10.2.3	Movie window size is adequate	never always 1 2 3 4 5 6 7 8 9	NA
10.3	Sound output	inaudible audible 1 2 3 4 5 6 7 8 9	NA
10.3.1	Sound output	choppy smooth 1 2 3 4 5 6 7 8 9	NA
10.3.2	Sound output	garbled clear 1 2 3 4 5 6 7 8 9	NA
10.4	Colors used are	unnatural natural 1 2 3 4 5 6 7 8 9	NA
10.4.1	Amount of colors available	inadequate adequate 1 2 3 4 5 6 7 8 9	NA

Summary of TVmed
For categories in Group

Appendix 23-9.6.F: Statistical Analyses for Studies 9.1, 9.2 and 9.3

Study 9.1 Experiment 1 Statistical Analysis Cd (n = 168)

Results of Study 9.1 Experiment 1

```
                Df  Sum Sq Mean Sq F value    Pr(>F)
factor(Group)   2 1595.04  797.52  109.96 < 2.2e-16 ***
Residuals     165 1196.74    7.25
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
  Welch Two Sample t-test
data:  J1S[Group == 1 | Group == 2] by Group[Group == 1 | Group ==
2]
t = -2.4719, df = 55.916, p-value = 0.01651
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:( -4.6103275 -0.4827262)
sample estimates:
mean in group 1 mean in group 2
     8.579996       11.126523
```

```
  Welch Two Sample t-test
data:  J1S[Group == 1 | Group == 3] by Group[Group == 1 | Group ==
3]
t = 6.3387, df = 25.971, p-value = 1.039e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (3.210651 6.292531)
sample estimates:
mean in group 1 mean in group 3
     8.579996       3.828405
```

```
  Welch Two Sample t-test
data:  J1S[Group == 2 | Group == 3] by Group[Group == 2 | Group ==
3]
t = 9.9034, df = 36.889, p-value = 6.173e-12
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (5.804798 8.791439)
sample estimates:
mean in group 2 mean in group 3
    11.126523       3.828405
```

These t tests check for significant differences between groups when the learning effect is compensated for by removing the first ten junctions from the analysis.

Results of Study 9.1 Experiment 2

Df Sum Sq Mean Sq F value Pr(>F)
factor(Group) 2 482.55 241.28 36.031 1.039e-13 ***
Residuals 165 1104.92 6.70

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Welch Two Sample t-test
data: J2S[Group == 1 | Group == 2] by Group[Group == 1 | Group == 2]
t = -0.0658, df = 33.774, p-value = 0.948
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-0.7662471 0.7181888)
sample estimates:
mean in group 1 mean in group 2
3.839948 3.863977

Welch Two Sample t-test
data: J2S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]
t = -7.9739, df = 67.469, p-value = 2.567e-11
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-4.439797 -2.662257)
sample estimates:
mean in group 1 mean in group 3
3.839948 7.390975

Welch Two Sample t-test
data: J2S[Group == 2 | Group == 3] by Group[Group == 2 | Group == 3]
t = -10.6303, df = 139.989, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-4.182962 -2.871033)
sample estimates:
mean in group 2 mean in group 3
3.863977 7.390975

Results of Study 9.1 Experiment 3

```
            Df Sum Sq Mean Sq F value    Pr(>F)
factor(Group)   2 636.04  318.02  76.454 < 2.2e-16 ***
Residuals     165 686.34    4.16
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
  Welch Two Sample t-test
data:  J3S[Group == 1 | Group == 2] by Group[Group == 1 | Group ==
2]
t = 11.9533, df = 35.265, p-value = 5.847e-14
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (3.787154 5.336220)
sample estimates:
mean in group 1 mean in group 2
      7.091644        2.529957
```

```
  Welch Two Sample t-test
data:  J3S[Group == 1 | Group == 3] by Group[Group == 1 | Group ==
3]
t = -0.6714, df = 46.994, p-value = 0.5053
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-1.1006348  0.5498247)
sample estimates:
mean in group 1 mean in group 3
      7.091644        7.367049
```

```
  Welch Two Sample t-test
data:  J3S[Group == 2 | Group == 3] by Group[Group == 2 | Group ==
3]
t = -17.298, df = 131.626, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-5.390248 -4.283936)
sample estimates:
mean in group 2 mean in group 3
      2.529957        7.367049
```

Results of Study 9.1 Experiment 4

```
            Df  Sum Sq Mean Sq F value    Pr(>F)
factor(Group)   2 117.931  58.965  61.578 < 2.2e-16 ***
Residuals     165 157.999   0.958
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
  Welch Two Sample t-test
data:  J4S[Group == 1 | Group == 2] by Group[Group == 1 | Group ==
2]
t = -6.2492, df = 55.939, p-value = 5.983e-08
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-2.574804 -1.324731)
sample estimates:
mean in group 1 mean in group 2
      3.013364        4.963131

  Welch Two Sample t-test
data:  J4S[Group == 1 | Group == 3] by Group[Group == 1 | Group ==
3]
t = 0.7109, df = 32.558, p-value = 0.4822
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-0.2524230  0.5233843)
sample estimates:
mean in group 1 mean in group 3
      3.013364        2.877883

  Welch Two Sample t-test
data:  J4S[Group == 2 | Group == 3] by Group[Group == 2 | Group ==
3]
t = 7.7983, df = 39.505, p-value = 1.617e-09
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (1.544603 2.625893)
sample estimates:
mean in group 2 mean in group 3
      4.963131        2.877883
n=168
```

Appendix 24-9.6.F CD Study 9.2 Experiment 1 Analysis (n = 40)

Study 9.2 Experiment 1 Results

```
                Df Sum Sq Mean Sq F value    Pr(>F)
factor(Group)  2 335.13  167.56  22.373 6.299e-07 ***
Residuals     34 254.64    7.49
```

```
  Welch Two Sample t-test
data:  J1S[Group == 1 | Group == 2] by Group[Group == 1 | Group ==
2]
t = -2.1723, df = 20.284, p-value = 0.04185
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-5.0965358 -0.1056380)
sample estimates:
mean in group 1 mean in group 2
       7.118242        9.719329
```

```
  Welch Two Sample t-test
data:  J1S[Group == 1 | Group == 3] by Group[Group == 1 | Group ==
3]
t = 7.0348, df = 16.783, p-value = 2.156e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (3.316633 6.162269)
sample estimates:
mean in group 1 mean in group 3
      7.118242        2.378791
```

```
  Welch Two Sample t-test
data:  J1S[Group == 2 | Group == 3] by Group[Group == 2 | Group ==
3]
t = 6.7437, df = 15.4, p-value = 5.755e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (5.025671 9.655404)
sample estimates:
mean in group 2 mean in group 3
      9.719329        2.378791
```

Appendix 25-9.6.F Study 9.2 Experiment 2 Results

```
             Df  Sum Sq Mean Sq F value Pr(>F)
factor(Group)  2   5.079   2.540  0.7412 0.4841
Residuals     34 116.494   3.426

  Welch Two Sample t-test
data:  J2S[Group == 1 | Group == 2] by Group[Group == 1 | Group ==
2]
t = 0.0082, df = 21.259, p-value = 0.9936
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: ( -0.9816664  0.9894021)
sample estimates:
mean in group 1 mean in group 2
      3.488725        3.484857

  Welch Two Sample t-test
data:  J2S[Group == 1 | Group == 3] by Group[Group == 1 | Group ==
3]
t = -0.8545, df = 13.722, p-value = 0.4075
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-2.841906  1.224828)
sample estimates:
mean in group 1 mean in group 3
              3.488725        4.297264

  Welch Two Sample t-test
data:  J2S[Group == 2 | Group == 3] by Group[Group == 2 | Group ==
3]
t = -0.8898, df = 12.177, p-value = 0.3908
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-2.798546  1.173733)
sample estimates:
mean in group 2 mean in group 3
      3.484857        4.297264
```

Appendix 26-9.6.F Study 9.2 Experiment 3 Analysis

```
             Df  Sum Sq Mean Sq F value    Pr(>F)

factor(Group)  2 133.004  66.502  8.5595 0.0009755 ***
Residuals     34 264.160   7.769

  Welch Two Sample t-test
data:  J3S[Group == 1 | Group == 2] by Group[Group == 1 | Group ==
2]
t = 6.1728, df = 13.618, p-value = 2.739e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (2.748916 5.688210)
sample estimates:
mean in group 1 mean in group 2
      6.834492        2.615929

  Welch Two Sample t-test
data:  J3S[Group == 1 | Group == 3] by Group[Group == 1 | Group ==
3]
t = 0.4982, df = 14.425, p-value = 0.6258
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: ( -2.453052  3.943017)
sample estimates:
mean in group 1 mean in group 3
      6.834492        6.089509

  Welch Two Sample t-test
data:  J3S[Group == 2 | Group == 3] by Group[Group == 2 | Group ==
3]
t = -2.5414, df = 10.549, p-value = 0.02821
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-6.4976583 -0.4495028)
sample estimates:
mean in group 2 mean in group 3
      2.615929        6.089509
```

Appendix 27-9.6.F Study 9.2 Experiment 4 Analysis

```
              Df  Sum Sq Mean Sq F value  Pr(>F)

factor(Group)  2  30.551  15.276   5.057 0.01195 *
Residuals     34 102.703   3.021

  Welch Two Sample t-test
data:  J4S[Group == 1 | Group == 2] by Group[Group == 1 | Group ==
2]
t = -2.5259, df = 16.245, p-value = 0.02229
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-3.2885462 -0.2893729)
sample estimates:
mean in group 1 mean in group 2
      3.220883        5.009843

  Welch Two Sample t-test
data:  J4S[Group == 1 | Group == 3] by Group[Group == 1 | Group ==
3]
t = 0.3858, df = 17.647, p-value = 0.7043
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-0.7428851  1.0765063)
sample estimates:
mean in group 1 mean in group 3
      3.220883        3.054073

  Welch Two Sample t-test
data:  J4S[Group == 2 | Group == 3] by Group[Group == 2 | Group ==
3]
t = 2.5812, df = 19.534, p-value = 0.01805
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (0.3728055 3.5387348)
sample estimates:
mean in group 2 mean in group 3
      5.009843        3.054073
```

Appendix 28-9.6.F Study 9.3 Web version of Route Simulator (n = 30)

Results of Study 9.3 Web version Experiment 1

Df Sum Sq Mean Sq F value Pr(>F)
factor(Group) 2 329.4 164.7 2.0601 0.1407
Residuals 40 3197.4 79.9

Welch Two Sample t-test
data: T1S[Group == 1 | Group == 2] by Group[Group == 1 | Group == 2]
t = -1.6526, df = 19.938, p-value = 0.1141
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-11.23873 1.30381)
sample estimates:
mean in group 1 mean in group 2
11.37143 16.33889

Welch Two Sample t-test
data: T1S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]
t = 0.4979, df = 22.973, p-value = 0.6233
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-2.678970 4.377383)
sample estimates:
mean in group 1 mean in group 3
11.37143 10.52222

Welch Two Sample t-test
data: T1S[Group == 2 | Group == 3] by Group[Group == 2 | Group == 3]
t = 1.809, df = 25.164, p-value = 0.08242
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: (-0.8034276 12.4367610)
sample estimates:
mean in group 2 mean in group 3
16.33889 10.52222

Results of Study 9.3 Web version Experiment 2

Df Sum Sq Mean Sq F value Pr(>F)
factor(Group) 2 141.77 70.88 4.2969 0.02040 *
Residuals 40 659.86 16.50

Welch Two Sample t-test

data: T2S[Group == 1 | Group == 2] by Group[Group == 1 | Group == 2]

t = -0.1049, df = 10.955, p-value = 0.9184

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: (-3.980423 3.618518)

sample estimates:

mean in group 1 mean in group 2

7.585714 7.766667

Welch Two Sample t-test

data: T2S[Group == 1 | Group == 3] by Group[Group == 1 | Group == 3]

t = -2.136, df = 12.162, p-value = 0.05368

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: (-7.68801847 0.07055816)

sample estimates:

mean in group 1 mean in group 3

7.585714 11.394444

Welch Two Sample t-test

data: T2S[Group == 2 | Group == 3] by Group[Group == 2 | Group == 3]

t = -2.6588, df = 33.603, p-value = 0.01192

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: (-6.4018382 -0.8537173)

sample estimates:

mean in group 2 mean in group 3

7.766667 11.394444

Results of Study 9.3 Web version Experiment 3

Df Sum Sq Mean Sq F value Pr(>F)

factor(Group) 2 100.90 50.45 3.6499 0.03499 *

Residuals 40 552.91 13.82

Welch Two Sample t-test

data: T3S[Group == 1 | Group == 2] by Group[Group == 1 | Group == 2]

t = 1.2994, df = 9.004, p-value = 0.2261

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: (-1.841881 6.814897)

sample estimates:

```
mean in group 1 mean in group 2
      8.014286       5.527778

  Welch Two Sample t-test
data:  T3S[Group == 1 | Group == 3] by Group[Group == 1 | Group
== 3]
t = -0.4137, df = 8.991, p-value = 0.6888
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: ( -5.118832  3.536293)
sample estimates:
mean in group 1 mean in group 3
      8.014286       8.805556

  Welch Two Sample t-test
data:  T3S[Group == 2 | Group == 3] by Group[Group == 2 | Group
== 3]
t = -2.7696, df = 34, p-value = 0.009027
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: ( -5.6829392 -0.8726164)
sample estimates:
mean in group 2 mean in group 3
      5.527778       8.805556
```

Appendix 29-9.7.G Quis: Statistical Data for each of the Questions

Table 9.7.1 Quis Question Scores and Statistics

Summaries
No Selector

Mean	Median	StdDev	Range	StdErr
4.66892	5	2.40008	9	0.197285
4.2027	5	2.61591	9	0.215026
3.73649	4	2.48905	9	0.204599
5.70946	6.5	3.01076	9	0.247483
4.66216	5	2.89629	9	0.238073
3.4527	3	2.57496	9	0.21166
6.25676	7	3.07972	9	0.253151
4.94595	6	2.93879	9	0.241567
5.92568	7	2.96485	9	0.243709
4.52703	5	3.07159	9	0.252483
2.36486	0	2.78027	9	0.228537
2.33784	0	2.71943	9	0.223536
3.76351	5	3.23325	9	0.265771
5.01351	6	2.67003	9	0.219475
5.64189	7	2.84045	9	0.233484
5.48649	7	2.85592	9	0.234755
5.62162	7	2.839	9	0.233365
5.07432	6	2.72327	9	0.223852
2.5473	1.5	2.66582	9	0.219129
4.81081	6	2.99285	9	0.246011
6.45946	8	3.0869	9	0.253742
6.12162	7	3.019	9	0.24816
4.43243	6	3.68885	9	0.303222
6.40541	7	2.98713	9	0.245541
4.43243	5	3.30566	9	0.271723
3.91216	4	3.33795	9	0.274378
3.45946	3	3.33886	9	0.274453
5.31757	7	3.40875	9	0.280197
5.56757	7	3.43088	9	0.282016
6.34459	7	2.92238	9	0.240218
5.18919	6	2.9769	9	0.2447
5.84459	7	2.98913	9	0.245705
3.41892	3	2.74299	9	0.225472
3.94595	5	3.45581	9	0.284066
3.62162	4	3.74058	9	0.307473
4.93919	7	3.64818	9	0.299878
4.38514	5	3.43159	9	0.282075
4.87838	7	3.62888	9	0.298292
3.91216	5	3.5069	9	0.288265
4.11486	5	3.27124	9	0.268894
4.43243	6	3.36885	9	0.276918
5.04054	7	3.61945	9	0.297517
4.91892	6	3.58666	9	0.294822
5.11486	7	3.91495	9	0.321807
4.90541	6	3.64279	9	0.299435
4.40541	5	3.62313	9	0.297819
4.06757	4	2.6021	9	0.213891
3.89865	4	2.59445	9	0.213262
4.67568	5	2.51043	9	0.206356
1.4527	0	2.68616	9	0.220801
1.37162	0	2.53219	9	0.208145
1.44595	0	2.60111	9	0.21381
1.58108	0	2.77014	9	0.227704
1.71622	0	2.95785	9	0.243133
1.62162	0	2.81494	9	0.231387
1.70946	0	2.90731	9	0.238979
4.33784	5	3.21889	9	0.264591
4.43243	5	3.2788	9	0.269515